Human small nuclear RNA gene transcription by RNA polymerases II and III depends upon promoter recognition by the SNAPC general transcription factor. DNA binding by SNAPC involves direct DNA contacts by the SNAP190 subunit in cooperation with SNAP50 and SNAP43. The data presented herein shows that SNAP50 plays an important role in DNA binding by SNAPC through its zinc finger domain. The SNAP50 zinc finger domain contains 15 cysteine and histidine residues configured in two potential zinc coordination arrangements. Individual alanine substitution of each cysteine and histidine residue demonstrated that eight sites are important for DNA binding by SNAPC. However, metal binding studies revealed that SNAPC contains a single zinc atom indicating that only one coordination site functions as a zinc finger. Of the eight residues critical for DNA binding, four cysteine residues were also essential for both U1 and U6 transcription by RNA polymerase II and III, respectively. Surprisingly, the remaining four residues, although critical for U1 transcription could support partial U6 transcription. DNA binding studies showed that defects in DNA binding by SNAPC alone could be suppressed through cooperative DNA binding with another member of the RNA polymerase III general transcription machinery, TFIIIB. These results suggest that these eight cysteine and histidine residues perform different functions during DNA binding with those residues involved in zinc coordination likely performing a dominant role in domain stabilization and the others involved in DNA binding. These data further define the unorthodox SNAP50 zinc finger region as an evolutionarily conserved DNA binding domain.

The human snRNA$^3$ gene family is unusual in that related member genes are transcribed by either RNA polymerase II or III depending upon their core promoter structures, and they thus serve as an interesting model to understand principles of polymerase selection and activity both during normal and deregulated growth (reviewed in Refs. 1 and 2). Human snRNA genes are defined by the presence of a diagnostic promoter element called the proximal sequence element (PSE). For both polymerase systems, the PSE recruits the general transcription factor called SNAPC$^3$, which is also known as PTF (4).

SNAPC$^3$ is composed of at least five subunits SNAP190, SNAP50, SNAP45, SNAP43, and SNAP19 (5–9). SNAP190 can interact with all the other subunits, and provides a central architectural backbone to the complex (10). SNAP190 directly interacts with the transcriptional activator protein Oct-1 during stimulated transcription of human snRNA genes (11, 12). In the RNA polymerase III pathway, SNAP190 also interacts with TBP to recruit Brf2-TFIIIB to the TATA box of human U6 snRNA genes (13, 14). Thus, the upstream signal from Oct-1 is conveyed through SNAP190 stimulating preinitiation complex assembly with other general transcription factors as a prerequisite for RNA polymerase III recruitment. Interestingly, protein kinase CK2 can phosphorylate SNAP190 to impede DNA binding; however, promoter recognition by SNAPC$^3$ can be restored by cooperative binding of TBP to those promoter sequences containing both a PSE and TATA box, but not sequences lacking the TATA element (15). This last observation suggests that CK2 can differentially influence RNA polymerase II and III snRNA transcription by covalent modification of SNAPC$^3$ to alter its DNA binding properties even though SNAPC$^3$ is shared between both polymerase systems.

The specific binding of SNAPC$^3$ to the PSE is mediated by a Myb-like DNA binding domain within the NH$_2$-terminal region of SNAP190. Other SNAPC$^3$ subunits may additionally function in PSE recognition and provide stabilizing contacts with DNA. Indeed, UV cross-linking experiments suggest that both SNAP190 (16) and SNAP50 (6) are in close contact with DNA. In addition, SNAP50 contains an unusual arrangement of cysteine and histidine residues at the carboxyl-terminal region of the protein that is evolutionarily conserved. In other transcription factors, zinc finger motifs function for nucleic acid binding and/or protein-protein interactions (17). In the current study, we have focused on the function for the zinc finger region of SNAP50, and we show that this region is critical for DNA binding by SNAPC$^3$. The arrangement of cysteine and histidine residues within SNAP50 further defines this region as an unorthodox zinc finger domain that functions in divergent preinitiation complex assembly pathways for RNA polymerase II and III transcription.

---

* This work was supported by National Institutes of Health Grants R01-GM59805 (to R. W. H.) and R01-GM063894 (to J. H. G.). 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 may be addressed. Tel.: 517-353-9715 (ext. 234).

‡ To whom correspondence may be addressed. Tel.: 517-353-9715; Fax: 517-353-9334; E-mail: henryrw@msu.edu.

§ The abbreviations used are: snRNA, small nuclear RNA; PSE, proximal sequence element; EMSA, electrophoretic mobility shift assays; TBP, TATA binding protein; TF, transcription factor; GST, glutathione S-transferase; HA, hemagglutinin; ICP-MS, inductively coupled plasma-mass spectrometry.
DNA Binding by the SNAP50 Zinc Finger Domain

Expression and Purification of Recombinant Proteins—For the EMSA shown in Fig. 1A, mini-SNAP_C was assembled from subunits individually expressed in Escherichia coli and purified as described before (14). The partial complex SNAP_C-Y4 was obtained by co-expression in E. coli of SNAP190-(1–505), SNAP43, SNAP19, and either wild-type or mutant SNAP50. The complex was purified as described (18). TBP was expressed as a GST fusion protein in E. coli and purified as described (14). The Brf2 and Bdp1-(1–470) proteins were expressed in E. coli and purified as described before (19). In all cases, GST tags were removed prior to use in functional assays by thrombin digestion during purification.

Electrophoretic Mobility Shift Assay—EMSA was performed in a 20-μl total volume using DNA probes containing a wild-type or mutant mouse U6 PSE with a mutant human U6 TATA box, or containing the wild-type human U6 sequence, as described previously (14). DNA binding reactions using only SNAP_C were performed in a buffer containing 60 mM KCl, 20 mM HEPES, pH 7.9, 5 mM MgCl2, 0.2 mM EDTA, 10% glycerol, 0.5 μg of poly(dI-dC), and 0.5 μg of pUC119 plasmid. Reactions were incubated for 20 min at room temperature after which 5,000 cpm of probe was added, and reactions were incubated an additional 20 min. Samples were fractionated on a 5% nondenaturing polyacrylamide gel (39:1) in TGE running buffer (50 mM Tris, 380 mM glycine, 2 mM EDTA). Reactions containing both SNAP_C-Y4 and TBP were performed in a buffer containing 100 mM KCl, 20 mM HEPES, pH 7.9, 5 mM MgCl2, 0.2 mM EDTA, 10% glycerol, 1 mM dithiothreitol, 0.07% Tween 20, 0.2 μg of poly(dG-dC), and 0.2 μg of pUC119 plasmid. The samples were fractionated on a 5% nondenaturing polyacrylamide gel (39:1) in TGE running buffer (50 mM Tris, 380 mM glycine, 2 mM EDTA, 5 mM MgCl2). Approximately 1 ng of SNAP_C-Y4, 50 ng of human TBP, 240 ng of Brf2, and 40 ng of Bdp1-(1–470) was used, as indicated.

GST Pulldown and Immunoprecipitation Assays—GST pull-down assays were done as described before (14). For the immunoprecipitation assays performed in Fig. 2B, ~20 ng of each complex containing SNAP50 with the indicated point mutation was used with 1 μl of α-SNAP43 antibody, preimmune sera, or buffer alone. Immunoprecipitation reactions were carried out in HEMGT-150 buffer (25 mM HEPES, 0.1 mM EDTA, 12.5 mM MgCl2, 10% glycerol, 0.1% Tween 20, 150 mM KCl) containing protease inhibitors and 1 mM dithiothreitol at 4 °C for 2 h. The immunoprecipitated material was collected by incubation with Protein G-agarose beads at 4 °C for 2 h. Beads were then collected, washed 5 times with HEMGT-150, and the bound protein was released by boiling in Laemmli buffer. Recovered proteins were resolved by 15% SDS-PAGE and transferred to a nitrocellulose membrane for Western blotting using anti-HA antibodies.

In Vitro Transcription Assays—Endogenous SNAP_C was removed from HeLa extracts by anti-SNAP43 antibody immunodepletion as described before (8). In vitro transcription of human U1 and U6 snRNA genes was performed for 1 h at 30 °C using the depleted extract. Approximately 5 ng of wild-type and mutant SNAP_C-Y4 complexes containing point mutations in SNAP50 were used for the reconstitution of both U1 and U6 transcription.

Zinc Binding Studies—The amount of zinc associated with recombinant SNAP_C was determined by two methods: 1) inductively coupled plasma mass spectrometry (ICP-MS) and 2) flame atomic absorption. For the ICP-MS studies, SNAP_C protein concentration was determined by Bradford assay using bovine serum albumin as a standard. The molecular weights of recombinant SNAP-C-Y4 and SNAP-C-Y3 were calculated as 159,735 and 148,432 g/mol, respectively. To measure zinc concentration, sample solutions were transferred to a Teflon vial and brought to dryness on a hot plate. Concentrated nitric acid was added and the sample was placed on the hot plate and hydrolyzed for 30 min. After cooling, the sample was diluted with water to bring the acid concentration to 2%, and 2% nitric acid was added to bring the solution to the desired volume. Nickel and zinc standards (Spex Certiprep) in the concentration range from 0 to 1000 ppb were prepared in 2% nitric acid. Samples and standards were each mixed with 20 ppb indium and bismuth standard solution (Spex Certiprep) as internal standards. Samples were analyzed on an ICP-MS instrument (GV Instruments) with a flow rate of 0.5 ml/min. 66Zn, 68Zn, and 60Ni isotopes were measured and quantified. The responses for zinc and nickel were corrected according to indium and bismuth response. For the flame atomic absorption studies, SNAP_C protein concentration was determined by UV absorbance in 6 M urea at 280 nm. The molar extinction coefficient of 176,950 cm⁻¹ M⁻¹ was used to calculate the molar concentration. SNAP_C samples were transferred to a crucible, dried, and ashed at 260 °C. Concentrated nitric acid was added to the ash and reheated until only white powder remained. The sample was reconstituted with 5% nitric acid. Zinc standards in the concentration range from 10 to 1000 ppb were prepared from zinc metal (Spectrum, 99.9%) dissolved in concentrated nitric acid. The samples and standards were analyzed on the flame atomic absorption instrument (Varian SpectraAA-200) equipped with a zinc hollow cathode lamp operating at 213.9 nm. Aspiration rate was 1 ml/min.

For the zinc analysis shown in Fig. 6, wild-type or mutant SNAP50-(315–411) was expressed as a NH2-terminal His-SUMO fusion protein from pET28a in E. coli. Cells were grown in LB broth containing kanamycin (100 μg/ml) and chloramphenicol (50 μg/ml) supplemented with 50 μM ZnCl2. Approximately 1 liter of culture was used for the expression and purification of each protein. Cells were ruptured by sonication in buffer A (50 mM Tris, pH 8.5, 350 mM NaCl, 10% glycerol, 10 mM imidazole, 1 mM β-mercaptoethanol) supplemented with 10 μM ZnCl2. His-tagged recombinant proteins were purified by nickel column chromatography followed by elution with buffer A containing 400 mM imidazole without added zinc. Protein containing fractions were dialyzed against buffer A containing 1 mM dithiothreitol and no added zinc for 10 h at 4 °C for subsequent flame atomic absorption analysis, which was performed as above using ~2 ml of each protein solution (~1 mg/ml) with an aspiration rate of 5 ml/min. Protein concentration was determined in 6 M urea by UV absorbance at 280 nm. The molar extinction coefficient of 27,305 cm⁻¹ M⁻¹ was used to calculate the molar concentration.
of the wild type and mutant His-SUMO-SNAP50-(315–411) proteins. Zinc measurements and protein determination were done in triplicate.

**Structural Modeling**—Two protocols were followed to obtain *ab initio* predictions of SNAP50-(301–411). First, the amino acid sequence of SNAP50 was submitted to the Robetta server (21), and 10 models were obtained. Second, lattice-based sampling from extended chains with MONSSTER (22) and the MMTSB Tool Set (23) were carried out. 2000 structures were generated in independent runs. The resulting structures were subjected to a short minimization with CHARMM (24) and evaluated with the scoring function DFIRE (25). Correlation-based scores were obtained from the original DFIRE scores according to a recently published method for enhancing scoring functions in protein structure prediction applications (26). The structures were then clustered and average correlation-based scores were compared between clusters. The structure with the highest correlation-based score from the cluster with the highest average correlation-based score and more than one member was then examined and subjected to further refinement through energy minimization and constrained short molecular dynamic simulations with CHARMM. The electrostatic potential on the surface of the final structure was calculated from solutions of the Poisson equation with the PBEQ module (27) in CHARMM. The program VMD was used for visualization of the final model and electrostatic surface maps (28).

**RESULTS**

The COOH-terminal Region of SNAP50 Is Required for DNA Binding by SNAPC—Promoter recognition by SNAPC is a common early event in both RNA polymerase II- and III-specific pathways of pre-initiation complex formation at human snRNA genes. Whereas the mechanism for PSE binding by SNAPC is not well understood, this initial event requires extensive cooperation among SNAP43, SNAP50, and SNAP190, as indicated by the data in Fig. 1A. EMSA reactions containing all three subunits supported robust binding by mini-SNAPC to wild-type (lane 7) but not mutant PSE probes (lane 8), whereas no detectable DNA binding was observed in reactions that lacked all three complex members (lanes 1–6), consistent with results previously described (14). Thus, even though SNAP190 contains a bona fide Myb DNA binding domain, additional components of the complex are required, perhaps serving to directly recognize the PSE alongside SNAP190.

Whereas the previous experiment was performed with three factors, we postulate that the recognition of DNA by SNAPC likely involves direct DNA contacts provided by SNAP50 in addition to the contributions made by SNAP190. First, SNAP50 was UV cross-linked to the PSE during DNA binding by endogenous SNAPc (6), indicating that SNAP50 is in close proximity to DNA during promoter recognition. Second, SNAP50 contains a putative zinc finger domain within the COOH-terminal region, and as zinc finger domains are typically involved in nucleic acid binding and/or protein-protein interactions, we hypothesized that this region of SNAP50 is involved in DNA binding by SNAPC. To test this idea, EMSA were performed using mini-SNAPC containing SNAP50 that lacked the COOH-terminal cysteine/histidine-rich region. In support of the hypothesis, DNA binding by mini-SNAPC was completely ablated in reactions performed with SNAP50-(1–300) lacking the cysteine/histidine-rich region (lanes 15–17), as compared with comparable amounts of mini-SNAPC containing wild-type SNAP50 (lanes 12–14). In these reactions, the assembly of
SNAP50-(1–300) into a complex along with SNAP43 and SNAP190-(1–505) was as efficient as that observed for wild-type SNAP50 (data not shown), suggesting that those protein-protein interactions required for complex formation are not seriously jeopardized by removal of the cysteine/histidine-rich region. Indeed, strong pairwise interactions between SNAP50 and SNAP43, its major partner in SNAPC4, were maintained for truncated SNAP50 molecules lacking this region (Fig. 1B). In GST pulldown experiments, GST-SNAP43 interacted well with full-length SNAP50-(1–411), SNAP50-(1–300), and SNAP50-(1–199), but not with SNAP50-(1–124) nor with the cysteine/histidine-rich region alone (SNAP50–301-411). These data support the idea that the central region of SNAP50 participates in complex assembly with SNAP43, whereas the COOH-terminal cysteine/histidine-rich region constitutes a DNA binding domain within SNAP50.

The requirement for the SNAP50 cysteine/histidine-rich region for DNA binding by SNAPC4 suggests that this region of SNAP50 constitutes a zinc finger domain. To determine whether zinc is indeed bound by SNAPC4, the ratio of zinc associated with a recombinant SNAPC4-containing HA epitope-tagged SNAP50 along with SNAP190-(1–505), SNAP43, and SNAP19 (hereafter referred to as SNAPCγ4) was determined by ICP-MS. SNAPCγ4 was chosen for this study because this recombinant complex, assembled by co-expression of each subunit in E. coli, is fully functional for PSE-specific DNA binding and for snRNA gene transcription by both RNA polymerases II and III (18), and zinc binding under functional conditions could be examined. As shown in Table 1, ~1.2 mol of zinc, as measured for both 66Zn and 68Zn, were found associated with each mole of SNAPCγ4, whereas little detectable nickel was associated with SNAPCγ4. Similar data were obtained for two separate SNAPCγ4 preparations. These results are comparable with those obtained by atomic absorption indicating a molar ratio (number of zinc/SNAPCγ4) of 0.9 for the four-subunit SNAPCγ4 complex, and for the three-subunit SNAPCγ3 complex lacking SNAP19. As no other SNAPC subunits besides SNAP50 contain suitable arrangements of cysteines and/or histidines for zinc coordination, the parsimonious explanation of these data is that SNAPCγ4 contains a single SNAP50 subunit that binds a single zinc atom.

Table 1

| Sample | [Zn] | [Ni] | [Protein] | No. (Zn)/No. (Ni) |
|--------|------|------|------------|-----------------|
| SNAPCγ4 (1) | 1250.4a | 2.6 | 1.2 | |
| SNAPCγ4 (2) | 1415.5a | 2.9 | 1.2 | 0.02 |
| SNAPCγ3 | 1443.6b | 2.9 | 1.2 | 0.03 |

* Total zinc calculated from Zn56; MW: SNAPCγ4 = 159,735 g/mol.
* Total zinc calculated from Zn58; MW: SNAPCγ4 = 148,432 g/mol.

An Evolutionarily Conserved Zinc Finger Domain in Human SNAP50 Functions for DNA Binding—The amino acid sequence for the human SNAP50 zinc finger domain is shown in Fig. 2A, which reveals an unorthodox arrangement of cysteines and histidines within the COOH-terminal 110 amino acids. Typically, zinc finger domains that function for DNA binding are constituted by multiple repeats of related zinc finger motifs (17). However, the putative DNA binding domain of human SNAP50 contains nine cysteines and six histidines that can be grouped into region 1 loosely resembling a TFIIIA-like C2H2 zinc finger and region 2 resembling a glucocorticoid-like C2C2 zinc finger. Whereas both types of motifs in other proteins are involved in direct DNA contacts, the discordant arrangement of zinc finger motifs suggests that DNA binding by SNAP50 is distinct from mechanisms employed by other zinc finger proteins. In addition, as SNAPCγ4 contains only a single zinc atom, it is likely that only one of the potential zinc fingers actually coordinates zinc. Interestingly, the results of a BLAST homology search show that of the 15 potential zinc coordination sites in this region, eight are highly conserved among putative SNAP50 homologues from mammals, insects (Drosophila and Anopheles), fish (Danio), worms (Caenorhabditis elegans), plants (Arabidopsis), slime mold (Dictyostelium), and parasites from Trypanosoma, Entamoeba, Plasmodium, and Leishmania genera. This comparison also shows that the critical zinc coordination sites within region 2 that constitute the glucocorticoid-like finger motif are well conserved, whereas the TFIIIA-like C2H2 zinc finger motif within region 1 is not. Nonetheless, two histidines and one cysteine residue within region 1 are conserved raising the possibility that a function performed by this region is also maintained across divergent species. The consensus motif derived from this comparison (L3C20–23YPX11–17C2C3–4CFX6–8H1–G) is distinct from any other family of zinc finger motifs (29), suggesting that this domain represents a novel zinc finger fold, and is hereafter referred to as the “SNAP finger” domain.

To examine the function of the cysteine and histidine residues within the COOH-terminal SNAP finger domain of SNAP50, each of these residues was changed to alanine for functional testing in the context of a partial SNAPC that contains SNAP190-(1–505), SNAP43, SNAP19, and the various derivatives of full-length HA-SNAP50. In addition, an arginine at position 385 was also substituted with alanine. First, to determine whether the targeted amino acids are critical for complex assembly, a two-step affinity purification of wild-type and mutant complexes were performed. Complexes were assembled by co-expressing four SNAPC subunits in E. coli followed by affinity purification of the complex via the GST tag contained on the amino terminus of SNAP190-(1–505). Complexes were liberated from the glutathione-agarose beads by thrombin digestion, which recognizes its cognate site between the GST tag and the SNAP190 coding region. The soluble complexes were further purified by anti-SNAP43 immunoprecipitation, and the extent of SNAP50 association, as a measure of complex integrity, was determined by Western blot analysis against the HA tag contained at the NH2 terminus of SNAP50.
FIGURE 2. Single amino acid changes in the SNAP50 subunit abolish DNA binding by SNAPC.

A, sequence alignment of human SNAP50 COOH-terminal amino acids (301–411) with corresponding regions from SNAP50 homologues of other species. Putative zinc fingers similar to TFIIIA (His-X5-His-X10-His-X3-Cys) and steroid receptors (Cys-X2-Cys-X22-Cys-X2-Cys) are indicated as regions 1 and 2, respectively. The sequence of highly conserved amino acids derived from this alignment corresponds to Lx4Gx6Hx3CxHx20–23YPx11–12Cx2Cx18Px3–4Cx2CFx3Hx1–4G. This alignment was performed using the Clustal W program.

Homo sapiens (NP003075), Canis familiaris (XP653813), Bos taurus (AAO08912), Mus musculus (NP084225), Rattus norvegicus (NP001013230), Drosophila melanogaster (NP724647), Drosophila pseudoobscura (EAL25490), Typhimurium brucel (XP827395), Arabidopsis thaliana (AAD30067), C. elegans 1 (NP500819), C. elegans 2 (NP497807), Plasmodium falciparum (NP702469), Danio rerio (XP694501), Leishmania major (XP835739), Anopheles gambiae (XP310411), Dictyostelium discoideum (XP644064), Entamoeba histolytica (XP653151). B, mutations in the SNAP50 zinc finger domain do not disrupt SNAPC assembly. Approximately 20 ng of each of the SNAPC-y4 complex containing substitution mutations in HA-SNAP50 were affinity purified first using glutathione-agarose to pull down GST-SNAP190 (1–505) followed by immunoprecipitation with SNAP43 antibodies. Associated wild-type or mutant HA-SNAP50 was detected by Western analysis (lanes 6–22). Lanes 4 and 5 contain wild-type SNAPC-y4 recovered with the protein G-agarose beads alone or with preimmune serum, respectively. The bottom panel represents 4% of the input material directly analyzed by Western analysis.

C, mutations in the SNAP50 zinc finger domain disrupt DNA binding by SNAPC. Increasing amounts (3 and 10 ng) of SNAPC-y4 with wild-type (lanes 2 and 3) or mutant HA-SNAP50 (lanes 4–35) containing the indicated substitution mutations was tested in an EMSA for binding to a radiolabeled DNA probe containing a high affinity PSE and TATA box. Lane 1 shows the probe alone with no added proteins.
DNA Binding by the SNAP50 Zinc Finger Domain

In this approach, any SNAP50 detected in this analysis had to be in a complex with both SNAP43 and SNAP190-(1–505). As shown in Fig. 2B, the amounts of SNAP50 recovered after the two-step affinity purification were similar for the wild-type and mutant complexes, suggesting that SNAP50 assembly into SNAP$_C$ is not markedly dependent upon any individual cysteine or histidine in this domain.

Next, partially purified wild-type and mutant complexes were then tested for PSE binding function in EMSA to determine whether the cysteine and histidine residues within the SNAP finger domain are important for DNA binding by SNAP$_C$. The results revealed three categories of DNA binding effects, as shown in Fig. 2C. The first category of mutants exhibited either wild-type DNA binding (H330A, H347A, C377A, R385) or modest defects (C302A, C312A, H331A, C334A) as compared with the complex containing wild-type SNAP50. The second category of mutant complexes exhibited severe defects in DNA binding (H313A, C317A, H319A, H388A), whereas complexes in the third category were completely impaired for DNA binding ability (C354A, C357A, C380A, C383A). Thus, the cysteine and histidine residues within the latter two categories are both evolutionarily conserved and critical for DNA binding by SNAP$_C$.

The SNAP50 Zinc Finger Domain Is Required for RNA Polymerase II and III Transcription—We next tested whether mutations that affected PSE recognition by SNAP$_C$ were also critical for human U1 snRNA gene transcription by RNA polymerase II and U6 snRNA gene transcription by RNA polymerase III. For these experiments, we predicted that those complexes that exhibited no DNA binding ability would also be defective in transcription, confirming that DNA binding by SNAP$_C$ is an essential aspect of snRNA gene transcription. Second, any previously unappreciated function for SNAP50 in communication with other components of the general transcription machinery would be revealed as a defect in transcription for those complexes with wild-type DNA binding activity. As the general transcription machinery for RNA polymerase II and III transcription are different, it was possible that certain mutations would result in transcriptional defects in only one polymerase system. However, as shown in Fig. 3A, all complexes that were capable of DNA binding also functioned well for RNA polymerase II and III transcription with the extent of activity essentially parallel for both systems. Thus, those residues that are not evolutionarily conserved are also not critical for human snRNA gene transcription by either polymerase. In contrast, those mutant complexes that were crippled for DNA binding activity (C354A, C357A, C380A, C383A) did not support either RNA polymerase II or III transcription. We consider it likely that these critical cysteine residues within region 2 of the SNAP finger play a structural role for SNAP50 function and are likely important for zinc binding. Some unexpected differences between RNA polymerase II and III transcription were apparent for those mutant complexes that exhibited reduced, not ablated, DNA binding. Some complexes (H313A, C317A, H319A) were moderately active for RNA polymerase III transcription but did not support RNA polymerase II transcription. Thus, the DNA binding defects caused by altering these conserved residues, especially for the H388A substitution are suppressed during RNA polymerase III transcription.

Suppression of DNA Binding Defects by Cooperation between TFIIIB and SNAP$_C$—Based on the previously described observations, we hypothesized that the evolutionarily conserved residues in region 1 (His$^{313}$, Cys$^{317}$, and His$^{319}$), as well as His$^{388}$ in region 2 facilitate DNA contacts by SNAP$_C$, either directly or indirectly, and during RNA polymerase III transcription their activity was restored because TFIIIB stabilized mutant SNAP$_C$ binding to the PSE. To determine whether TFIIIB was capable of restoring DNA binding activity to SNAP$_C$, it was first necessary to establish an assay that measured DNA binding by TFIIIB and SNAP$_C$. Human U6 snRNA gene transcription relies on a variant of TFIIIB called Brf2-TFIIIB, which is composed of TBP, Brf2, and Bdp1 (19). In the following experiments, a truncated form of Bdp1-(1–470) was used because it supports wild-type U6 transcription (30) and can be expressed and purified more easily than the full-length Bdp1-(1–1338). As shown in Fig. 4A (left panel), weak binding of Brf2-TFIIIB to U6 pro-
DNA Binding by the SNAP50 Zinc Finger Domain

FIGURE 4. SNAP$_C$, Brf2-TFIIIB, and TBP recruit DNA in vitro. A, DNA binding by Brf2-TFIIIB is TATA-box dependent. Electrophoretic mobility shift assays were performed using a U6 promoter DNA probe and a wild-type TATA box. DNA binding was carried out in the absence (lanes 1–8) or presence (lanes 9–12) of wild-type SNAP$_C$-4. Reactions containing individual TBP, Brf2, and Bdp1-(1–470) subunits are shown in lanes 2–4. Reactions containing pairwise combinations of TBP with Brf2, TBP with Bdp1-(1–470), and Bdp1-(1–470) with Brf2 are shown in lanes 5, 6, and 8. DNA binding by the complete Brf2-TFIIIB complex in the absence of SNAP$_C$ is shown in lane 7. Additional reactions were performed with SNAP$_C$-H$_3$13A alone (lane 9) or in combination with Brf2 or TBP (lanes 10–12), as indicated. Lane 1 shows migration of the probe alone. The relative positions of the various SNAP$_C$-4/Brd2-TFIIIB complexes are shown on the right. B, coordinated DNA binding by SNAP$_C$ and TBP facilitates higher order complex assembly with Brf2 and Bdp1. DNA binding reactions were performed with the indicated combinations of SNAP$_C$-4 and Brf2-TFIIIB subunits. These results suggest that the preinitiation complex assembly follows the order SNAP$_C$ > TBP > Brf2 > Bdp1.

...SNAP$_C$ (right panel), and for exposures that exhibited prominent DNA complexes containing SNAP$_C$, the DNA binding by Brf2-TFIIIB alone was essentially undetectable (not shown and Fig. 4B). Thus, at least under these conditions, SNAP$_C$ plays a dominant role for TBP recruitment to U6 promoter DNA. Interestingly, serial addition of Brf2 then Bdp1 resulted in complexes that migrated incrementally more slowly, consistent with the idea that increasingly larger complexes are being assembled on the DNA. The amount of the SNAP$_C$-4-TBP-DNA complex was not further affected by Brf2 and Bdp1. In contrast, the amount of the SNAP$_C$-4/TBP-DNA complex was diminished by Brf2 addition, and the SNAP$_C$-4/TBP-Brd2-DNA complex was diminished by Bdp1 addition. These last observations suggest a substrate-product relationship during complex assembly, and are consistent with the idea that Brf2 and Bdp1 both exhibit a binding preference for the higher order complexes containing TBP than that complex containing only SNAP$_C$.

To further refine the pathway for assembly of SNAP$_C$ and Brf2-TFIIIB on U6 promoter DNA, combinations of each factor were tested for DNA binding ability in EMSA. As shown in Fig. 4B, in the absence of SNAP$_C$, none of the Brf2-TFIIIB components bound to DNA when tested singly or in combination under conditions that support robust DNA binding by SNAP$_C$-4. Interestingly, SNAP$_C$-4 was capable of recruiting TBP, but not Brf2 or Bdp1, when tested in pairwise combination, whereas Brf2 could be incorporated into the complex only in reactions containing SNAP$_C$-4 and TBP. This observation further supports the premise that Brf2 preferentially recognizes the SNAP$_C$/TBP promoter bound complex. Similarly, Bdp1 was not recruited to the SNAP$_C$-TBP complex, but again it required the presence of Brf2 for DNA association. Based on these results, we propose a sequential assembly pathway with the initial promoter recognition performed by SNAP$_C$ and TBP followed by Brf2 and Bdp1. The eventual recruitment of Bdp1 is predicted to enable RNA polymerase III recruitment under transcription conditions.

To test the hypothesis that Brf2-TFIIIB can suppress DNA binding defects in SNAP$_C$, recombinant SNAP$_C$ containing wild-type or mutant SNAP50 were tested for DNA binding in the absence or presence of each component of Brf2-TFIIIB (Fig. 5). For this assay, examples from each category of mutant SNAP50 were tested including H388A and H313A that were fully or partially active in U6 transcription, respectively, as well as C383A that was devoid of measurable activity for either DNA binding or transcription. In these assays, the same U6 promoter sequence was used for DNA binding as that contained on the reporter plasmids that were previously used for in vitro U6 transcription. As expected, wild-type recombinant SNAP$_C$ bound DNA well and supported robust recruitment of Brf2-TFIIIB. In contrast, mutant SNAP$_C$ containing SNAP50 (H388A), SNAP50 (C383A), or SNAP50 (H313A) were each inactive when tested alone for DNA binding. Thus, both wild-type and mutant SNAP$_C$ bound with reduced affinity to the U6 promoter sequences tested in this assay relative to those experiments done with the artificial high affinity PSE promoter probes as shown in Figs. 2 and 4. Importantly DNA binding by the SNAP50 (H388A)-containing complex was restored by TBP,
although not to levels seen with the wild-type SNAP50 complex, and higher order complex formation with Brf2 and Bdp1 occurred at nearly wild-type levels. Thus, Brf2-TFIIIB was capable of restoring DNA binding activity to mutant SNAPC. This result stands in contrast to the SNAP50 (C383A) containing complex that was unable to bind DNA under any conditions. Interestingly, TBP did not restore DNA binding by the SNAP50 H313A-containing complex but Brf2 plus TBP did, suggesting that the H313A substitution presents a more dramatic defect to DNA binding by SNAPC. Nonetheless, Bdp1 could be recruited at reduced levels by this SNAP(C(H313A), TBP-Brf2 complex, consistent with the markedly reduced transcription supported by this complex for in vitro U6 transcription.

The COOH-terminal SNAP50 Zinc Finger Binds Zinc—As the previous analyses revealed that certain amino acids within the SNAP finger domain are differentially required for DNA binding and snRNA gene transcription, we postulated that those amino acids that are absolutely critical for both functions are also important for zinc binding. To test this hypothesis, zinc binding studies of wild-type and mutant SNAP50 were undertaken. Whereas our initial analysis of zinc content was performed for the four-member complex (SNAPC,4), suitable amounts of mutant complexes for zinc analysis were not obtained in this context. Nor were we able to obtain sufficient amounts of full-length SNAP50-(1–411) or truncated SNAP50-(315–411) for these studies when expressed individually in E. coli. However, suitable amounts of truncated SNAP50-(315–411) were obtained when expressed as fusion protein with a His-SUMO NH2-terminal tag, and zinc binding studies of this protein were therefore pursued.

As shown in Fig. 6, analysis of SNAP50-(315–411) by flame atomic absorption showed that this region of SNAP50 bound substantial amounts of zinc (~0.7 mol of zinc/mol of protein). This level of zinc binding by the isolated SNAP50 zinc finger domain is comparable with that seen for full-length SNAP50-(1–411) in the context of the four-protein complex SNAPC,4 (~0.9 mol of zinc/mol of protein), indicating that the SNAP finger domain is likely responsible for zinc binding by SNAPC. Zinc binding by SNAP50-(315–411) was markedly reduced by individual alanine substitution at positions Cys354, Cys357, Cys380, and Cys383, whereas alanine substitutions of Cys377 and His388 did not substantially affect zinc binding. Overall, these data indicate that those cysteines within the Cys354-X2-Cys357-X22-Cys380-X2-Cys383 motif are important for zinc binding by SNAPC, whereas the adjacent Cys377 and His388 residues are not. Moreover, SNAP50-(315–411) harboring either the C354A/C380A or C354A/C383A double alanine substitutions was further incapacitated as compared with SNAP50 C354A or C383A, but zinc levels were comparable with the reduced levels observed for SNAP50 harboring the single C380A substitution. Thus, Cys380 plays a more critical role in zinc binding than Cys354 and Cys383.

DISCUSSION

DNA binding by SNAPC is a cooperative event wherein SNAP190, SNAP50, and SNAP43 are all required for promoter recognition (14). Our data demonstrates that an unorthodox zinc finger domain in SNAP50 plays a critical role in...
this process. A comparison of this region with SNAP50 homologues from other species revealed that the arrangement of many cysteine and histidine residues within the SNAP50 COOH-terminal region is remarkably well conserved (Fig. 2A), and a mutational analysis of all histidine and cysteine amino acids throughout this region showed that these highly conserved residues are critical for DNA binding by SNAPC.

The high degree of sequence conservation within the SNAP finger domain throughout evolution suggests that SNAP50 function is conserved in other species. Indeed, *Drosophila* PBP50 (proximal element-binding protein 50 kDa), a homologue of human SNAP50, makes direct DNA contacts within the U6 and U1 promoters although the promoter sequences recognized by human SNAPC and *Drosophila* proximal element-binding protein are different (33). Besides SNAP50, *Drosophila* also maintains homologues of SNAP43 and SNAP190, but not of SNAP45 or SNAP19. The identification of SNAP50- and SNAP43-related proteins in Trypanosoma and Leishmania suggests an ancient function for SNAPC in non-translated RNA production, in this case RNA polymerase II transcription of spliced leader RNA. *Trypanosoma* SNAPC also contains a Myb domain-containing protein (34, 35) reminiscent of human SNAP190, which contains an unusual Myb DNA binding domain (9). The conservation of these three subunits throughout evolution remarkably parallels the experimental definition of a minimal human SNAPC composed of SNAP190-(1–505), SNAP50, and SNAP43 that retains full activity in RNA polymerase II and III transcription (12, 18).

Interestingly, the overall spacing of potential zinc coordination sites within SNAP50 does not resemble other known zinc finger motifs, although the arrangement for a subset of these
cysteine residues resembles that of hormone receptor DNA binding domains (29), consistent with a role in DNA binding for this region in SNAP50. Nonetheless, the unusual arrangement of zinc coordination sites combined with secondary structure predictions suggest that human SNAP50 is the founding member for a novel class of zinc finger domains that we refer to as the SNAP finger domain.

To date, an experimental structure of SNAP50 has not been determined. Therefore, computational methods were employed to predict model structures to assist our understanding of the mechanism for DNA binding by SNAP50. Comparative modeling, which is often successful in other cases, was not possible because no sequence homologs of SNAP50 with known structures are available. However, as the SNAP finger domain is sufficiently short (~100 residues) ab initio modeling based only on the amino acid sequence and the predicted secondary structure was performed. Although ab initio structure prediction methods in general cannot accurately predict protein structures at the level of experimental structures, it is often possible to obtain approximate models of relatively small domains (~100 residues) with an overall root mean square deviation of 5–10 Å from the correct, native structure.

The first round of ab initio prediction with the Robetta server resulted in 10 models for the SNAP finger domain. The resulting models were substantially different; however, the conserved cysteines 354, 357, 380, and 383 were in sufficiently close proximity to serve as zinc coordination sites in 5 of the 10 models. Such a result is non-trivial given that the pair Cys354/Cys357 is in proximity to serve as zinc coordination sites in 5 of the 10 models. The best scoring model consists mainly of β-sheets and a small α-helical segment according to the predicted secondary structure (Fig. 7A), and the predicted structure is shown in Fig. 7B. Submission of the model shown to the DALI server (36, 37) resulted in two known structures with remotely similar topology: domain II from calpain, a cysteine protease (Protein Data Bank code 1KXR) and a β-propeller domain of sialidase (PDB code 1EUT). However, the structural similarity is sufficiently low to suggest that these four cysteines coordinate zinc in a novel zinc finger fold topology. To examine a wider range of possible structures and arrive at a model for the entire SNAP finger domain, additional ab initio sampling was carried out under the constraint that the two pairs of residues, Cys354/Cys380 and Cys357/Cys383, are each in close proximity. The best scoring model main consists of β-sheets and a small α-helical segment according to the predicted secondary structure (Fig. 7A), and the predicted structure is shown in Fig. 7B. Submission of the model shown to the DALI server (36, 37) resulted in two known structures with remotely similar topology: domain II from calpain, a cysteine protease (Protein Data Bank code 1KXR) and a β-propeller domain of sialidase (PDB code 1EUT). However, the structural similarity is sufficiently low to suggest that the predicted SNAP finger fold has a novel architecture. The current model for the SNAP finger domain highlights the zinc coordination by Cys354, Cys357, Cys380, and Cys383, and is independently supported by experimental data indicating that these four cysteine residues are involved in zinc binding. The proposed model is vaguely reminiscent of the GATA-1 zinc finger motif wherein an α-helix plays a critical role in DNA recognition (38), but is substantially different because of a much longer inserted sequence between the two pairs of cysteines. In the SNAP50 model, other highly conserved residues, Gly315, Pro341, and Pro379, are located at critical turn regions for stabilization of the proposed structure.

Interestingly, this SNAP50 model also brings the conserved residues His313, Cys317, and Cys319 in proximity with His388, which could provide an alternate metal binding site. We note, although, that the binding of an additional metal atom to SNAP_c is not supported by the experimental data, and the exact function of these residues remains unclear. These residues may stabilize structures involved in DNA binding by SNAP50, or instead participate directly in DNA contacts. Of note, these residues are located adjacent to the α-helix, which in other zinc finger proteins is frequently used to make specific base contacts within the major groove during DNA binding (29). However, only the COOH terminus of the corresponding α-helix within the SNAP finger domain model is fully exposed, which would likely limit major groove contacts. Thus, stable DNA binding through this region may require additional contacts by flanking residues. Interestingly, the electrostatic surface potential for the SNAP finger domain (Fig. 7C) shows that the presented model clearly distinguishes between positively and negatively charged faces. A large well defined positively charged surface patch surrounding the α-helix suggests the potential for DNA binding, although a specific mode of interaction between SNAP50 and DNA cannot yet be predicted because of the uncertainty associated with this working model. Nonetheless, the current study provides insight into the DNA binding properties of SNAP_c, and identifies the evolutionarily conserved zinc finger domain of SNAP50 as critical for promoter recognition and human snRNA gene transcription.

Acknowledgments—We thank Craig Hinkley, Heather Hirsch, Liping Gu, Dorothy Tappenden, and Justin Bammer for technical assistance, and Tharakeswari Selvakumar for critical reading of the manuscript. We thank David Szymanski for assistance with the ICP-MS.

REFERENCES
1. Hernandez, N. (2001) J. Biol. Chem. 276, 26733–26736
2. Henry, R. W., Ford, E., Mital, R., Mittal, V., and Hernandez, N. (1998) Cold Spring Harbor Symp. Quant. Biol. 63, 111–120
3. Sadowski, C. L., Henry, R. W., Lobo, S. M., and Hernandez, N. (1993) Genes Dev. 7, 1535–1548
4. Murphy, S., Yoon, J. B., Gerster, T., and Roeger, R. G. (1992) Mol. Cell. Biol. 12, 3247–3261
5. Henry, R. W., Sadowski, C. L., Kobayashi, R., and Hernandez, N. (1995) Nature 374, 653–656
6. Henry, R. W., Ma, B., Sadowski, C. L., Kobayashi, R., and Hernandez, N. (1996) EMBO J. 15, 7129–7136
7. Sadowski, C. L., Henry, R. W., Kobayashi, R., and Hernandez, N. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 4289–4293
8. Henry, R. W., Mittal, V., Ma, B., Kobayashi, R., and Hernandez, N. (1998) Genes Dev. 12, 2664–2672
9. Hong, W. M., Henry, R. W., Ma, B., Kobayashi, R., Kanges, N., Matthis, P., Strubin, M., and Hernandez, N. (1998) Mol. Cell. Biol. 18, 368–377
10. Ma, B., and Hernandez, N. (2001) J. Biol. Chem. 276, 5027–5035
11. Ford, E., Strubin, M., and Hernandez, N. (1998) Genes Dev. 12, 3528–3540
12. Mittal, V., Ma, B., and Hernandez, N. (1999) Genes Dev. 13, 1807–1821
13. Ma, B., and Hernandez, N. (2002) Mol. Cell. Biol. 22, 8067–8078
14. Hinkley, C. S., Hirsch, H. A., Gu, L., LaMere, B., and Henry, R. W. (2003) J. Biol. Chem. 278, 18649–18657
15. Gu, L., Esselman, W. J., and Henry, R. W. (2005) J. Biol. Chem. 280, 27697–27704
16. Yoon, J. B., Murphy, S., Bai, L., Wang, Z., and Roeger, R. G. (1995) Mol. Cell. Biol. 15, 2019–2027
17. Krishna, S. S., Majumdar, I., and Grishin, N. V. (2003) Nucleic Acids Res. 31, 532–550
18. Hanzlowsky, A., Jelencic, B., Jawdekar, G. W., Hinkley, C. S., Geiger, J. H., and Henry, R. W. (2006) Protein Expression Purif. 48, 215–223
DNA Binding by the SNAP50 Zinc Finger Domain