

# Two Singular Types of CCCH Tandem Zinc Finger in Nab2p Contribute to Polyadenosine RNA Recognition

Santiago Martínez-Lumbreras,<sup>1,2</sup> Clara M. Santiveri,<sup>1,2,3</sup> Yasmina Mirassou,<sup>1</sup> Silvia Zorrilla,<sup>1,4</sup> and José Manuel Pérez-Cañadillas<sup>1,\*</sup>

<sup>1</sup>Department of Biological Physical Chemistry, Instituto de Química-Física “Rocasolano,” CSIC, Serrano 119, 28006 Madrid, Spain

<sup>2</sup>These authors contributed equally to this work

<sup>3</sup>Present address: Department of Chemical and Physical Biology, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

<sup>4</sup>Present address: Department of Cellular and Molecular Biology, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

\*Correspondence: [jmperez@iqfr.csic.es](mailto:jmperez@iqfr.csic.es)

<http://dx.doi.org/10.1016/j.str.2013.07.019>

## SUMMARY

The seven C-terminal CCCH-type zinc fingers of Nab2p bind the poly(A) tail of mRNA ( $\sim A_{25}$ ). Using NMR, we demonstrated that the first four (Zf1–Zf4) contain two structurally independent tandems (TZF<sub>12</sub> and TZF<sub>34</sub>) and bind A<sub>12</sub> with moderate affinity ( $K_D = 2.3 \mu\text{M}$ ). Nab2p TZF<sub>12</sub> contains a long  $\alpha$  helix that contacts the zinc fingers Zf1 and Zf2 to arrange them similarly to Zf6–7 in the Nab2p Zf5–7 structure. Nab2p TZF<sub>34</sub> exhibits a distinctive two-fold symmetry of the zinc centers with mutual recognition of histidine ligands. Our mutagenesis and NMR data demonstrate that the  $\alpha$  helix of TZF<sub>12</sub> and Zf3 of TZF<sub>34</sub> define the RNA-binding interface, while Zf1, Zf2, and Zf4 seem to be excluded. These results further our understanding of polyadenosine RNA recognition by the CCCH domain of Nab2p. Moreover, we describe a hypothetical mechanism for controlling poly(A) tail length with specific roles for TZF<sub>12</sub>, TZF<sub>34</sub>, and Zf5–7 domains.

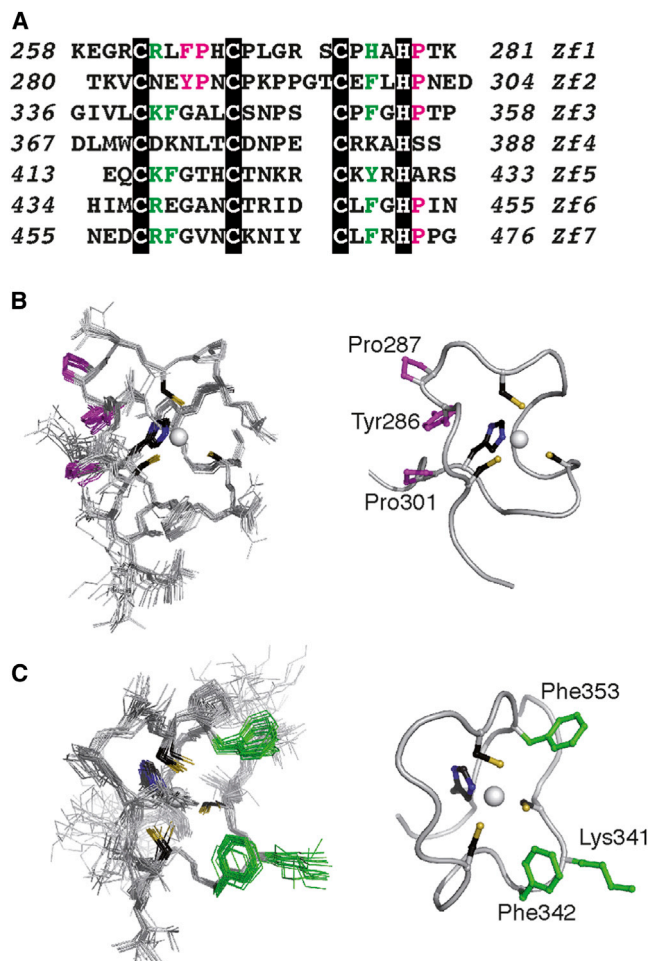
## INTRODUCTION

The 3'-end of eukaryotic mRNAs is processed by endonucleolytic cleavage in the 3'-UTR, followed by the addition of a poly(A) tail (Proudfoot and O'Sullivan, 2002). The length of this tail is 60–80 nucleotides in yeast and approximately 250 nucleotides in high eukaryotes and determines the effectiveness of mRNA export, stability, and translation in the cytoplasm (Eckmann et al., 2011). Cleavage and polyadenylation machineries operate cotranscriptionally and in coordination with other mRNA processing mechanisms to ensure that only properly transcribed and processed mRNAs are endowed with a functional 3'-end tail. Accordingly, control of the poly(A) tail length represents an important mRNA quality control mechanism (Eckmann et al., 2011; Soucek et al., 2012), and transcripts with shorter tails are rapidly degraded by the nuclear exosome (Lebreton and Séra-

phin, 2008). Poly(A)-binding proteins (PABPs) protect against this nuclease activity and are cotranscriptionally recruited to the nascent poly(A) tail.

The three major families of PABPs are represented by PABPC1, which contains four RRM domains in tandem; PABPN1, which has a single RRM; and *Saccharomyces cerevisiae* Nab2p, which contains seven CCCH-type zinc finger domains (Kühn and Wahle, 2004). The mechanisms of polyadenosine RNA recognition differ among each of these PABPs. The cytoplasmic PABPC1 forms linear deposits on the poly(A) tail in which the first two RRMs of each protein copy protect approximately 12 nucleotides of mRNA (Deo et al., 1999). The nuclear PABPs, PABPN1 and Nab2p, also interact with components of the polyadenylation machinery (Dheur et al., 2005; Hector et al., 2002; Kühn et al., 2009) and the nuclear exosome (Lemay et al., 2010; Schmid et al., 2012). These interactions are thought to be crucial for the control of poly(A) tail length (Eckmann et al., 2011; Soucek et al., 2012), and although they have been characterized biochemically, they have yet to be elucidated at the structural level.

Nab2p is a 525-residue protein with a multidomain organization and consists of an N-terminal  $\alpha$ -helical (PWI)-like domain (1–97) that interacts with components of the nuclear pore (Marfatia et al., 2003); a Q-rich domain (104–169), to which no specific functions have been ascribed to date; an RGG domain (201–250), which is involved in nuclear import; and a C-terminal domain that contains seven CCCH-type zinc fingers (251–525) responsible for polyadenosine RNA binding (Anderson et al., 1993; Kelly et al., 2007, 2010; Marfatia et al., 2003). Deletion mutations of the complete CCCH domain or parts thereof result in nonviable cells (Marfatia et al., 2003), underscoring its importance. Moreover, the last three zinc fingers Nab2p Zf5–7 are necessary and sufficient for polyadenosine RNA binding (Kelly et al., 2007, 2010; Marfatia et al., 2003). The structure of this fragment features a novel arrangement of zinc fingers (Brockmann et al., 2012). The polyadenosine RNA binding affinity of Nab2p Zf5–7 ( $K_D = 0.5 \mu\text{M}$  for A<sub>8</sub>; Brockmann et al., 2012) is about 20-fold weaker than that of the full-length protein ( $\sim 20$ – $30 \text{ nM}$  for A<sub>25</sub>; Hector et al., 2002; Kelly et al., 2007, 2010; Tran et al., 2007), and the length of the optimal RNA segment to which this construct binds ( $\sim 10 \text{ nt}$ ) is half that of



**Figure 1. Nab2p Peptides**

(A) Sequence alignment of Nab2p CCCH-type zinc fingers. Numbering corresponds to *S. cerevisiae* Nab2p sequence. Zinc-coordinating residues are shaded in black, and the characteristic Aro-Pro-Pro and cation-Aro-Aro motifs are shown in magenta and green, respectively.

(B and C) Also shown are NMR structures of Nab2p Zf2 (B) and Zf3 (C). On the left are the superpositions of the 20 conformers, and on the right are the ribbon representation of one conformer. Side chains of key residues are colored as in (A).

See also Figures S1–S3, S7, S11, and S12 and Tables S1 and S2.

the full-length Nab2p footprint (Viphakone et al., 2008). Moreover, the Nab2p  $\Delta$ Zf5–7 mutant retains some poly(A) RNA binding affinity and specificity ( $K_D = 1.67 \mu\text{M}$ ; Kelly et al., 2010). Together, these observations suggest the existence of other regions that contribute to polyadenosine RNA recognition in Nab2p, a role that has been tentatively attributed to Nab2p Zf1–4 (Brockmann et al., 2012).

In the present study, we determined the structure, RNA recognition mode, and binding affinities of the Nab2p Zf1–4 construct. This region contains two structurally independent zinc finger tandems (TZF): Nab2p TZF<sub>12</sub> and Nab2p TZF<sub>34</sub>. The first tandem shares certain structural features with Nab2p Zf5–7 and contains a long C-terminal  $\alpha$  helix. The second tandem exhibits a singular interaction between zinc centers. Using fluorescence anisotropy, we determined that Nab2p Zf1–4 can bind A<sub>12</sub> RNA with

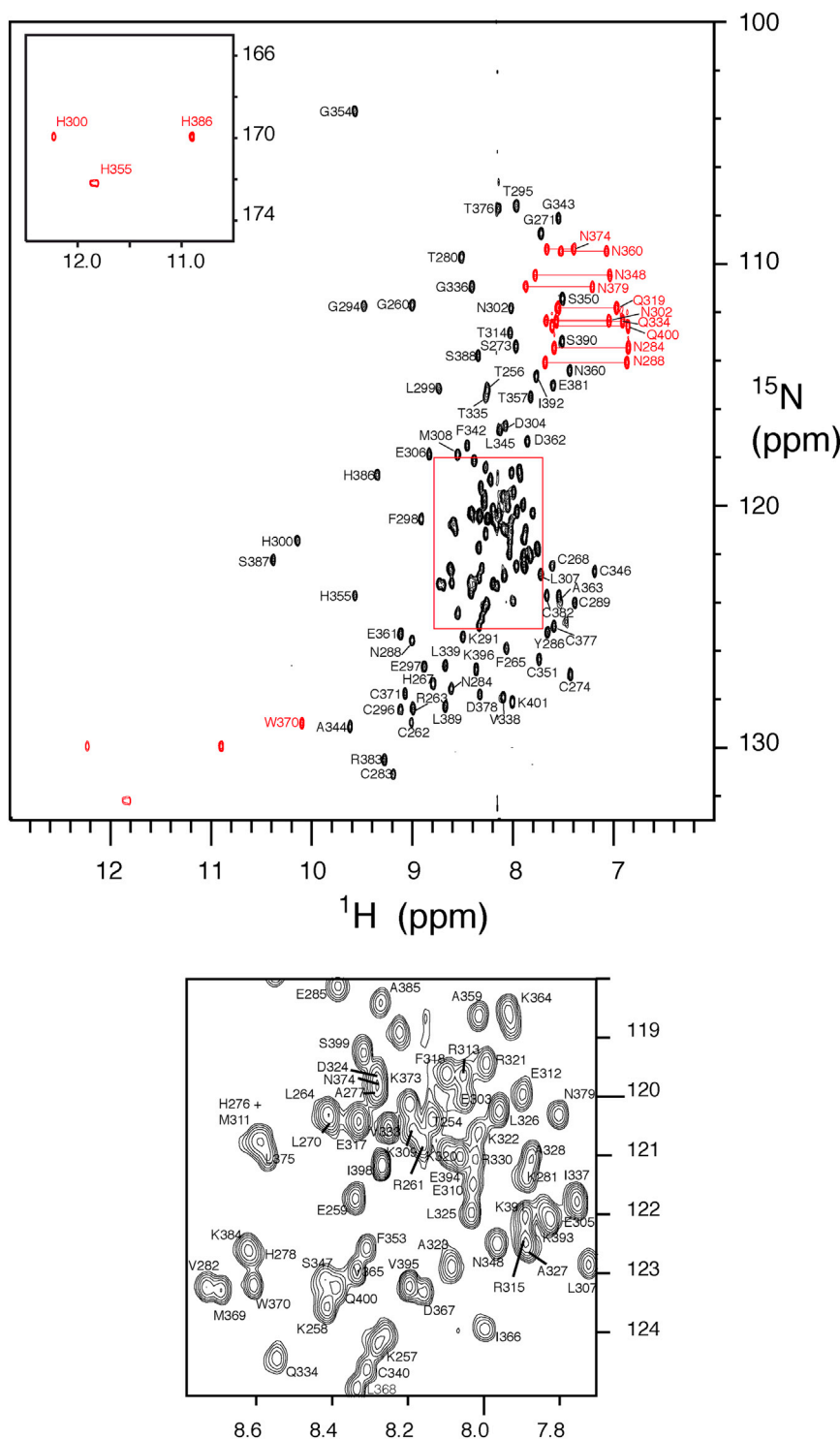
a  $K_D$  of  $2.3 \pm 0.2 \mu\text{M}$ . An extensive mutagenesis analysis identified the RNA interfaces in the two tandems (the  $\alpha$  helix in Nab2p TZF<sub>12</sub> and Zf3 in Nab2p TZF<sub>34</sub>). Together, our findings suggest that Nab2p Zf1–4 cooperates with Nab2p Zf5–7 to reconstitute the poly(A) binding profile of the full-length protein, in agreement with the results of a recently published structural study of this region (Brockmann et al., 2012).

## RESULTS

### The Structures of Individual Zinc Fingers in Nab2p Zf1–4 Resemble Those of Nab2p Zf5–7

We began the structural characterization of Nab2p Zf1–4 by studying the peptides containing each of these zinc fingers within this domain. Sequence analysis revealed that Zf3 closely resembles Zf5 and Zf7, with which it shares conserved side chains (two aromatic and one positively charged residues depicted in green in Figure 1A) that are important for RNA recognition (Brockmann et al., 2012). By contrast, none of these positions are conserved in Zf4, the only zinc finger lacking an aromatic side chain at position +2 of the third cysteine. Nab2p Zf1 and Zf2 differ from the others in that they display a greater separation between the second and third cysteines and, more important, contain two conserved prolines and one aromatic residue (Figure 1A in magenta). Peptides corresponding to these four zinc fingers appear to be intrinsically unstructured in the apo form according to their NMR spectra (Figures S1 and S2 available online). The NMR spectra of Nab2p Zf1 (at 1:1 titration point) showed line broadening indicative of chemical exchange, probably with a small proportion of the apo form (Figure S1). Nab2p Zf4 also showed line broadening, but in this case, the spectrum is similar to the apo form, suggesting that it binds  $\text{Zn}^{2+}$  less efficiently. By contrast, the spectra of Nab2p Zf2 and Zf3 did not show line broadening (Figures S1 and S2) upon  $\text{Zn}^{2+}$  binding, allowing us to calculate their high-resolution NMR structures (Figures 1B and 1C and Table S1) following the standard sequence-specific methodology (Wüthrich, 1986). Because we did not know a priori with which the nitrogen atom of the imidazole ring coordinates the zinc ion, we performed preliminary calculations with the two possible candidates (Ne2 and N $\delta$ 1) and found that coordination via Ne2 is more compatible with the experimental NOE restraints. The same configuration was found in all the protein structures containing CCCH-type zinc fingers deposited in the RCSB protein data bank (seven structures PfamID: PF00642).

The structures of Nab2p Zf2 and Zf3 (Figures 1B and 1C), as determined by NMR, are well ordered and characterized by the absence of regular secondary structure elements. The Nab2p Zf2 and Zf3 folds are very similar to those of Zf5, Zf6, and Zf7 (RMSD values in Table S2) in Nab2p Zf5–7 (Brockmann et al., 2012). The residues that are putatively important for RNA binding in Nab2p Zf3 (green in Figures 1A and 1C) are oriented toward the same face. By contrast, in Nab2p Zf2, the two conserved prolines sandwich the aromatic residue to form a linear arrangement (Figure 1B) on the face of the structure opposite of that which contains the conserved motif of Nab2p Zf3 (Figure 1C). The proline following the aromatic residue adopts a *cis* conformation characterized by the presence of a Tyr286 H $\alpha$ -Pro287 H $\alpha$  NOE cross-peak (Wüthrich, 1986).



**Figure 2. Assigned  $^1\text{H}$ - $^{15}\text{N}$  HSQC Spectrum of Nab2p Zf1–4 Construct with Cross-Peaks Corresponding to Side-Chain Correlations Colored in Red**

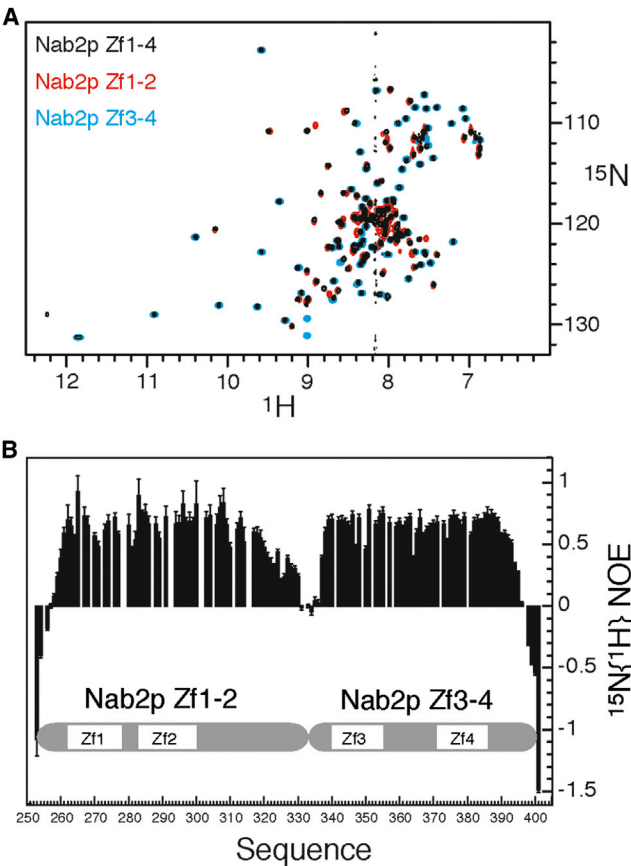
The left inset represents the real positions of the three folded peaks that correspond to  $^1\text{H}\delta_1$ - $^{15}\text{N}\delta_1$  correlations of  $\text{Zn}^{2+}$ -coordinating histidine residues. The assignment of the crowded region in the middle of the spectra (shown in red box) is shown below at higher magnification. See also [Figures S4](#) and [S5](#) and [Table S3](#).

that these four zinc fingers can be grouped into two clusters, Nab2p Zf1–2 and Zf3–4 ([Figure S3](#)), separated by a linker of variable length. We investigated the subdomain organization of the whole Nab2p zinc finger region by analyzing the NMR spectra of several protein constructs of Nab2p ([Table S3](#)). Unfortunately, Nab2p Zf1–7 and Nab2p Zf5–7 exhibited low solubility upon removal of the GST tag. By contrast, freshly made samples of Nab2p Zf1–4 were soluble at concentrations suitable for NMR (up to 300  $\mu\text{M}$ ) and displayed an excellent  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum ([Figure 2](#)). Nab2p Zf1–2 and Zf3–4 constructs also rendered good-quality  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra that matched with that of Nab2p Zf1–4 ([Figure 3A](#)). Nab2p Zf1–2 and Zf3–4 contain long sequence stretches that do not belong to the canonical zinc finger domains. To confirm that these regions are integral parts of the structures, we performed  $^{15}\text{N}$  relaxation experiments using the Nab2p Zf1–4 sample. The sequence profile of  $^{15}\text{N}\{^1\text{H}\}$  NOE values ([Figure 3B](#)) defines two plateaus separated by a narrow depression in the linker between Zf2 and Zf3. Values for the N and C termini are even lower as a result of the high mobility of these regions. The simplest interpretation of the data is that the plateaus, which roughly correspond to the Zf1–2 and Zf3–4 regions, define two individual domains with a slow tumbling rate, separated by highly mobile regions. The data confirm that the large sequence fragments outside the canonical zinc fingers (i.e., the region C-terminal to Zf2 and the linker between Zf3 and Zf4) are parts of the structures.

The  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of Nab2p Zf1–4 contains three folded peaks corresponding to  $^{15}\text{N}$   $\delta$  = 170–172 ppm ([Figure 2](#) inset). Two of these peaks are visible in the Nab2p Zf3–4 spectrum and could be assigned to  $\text{N}\delta_1$ - $\text{H}\delta_1$  correlations of

### Nab2p Zf1–4 Contains Two Structurally Independent Subdomains

The structure of Nab2p Zf5–7 reveals that the three zinc fingers form a unique arrangement ([Brockmann et al., 2012](#)). We thus investigated whether zinc fingers 1–4 can form similar associations. The alignment of Nab2p-like fungal proteins revealed



**Figure 3. NMR Data Show that Nab2p Zf1–4 Contains Two Structurally Independent Subdomains**

(A) Comparison of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of Nab2p Zf1–2 (red), Nab2p Zf3–4 (blue), and Nab2p Zf1–4 (black) constructs reveals a high degree of similarity. (B)  $^{15}\text{N}\{^1\text{H}\}$  NOE values determined for Nab2p Zf1–4 construct. The  $^{15}\text{N}\{^1\text{H}\}$  NOE profile is consistent with a model of two dynamically independent subdomains as defined below. See also Table S3.

the zinc-coordinating His355 and His386 by following an elaborate assignment strategy (Figure S4). The spectrum of the Nab2p Zf1–4 (and Nab2p Zf1–2) construct contains a third N $\delta$ 1–H $\delta$ 1 correlation (Figure 2), which we assigned to His300 N $\delta$ 1–H $\delta$ 1 on the basis of its NOE network (Figure S5). The observation of these resonances suggests that the structural environment somehow prevents the rapid exchange of protons with the solvent (the standard scenario).

#### Nab2p Zf1–2 and Zf3–4 Form Two Tandem Zinc Fingers

We calculated the structures of the Nab2p Zf1–2 and Zf3–4 constructs by NMR (statistics in Table 1). Both regions form compact structures with a fixed relative orientation of zinc fingers (Figure 4; Figure S6). The absolute chirality of the  $\text{Zn}^{2+}$  centers is Z (Berg, 1988) for all the conformers and other geometry values are similar to those of related structures (Table S4). The most prominent feature of the Nab2p TZF<sub>12</sub> global fold (Figure 4A) is a C-terminal  $\alpha$  helix that apparently fixes the orientation of the two zinc fingers. The Nab2p TZF<sub>34</sub> structure (Figure 4B) does not contain regular secondary structural elements other than a short  $\alpha$  helix

**Table 1. Structure Statistics for the Final 20 Structure Ensembles of Nab2p TZF<sub>12</sub> and Nab2p TZF<sub>34</sub>**

	Nab2p TZF <sub>12</sub>	Nab2p TZF <sub>34</sub>
<b>NMR Experimental Restraints</b>		
<b>NOE-derived</b>		
Intraresidue	266	215
Sequential	239	288
Medium-range <sup>a</sup>	191	141
Long-range <sup>b</sup>	282	430
Total	978	1,074
$^3\text{J}(\text{NH-H}\alpha)$ -derived $\phi$ angle	41	42
<b>Structure Statistics</b>		
<b>Mean AMBER energies (kcal/mol <math>\pm</math> SD)</b>		
Total	$-2,999 \pm 12$	$-1,849 \pm 9$
Van der Waals	$-547 \pm 7$	$-473 \pm 8$
Restraints (distance + angle)	$21 \pm 2$	$16 \pm 1$
<b>Violations</b>		
Distance <sup>c</sup>	$16 \pm 2$	$12 \pm 2$
Maximum distance violation ( $\text{\AA}$ )	0.23	0.21
Angle <sup>d</sup>	0–2	0–2
Maximum angle violation ( $^\circ$ )	4.03	6.77
<b>RMS from Ideal Geometry</b>		
Bond lengths ( $\text{\AA}$ )	$0.01699 \pm 0.00002$	$0.01998 \pm 0.00003$
Bond angles ( $^\circ$ )	$2.45 \pm 0.03$	$2.32 \pm 0.03$
Average pairwise rmsd (range)	(259–316)	(335–394)
N, C', C $\alpha$ ( $\text{\AA}$ ) ( $\pm$ SD)	$0.68 \pm 0.20$	$0.68 \pm 0.22$
All heavy ( $\text{\AA}$ ) ( $\pm$ SD)	$1.56 \pm 0.23$	$1.28 \pm 0.20$

<sup>a</sup> $a_1 < (\text{residue } i - \text{residue } j) < 4$ .

<sup>b</sup> $(\text{residue } i - \text{residue } j) > 4$ .

<sup>c</sup>Average number of distance violations  $> 0.1 \text{ \AA}$  per structure.

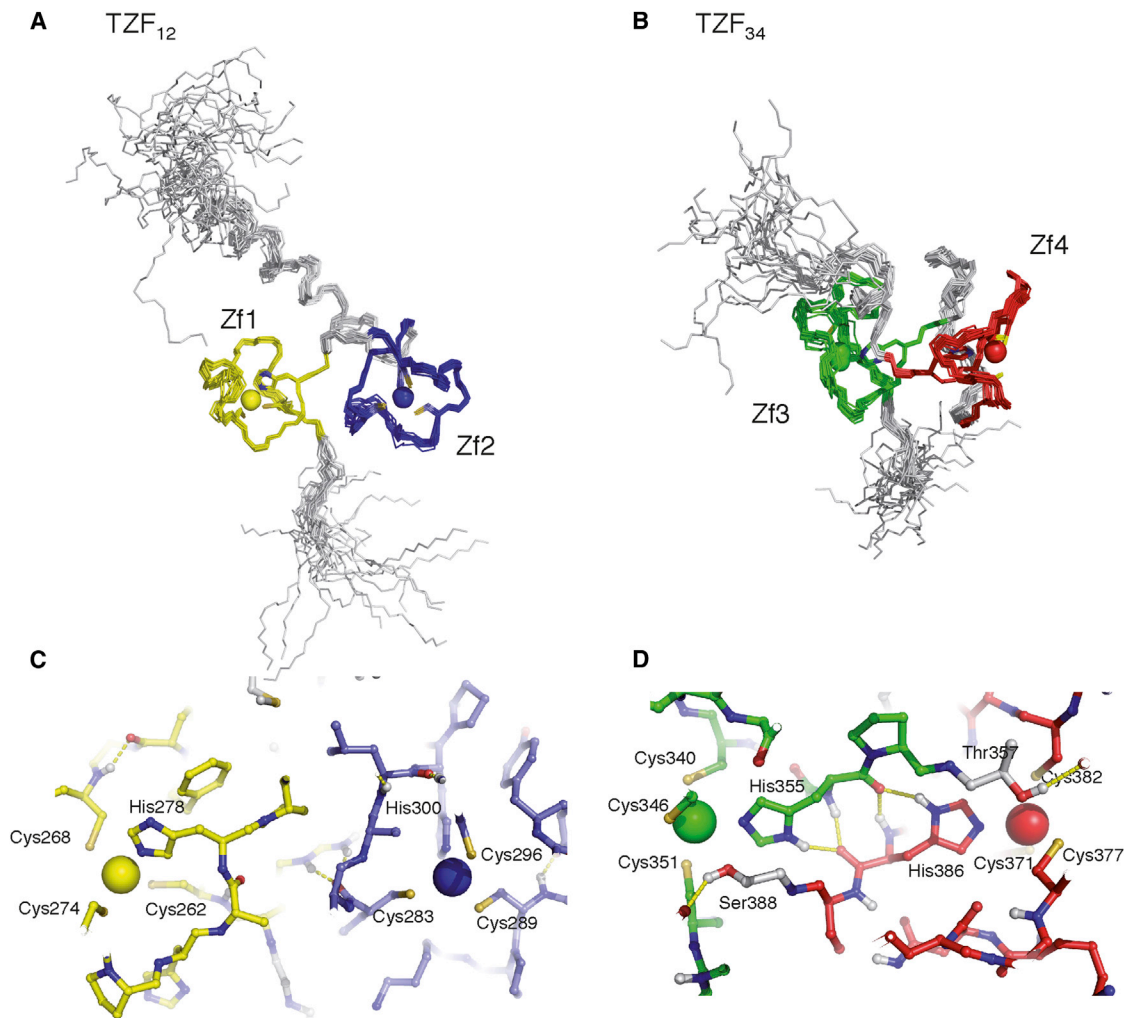
<sup>d</sup>Range of total angle violations ( $< 10^\circ$ ).

at the C terminus in some of the conformers. The linker between Zf3 and Zf4 is highly structured and contacts with Zf3 and Zf4, in agreement with the  $^{15}\text{N}$  relaxation data (Figure 3B).

The  $^1\text{H}$ - $^{15}\text{N}$  HMBC spectrum demonstrates that His355 and His386 are in the N $\delta$ 1-H tautomer form (N $\delta$ 1, 170–172 ppm; N $\epsilon$ 2  $\sim$ 220 ppm; Legge et al., 2004), and therefore metal coordination should take place via N $\epsilon$ 2. This hypothesis was confirmed (also for zinc fingers in TZF<sub>12</sub>) by comparing the compatibility of the NMR experimental restraints with alternative histidine coordination modes in preliminary structure calculations. The first coordination shell of the metal is completed by three cysteines displaying  $^{13}\text{C}\beta$  chemical shifts ( $\sim$ 30.0 ppm), consistent with their zinc-binding character (Kornhaber et al., 2006).

The structures of Nab2p TZF<sub>12</sub> and TZF<sub>34</sub> form an intricate network of hydrogen bonds that defines a second coordination shell. Sulfur atoms of the inner-shell cysteines are hydrogen-bound to backbone amide groups of residues at position +2 (Table S5). Similar interactions have been described for other CCCH structures (Hudson et al., 2004), which also exhibited donor-acceptor distances that were in general slightly longer than those of other sulfur-mediated hydrogen bonds in proteins (Zhou et al., 2009), indicating that these interactions are probably weak. The zinc-binding histidines also participate in the second





**Figure 4. Structures of the Nab2p TZF<sub>12</sub> and TZF<sub>34</sub> Tandem Zinc Fingers**

(A and B) Backbone superposition of the final 20 structures of Nab2p TZF<sub>12</sub> and TZF<sub>34</sub> are shown. Fitting was achieved using C', C $\alpha$ , and N backbone atoms and the residue ranges specified in Table 1. Zn<sup>2+</sup> ions are represented as spheres, and regions outside the canonical boundaries are shown in gray.

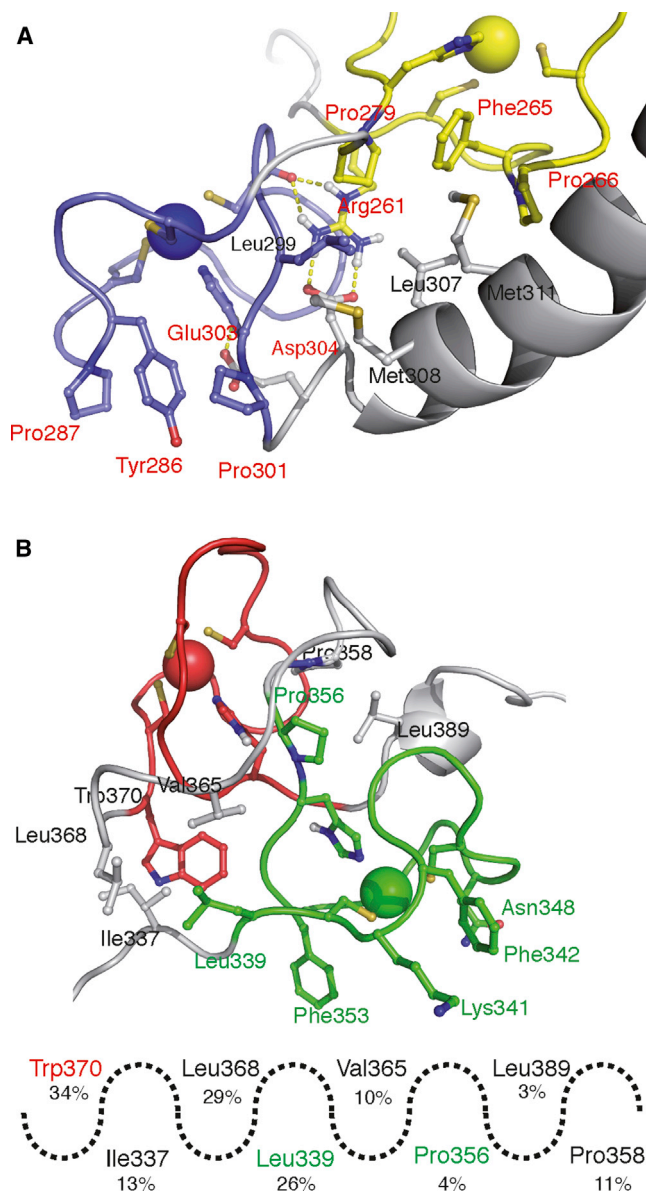
(C and D) Detailed views of the structure around the zinc centers are shown. Carbons are colored as in (A) and (B): oxygens in red, nitrogens in dark blue, and sulphurs in yellow. Protons involved in hydrogen bonds (yellow dashed lines) are depicted in white.

See also Figures S6–S8, S11, and S12, Tables S2–S5, and 3D Molecular Models S1 and S2.

coordination shell (Figure 4D). In Nab2p TZF<sub>34</sub>, the H $\delta$ 1 of His355 forms a hydrogen bond with His386 O that is equivalent to the bond between His386 H $\delta$ 1 and His355 O (Figure 4D). Overall, these histidine residues define a planar dyad with the two zinc centers organized around a two-fold symmetry axis, forming an arrangement with no resemblance to other metalloproteins in the PDB. The exceptionally low solvent exchange rates of the H $\delta$ 1 resonances can now be rationalized in terms of their buried character (ASAs ~0%) and their involvement in the hydrogen bonds described. His300 N $\delta$ 1–H $\delta$ 1 correlation is also observed in the TZF<sub>12</sub> spectrum, although there is evidence of a slightly increased solvent exchange rate because of a less restrictive geometrical environment. Indeed, His300 H $\delta$ 1 does not have a clear hydrogen bond acceptor and the most likely candidate (Glu303 side chain) only establishes such an interaction (Figure 5A) in less than 25% of the conformers of the

ensemble. Moreover, an NOE (or exchange) peak between water and His300 H $\delta$ 1 observed in the spectra (Figure S5) suggests that this site is still partially accessible to solvent or that a resident water molecule bridges the interaction between His300 and Glu303. Conversely, no water NOE (or exchange) peaks were observed for His355 and His386 H $\delta$ 1 of TZF<sub>34</sub>, in agreement with their buried character (Figure S5).

Additional contacts in Nab2p TZF<sub>12</sub> determine the relative orientations of the zinc fingers. The positively charged Arg261 (Zf1) is buried (ASA < 9%) and interacts with the Asp304 carboxyl group (Zf2) via the H $\eta$ 12 and H $\eta$ 22 protons (in more than 50% of structures) and with Cys296 O through a bifurcated hydrogen bond involving H $\epsilon$  and H $\eta$ 21 protons (Figure 5A). The methyl group of the conserved Thr295 (Zf2) forms a long-range contact with the aliphatic portion of Glu259 (just before Zf1) and its H $\gamma$ 1 exchanges slowly with the solvent, likely via the formation



**Figure 5. Detailed Views of Key Contacts in the Structures of Nab2p TZF<sub>12</sub> and Nab2p TZF<sub>34</sub>** Color-Coded as in Figure 4

(A) Interactions between the Phe265-Pro266-Pro279 motif in Zf1 and hydrophobic residues of the C-terminal  $\alpha$  helix together with the hydrogen bond network between Arg261 and Asp304 determine the relative orientations of Zf1, Zf2, and the C-terminal helix. The Tyr286-Pro287-Pro301 motif in Zf2 is solvent-exposed.

(B) The linker between Zf3 and Zf4 folds back onto the zinc fingers to form a zipper-like structure. A schematic view including averaged solvent accessibilities across the NMR ensemble is shown below. Residues potentially involved in RNA recognition according to mutagenesis results are Lys341, Phe342, and Phe353 (Table 2) and are solvent-accessible in the structure. See also 3D Molecular Models S1 and S2.

of a hydrogen bond with the main chain carbonyl group of Pro292. Perhaps the most interesting contact in Nab2p TZF<sub>12</sub> is the hydrophobic cluster formed between the Phe265-Pro266-Pro279 motif of Zf1, Leu299 of Zf2, and Leu307,

Met308, and Met311 of the C-terminal helix (Figure 5A). Interestingly, the twin motif in Zf2, Tyr286-Pro287-Pro301 (Figures 1A and 5A), has a similar three-dimensional arrangement but is solvent-exposed.

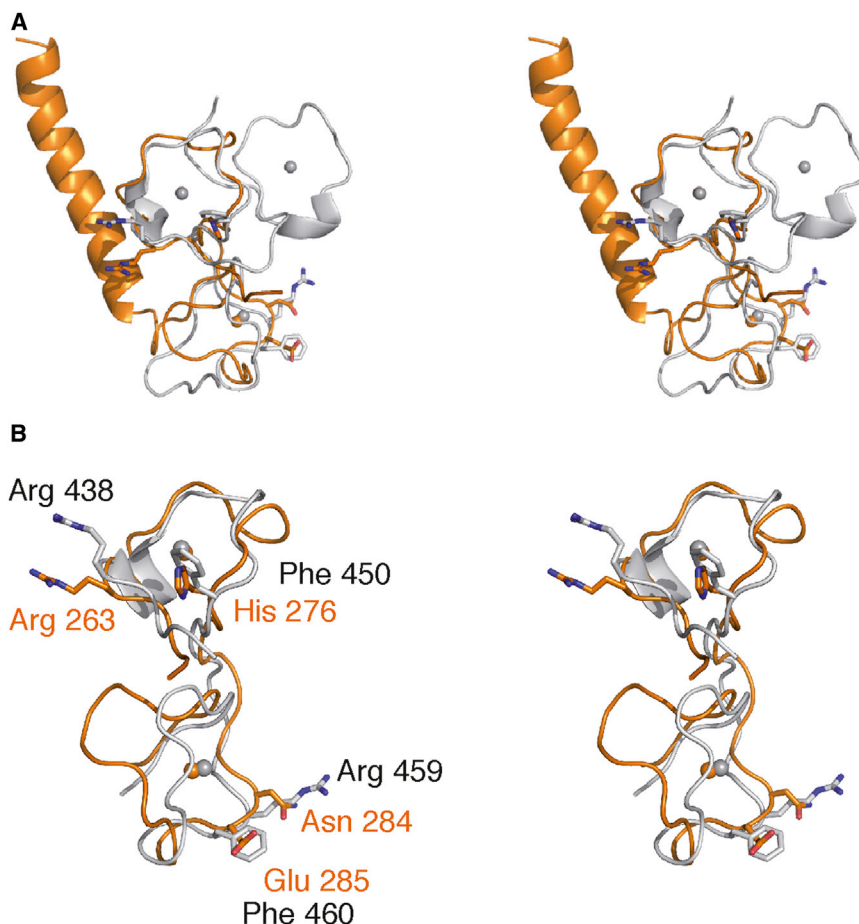
A complex network of interactions involving zinc centers (see above), the interdomain linker, and N- and C-terminal segments maintains the tertiary structure of Nab2p TZF<sub>34</sub>. The first three residues of this linker (Pro356-Thr357-Pro358) exhibit very low solvent accessibility (0% in the case of Thr357). Leu389 (at the C-terminal portion of TZF<sub>34</sub>) is sandwiched between these buried proline residues (Figure 5B), forming a well-defined hydrophobic cluster that propagates on the other side of Pro356 via a chain of residues (Val365, Leu339, Leu368, and Ile337) that interlock like the teeth of a zipper (Figure 5B). The interaction between Ile337 and the totally conserved Trp370 at the C terminus of the linker further define its conformation.

Polar residues Thr357 and Ser388 are completely buried in the structure of Nab2p TZF<sub>34</sub> and their hydroxyl resonances, which form well-defined hydrogen bonds (Thr357 H $\gamma$ 1-Asn379 O and Ser388 H $\gamma$ -Asn348 O), can be observed in the 2D NOESY spectrum. These two buried hydroxyl groups occupy similar positions with respect to the tetrahedral metal center (Figure 4D) and exhibit O-Zn<sup>2+</sup> distances in the 3.5–4.0 Å range that suggest a role in the stabilization of the zinc center.

#### Nab2p TZF<sub>12</sub> Is Structurally Similar to the Zf6–7 Tandem in Nab2p Zf5–7

The structures of Nab2p Zf2 and Zf3 peptides are very similar to the equivalent regions in the tandems (Figure S7; Table S2). However, the loop between Cys289 and Cys296 in Nab2p Zf2 undergoes a small alteration to accommodate the head of the C-terminal helix in Nab2p TZF<sub>12</sub>. The extra two residues in this Zf2 loop (Figure 1A) may be necessary to optimize the interaction with the C-terminal helix.

The Nab2p TZF<sub>34</sub> tandem displays little similarity with Nab2p TZF<sub>12</sub> and with the three subpairs in the structure of Nab2p Zf5–7 (Brockmann et al., 2012; Table S2). This is because the interaction between the two zinc centers (Figure 4D) results in a shorter distance between the zinc ions than in other CCH-containing structures (Figure S8). By contrast, the global orientation of Nab2p TZF<sub>12</sub> zinc fingers is very similar to the structure of the zinc finger pair 6–7 within Nab2p Zf5–7 (mean pairwise RMSD value of 1.8 Å; Table S2). In Nab2p Zf5–7, the contacts between each pair of zinc fingers provide structural coherence (Brockmann et al., 2012). Similarly, Zf1 and Zf2 form a coherent structure in Nab2p TZF<sub>12</sub>, in which the C-terminal helix rather than a third zinc finger helps to maintain coherence (Figure 6A). This structural alignment identifies a number of residues in Nab2p TZF<sub>12</sub> that may be important for RNA binding if this tandem shares a similar recognition mechanism to that described for Nab2p Zf5–7 (Brockmann et al., 2012). Residues Arg438, Phe450, Arg459, and Phe460 of Zf6–7 are involved in polyadenosine RNA binding. Of these four residues, only Arg263 (in Zf1) is conserved in Nab2p TZF<sub>12</sub> (equivalent to Arg438). The aromatic character of His276 (Phe450 in Zf6) appears to be preserved, while the two positions in Zf7 (Arg459 and Phe460) are not conserved in Zf2 (Asn284 and Glu285). The effects of these variations on the RNA binding mode were investigated and are discussed in the following section.



**Figure 6. Stereoview of Structural Fitting between Nab2p TZF<sub>12</sub> and Zf5–7**

(A) Nab2p TZF<sub>12</sub> (orange) and Zf5–7 (gray; PDB: 2LHN) are shown in ribbon. Zf1, Zf2, Zf6, and Zf7 regions were used in the alignment.

(B), Labeled side chains (in sticks) of Zf5–7 residues involved in RNA binding and their equivalent positions in TZF<sub>12</sub>. Zf5 and the C-terminal helix of Nab2p TZF<sub>12</sub> have been omitted for clarity.

See also Figures S8, S11, and S12 and Table S2.

strength and most of the signals disappear under substoichiometric conditions at 800 MHz (proton field). The quality of the spectra improves when a 1:1 RNA:protein ratio is reached (Figure S10) and the spectra remain unchanged at a 2:1 RNA excess (data not shown), suggesting that the stoichiometry of the complex is equimolecular. However, the quality of the triple resonance data is insufficient to assign the complex spectra in full, as previously described for Nab2p Zf5–7 (Brockmann et al., 2012). RNA titrations monitored in the <sup>1</sup>H-<sup>15</sup>N HSQC spectrum of Nab2p Zf1–4 (Figure 7B) cause larger changes in the resonances of the first tandem ZF region (Nab2p TZF<sub>12</sub>), probably because they are close to the intermediate exchange regime, resulting in a low percentage of assignment (34% of the NHs in segment 253–333).

The impossibility of assigning C-terminal

#### Polyadenosine RNA Binding Mode of Nab2p Zf1–4

Full-length Nab2p binds polyadenosine RNA with high affinity and specificity ( $K_D \sim 20$ – $30$  nM for A<sub>25</sub>) (Hector et al., 2002; Kelly et al., 2007, 2010; Tran et al., 2007) via its C-terminal region, which contains seven CCHC zinc fingers. The Nab2p Zf5–7 construct has been reported to bind 8–10 nucleotides with a  $K_D$  of 0.5–0.1  $\mu$ M (Brockmann et al., 2012). To complement this study, we examined the RNA binding mode of the Nab2p Zf1–4 construct by fluorescence anisotropy and found that it binds to A<sub>12</sub> and A<sub>10</sub> probes with a  $K_D$  of  $\sim 2$   $\mu$ M and  $\sim 4$   $\mu$ M, respectively (Table 2; Figure 7A). These affinities were further confirmed by competition experiments against the A<sub>12</sub> probe with unlabeled RNAs of various lengths (8–20 nt) (Table 2; Figure S9A). These experiments identified 12 nucleotides as the optimal RNA length bound by Nab2p Zf1–4 (Table 2; Figure S9B). Nab2p Zf1–4 displays modest polyadenosine specificity as binding to (AU)<sub>6</sub> and U<sub>12</sub> is only between 3- and 5-fold weaker than binding to A<sub>12</sub> (Table 2; Figure 7A; Figure S9C). Of the two structural units contained in Nab2p Zf1–4, only TZF<sub>12</sub> produced measurable changes in the anisotropy of the A<sub>12</sub> probe, although binding was significantly weaker as compared with Nab2p Zf1–4 (Table 2).

The <sup>1</sup>H-<sup>15</sup>N HSQC spectrum of Nab2p Zf1–4 underwent significant line broadening upon titration with A<sub>12</sub> (Figure S10). The exchange regime is heavily dependent on the spectrometer field

helix cross-peaks of the bound state reveals that either it directly binds poly(A) or undergoes a global conformational change in the process. By contrast, most of the resonances of the TZF<sub>34</sub> region could be assigned in the bound form (72% of the NHs in segment 335–401). All side-chain NH<sub>2</sub> groups could be assigned in the free and bound forms (Figure 7C). The largest changes upon A<sub>12</sub> binding correspond to Asn348 (Zf3) and Asn284 (Zf2). The perturbations of Asn302 and Gln319 support the view that the  $\alpha$  helix of TZF<sub>12</sub> is involved in RNA binding. By contrast, Zf4 does not appear to participate in RNA binding because Asn374 and Asn379 side chains remain basically unchanged. Moreover, Asn379 is equivalent to Asn348 (Zf3), Asn423 (Zf5), and Asn466 (Zf7), all of which are sensitive to RNA binding (Figure 7C; Brockmann et al., 2012).

To further investigate the RNA binding mode, we measured the affinity of Nab2p Zf1–4 mutants to A<sub>12</sub> by fluorescence anisotropy (Table 2; Figures S9D–S9F). The protein variants were designed based on previous mutagenic studies of Nab2p Zf5–7 (Brockmann et al., 2012) and the structural similarities between Nab2p TZF<sub>12</sub> and Nab2p Zf5–7 (Figure 6) and between the zinc fingers of Nab2p Zf1–4 and those of the Tis11d:RNA complex (Hudson et al., 2004; Figures S11 and S12).

In the second tandem, Zf3 resembles Nab2p Zf5 and Zf7 (Figure 1A) and Tis11d (Figure S12). Consistent with this observation, K341A/F342A and F353A mutants display a significant



**Table 2. Energetic Parameters of Nab2p-RNA Interactions with Different RNA Probes and Nab2p Zf1–4 Mutants**

Protein	RNA	$\Delta G$ (kcal/mol)	$K_D$ ( $\mu$ M)
Nab2p Zf1–4	A <sub>12</sub>	7.56 $\pm$ 0.05 <sup>a</sup> (7.7 $\pm$ 0.3 <sup>b</sup> )	2.3 $\pm$ 0.2 (2 $\pm$ 1)
Nab2p Zf1–4	A <sub>10</sub>	7.18 $\pm$ 0.05 <sup>a</sup> (7.1 $\pm$ 0.2 <sup>b</sup> )	4.4 $\pm$ 0.4 (5 $\pm$ 2)
Nab2p Zf1–4	A <sub>20</sub>	7.6 $\pm$ 0.1 <sup>b</sup>	2.1 $\pm$ 0.4
Nab2p Zf1–4	A <sub>14</sub>	7.6 $\pm$ 0.2 <sup>b</sup>	2.1 $\pm$ 0.7
Nab2p Zf1–4	A <sub>8</sub>	6.6 $\pm$ 0.2 <sup>b</sup>	12 $\pm$ 4
Nab2p Zf1–4	U <sub>12</sub>	6.6 $\pm$ 0.2 <sup>a</sup> (6.7 $\pm$ 0.2 <sup>b</sup> )	12 $\pm$ 4 (10 $\pm$ 3)
Nab2p Zf1–4	(AU) <sub>6</sub>	7.2 $\pm$ 0.2 <sup>b</sup>	7 $\pm$ 2
Nab2p Zf1–2	A <sub>12</sub>	<5.0 <sup>a,c</sup>	>190
R263A	A <sub>12</sub>	7.6 $\pm$ 0.1 <sup>a</sup>	2.1 $\pm$ 0.4
H276A	A <sub>12</sub>	8.0 $\pm$ 0.2 <sup>a</sup>	1.0 $\pm$ 0.4
F298A	A <sub>12</sub>	7.7 $\pm$ 0.1 <sup>a</sup>	1.8 $\pm$ 0.3
N348A	A <sub>12</sub>	7.6 $\pm$ 0.1 <sup>a</sup>	2.1 $\pm$ 0.4
K341A/F342A	A <sub>12</sub>	5.7 $\pm$ 0.2 <sup>a</sup>	56 $\pm$ 19
F353A	A <sub>12</sub>	6.0 $\pm$ 0.1 <sup>a</sup>	33 $\pm$ 6
N284R/E285F	A <sub>12</sub>	8.6 $\pm$ 0.1 <sup>a</sup>	0.38 $\pm$ 0.06
D372K/K373F/K384F	A <sub>12</sub>	7.7 $\pm$ 0.1 <sup>a</sup>	1.8 $\pm$ 0.3
K309A	A <sub>12</sub>	7.2 $\pm$ 0.2 <sup>a</sup>	4 $\pm$ 1
K320A/R321A/K322A	A <sub>12</sub>	6.1 $\pm$ 0.2 <sup>a</sup>	28 $\pm$ 9

The concentration of labeled RNA was 22 nM, in 25 mM Tris-HCl (pH 8.1), 100 mM NaCl, 0.1 mM DTT. The uncertainties in the  $\Delta G$  values were calculated by rigorous confidence limit testing at the 67% level using BIOEQS software and those in the  $K_D$  values by error propagation. See also Figure S9 and Table S3.

<sup>a</sup>Free energy values obtained from fluorescence anisotropy-based titrations of the indicated fluorescein-labeled (at 3'-phosphates: [36FAM]) RNA probe with protein.

<sup>b</sup>Values determined from anisotropy titrations of the rA<sub>12</sub>-36FAM probe with protein in the presence of 20  $\mu$ M of the indicated unlabeled RNA.

<sup>c</sup>Values are more imprecise because of the low affinity of the interaction, which precluded reaching saturation.

drop in binding affinity, similar to that reported for K416A (Zf5), F417E (Zf5), R438A (Zf6), F450A (Zf6), R459A (Zf7), and F460E (Zf7) (Brockmann et al., 2012). Asn348 (Zf3) is conserved in Nab2p Zf5 and Zf7 and is structurally equivalent to Tis11d Tyr176 (Figure S12). Moreover, its side chain is strongly perturbed upon RNA binding (Figure 7C). However, Nab2p Zf1–4 N348A exhibits no variation in RNA binding affinity.

In the first tandem ZF region of Nab2p Zf1–4, Arg263 and His276 are structurally equivalent to Arg438 and Phe450 in Nab2p Zf5–7 (Figure 6B) and to Arg160 and Phe176 in Tis11d (Figure S11A). Surprisingly, the R263A and H276A mutations do not cause a decrease in RNA affinity, as observed with R438A and F450A in Nab2p Zf5–7, and H276A even exhibits a slight increase in RNA affinity. Strikingly, Phe298 (Zf2), which is structurally equivalent to Phe214 in Tis11d (Figure S11B), does not seem to be required for adenine recognition in Nab2p Zf1–4 (mutant F298A; Table 2). On the other hand, some critical side chains for RNA binding in Zf7 are not conserved in Zf2 (Figures 1A and 6B). Accordingly, restoring the configuration of the Zf6–7 tandem (N284R/E285F mutant; Table 2) increases the

affinity to values similar to those of Nab2p Zf5–7 (Brockmann et al., 2012). Overall our data suggest that the presumed RNA-binding interface of Nab2p is not optimally designed for this purpose. In an attempt to identify this interface in TZF<sub>12</sub>, we generated mutations in the C-terminal helix, whose involvement in RNA recognition is indirectly suggested by our NMR data. The replacement of conserved (and exposed) Lys and Arg residues by Ala decreases the overall affinity (K309A and K320A/R321A/K322A mutants; Table 2), indicating that this element participates in RNA binding, very likely through electrostatic interactions.

Finally, we redesigned Nab2p Zf4 (D372K/K373F/K384F) to resemble the Zf3 RNA-binding interface. Surprisingly, the affinity of this mutant is largely unchanged as compared with the wild type. The remodeled Zf4 interface thus cannot cooperate with the Zf3 interface, probably because they are located on opposite faces of the Nab2p TZF<sub>34</sub> tandem.

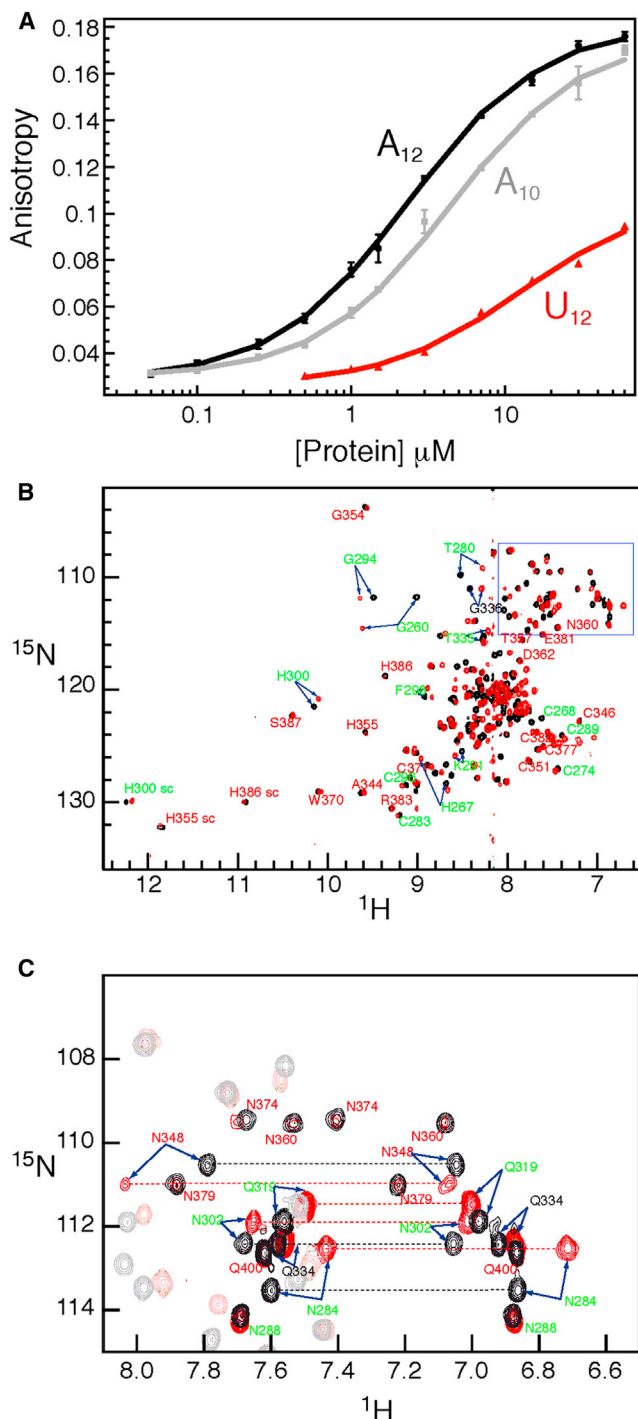
## DISCUSSION

A previous study of the structure and polyadenosine RNA binding mode of Nab2p Zf5–7 (Brockmann et al., 2012) described a novel arrangement of three CCCH-type zinc fingers and an RNA binding mode partially resembling that reported for other protein-RNA complexes containing CCCH-type zinc fingers (Hudson et al., 2004; Teplova and Patel, 2008). Here, we performed a similar analysis of the Nab2p Zf1–4 region and found that it contains two structurally independent tandems. Nab2p TZF<sub>12</sub> shares some structural similarity with Nab2p Zf5–7, contains a C-terminal  $\alpha$  helix that contacts Zf1 and Zf2, and probably maintains these zinc fingers in a fixed relative orientation. A similar role has been described for Zf6, which maintains the structural coherence of Nab2p Zf5–7 (Brockmann et al., 2012). Conversely, in Nab2p TZF<sub>34</sub> the hydrogen bond network between the two zinc-coordinating histidines appears to be the principal factor maintaining structural coherence. This arrangement is atypical among CCCH-type structures.

The Nab2p zinc finger domain binds polyadenosine RNA (A<sub>25</sub>) with a binding affinity in the 20–30 nM range (Hector et al., 2002; Kelly et al., 2007, 2010; Tran et al., 2007). Nab2p Zf5–7 binds only 10 nucleotides with high affinity (Brockmann et al., 2012), and we have demonstrated that Nab2p Zf1–4 can bind 12 nt with modest affinity in the micromolar range. We identified the RNA-binding interfaces (Figure 8) within this region: Zf3 and the C-terminal helix of Nab2p TZF<sub>12</sub>. The Zf3 binding pocket appears to employ a recognition mechanism similar to that of Nab2p Zf5–7 (affinity drops in K341A/F342A and F353A mutants), with which it may cooperate in the specific recognition of longer polyadenosine segments (Figure 8). By contrast, binding to the Nab2p TZF<sub>12</sub> helix probably occurs via unspecific electrostatic contacts.

We speculate that Nab2p Zf1–4 may bind RNA in the 3' to 5' direction (i.e., from the 5'-end of RNA to the C-terminal zinc finger; Figure 8) in a manner analogous to the binding modes of related protein-RNA complexes (Figure S8B; Hudson et al., 2004; Teplova and Patel, 2008). Based on this recognition model (Figure 8), we propose a mechanism of poly(A) tail length control that shares key features with the mammalian model (Eckmann et al., 2011; Kühn et al., 2009; Figure S13A). In our model, the nascent poly(A) tail first binds to Nab2p Zf5–7, the high-affinity





### Figure 7. Polyadenosine RNA Binding by Nab2p Zf1-4

(A) Fluorescence anisotropy binding titrations of fluorescein-labeled  $A_{12}$  (black),  $A_{10}$  (gray), and  $U_{12}$  (red) with Nab2p ZF1–4. Solid lines represent the fit of the model indicated in the main text to the experimental data and error bars the standard deviation of three measurements.

(B and C) Superposition of Nab2p21 ZH1-4  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of the free form (in black) and the 1:1 protein A<sub>12</sub> RNA titration point (in red; at 600 MHz). (B) Representative assignments are included and labels are colored in green (region corresponding to TZF<sub>12</sub>) and red (region corresponding to TZF<sub>34</sub>). Labels for residues in the boundary between the tandems are colored in black. Among the assigned peaks, those with chemical shift perturbations > 0.1 ppm

region of Nab2p (possibly in cooperation with Zf3). Subsequently, Nab2p TZF<sub>12</sub> binds weakly to the tail, protecting a sufficiently long section of RNA to later accommodate a second Nab2p Zf5–7 unit. Once released from RNA, Nab2p TZF<sub>12</sub> may participate in key protein-protein interactions and possibly in self-recognition as described for PABPN1, the mammalian protein that regulates poly(A) length (Eckmann et al., 2011; Kühn et al., 2009). Supporting this model, we obtained indirect experimental evidence suggesting that self-recognition ability resides in Nab2p TZF<sub>12</sub>. The <sup>1</sup>H-<sup>15</sup>N HSQC spectra of aged Nab2p Zf1–4 samples systematically lose the cross-peaks corresponding to TZF<sub>12</sub> (Figure S13B), probably because of aggregation (consistent with light scattering at 330 nm in the UV absorption spectra). This same effect is observed when samples are overconcentrated during purification or subjected to several freeze-thaw cycles. Whether self-association is biologically relevant to the Nab2p-mediated control of poly(A) tail length remains to be determined. However, our study provides tools, such as the N284R/E285F mutant, that alter the RNA binding mode of Nab2p TZF<sub>12</sub> and may be highly useful in testing these hypotheses in the future.

## EXPERIMENTAL PROCEDURES

## DNA Cloning

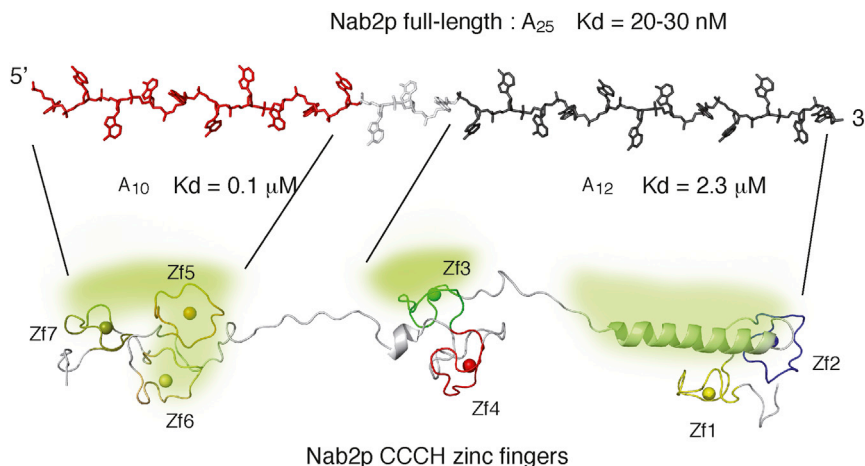
Nab2p constructs generated in the study are listed in [Table S3](#). DNA fragments were obtained by PCR from *Saccharomyces cerevisiae* genomic DNA (Novagen) using KOD DNA polymerase (Novagen) and DNA oligonucleotides (Integrated DNA Technologies; [Table S3](#)). Purified fragments were digested with restriction enzymes and subcloned into a modified version of pET28 (Novagen), which contains an N-terminal GST fusion and a TEV protease site. Nab2p Zf1–4 mutants ([Table S3](#)) were generated using a commercial kit (QuikChange, Stratagene) following the manufacturer's instructions. The entire plasmid was amplified by PCR with two DNA oligonucleotides (Integrated DNA Technologies) that incorporated the desired mutation, and the parental DNA was digested with Dnp1 prior to transformation in *E. coli* competent cells. Sequences of all clones were verified by DNA sequencing (Macrogen and Seugen).

## Protein Expression and Purification

Wild-type and mutant proteins were expressed in *E. coli* BL21(DE3) cells (Novagen), in either LB medium (for natural isotopic abundance) or in K-MOPS minimal media (Neidhardt et al., 1974) supplemented with  $^{15}\text{NH}_4\text{Cl}$  (1 g/l) and/or  $^{13}\text{C}$ -glucose (4 g/l) and 30  $\mu\text{g/l}$  of kanamycin (Sigma-Aldrich). The cultures were grown at  $37^\circ\text{C}$  until reaching an  $\text{OD}_{600}$  of 0.3 and were then supplemented with  $\text{ZnCl}_2$  (at 10  $\mu\text{M}$  final), cooled to  $20^\circ\text{C}$  and grown further until  $\text{OD}_{600}$  of 0.6 prior to overnight induction with 0.5 mM IPTG (Sigma-Aldrich). Cells were harvested  $\sim 20$  hr postinduction by centrifugation and directly processed or stored at  $-20^\circ\text{C}$ . Cell pellets were resuspended in 50 ml of PBS buffer (140 mM NaCl, 2.7 mM KCl, 10 mM  $\text{Na}_2\text{HPO}_4$ , 1.8 mM  $\text{KH}_2\text{PO}_4$  [pH 7.4]) plus one tablet of EDTA-free protease inhibitors (Roche) and were disrupted by French-press or sonication. The soluble fraction was separated by ultracentrifugation (30 min at  $40,000 \times g$ ) and rocked on ice for 1 hr after the addition of 1 volume of Glutathione Sepharose 4 Fast Flow (GE Healthcare) to 5 volumes of the soluble fraction. The unbound fraction was filtered out in a gravity column, the resin washed with 5 volumes of PBS, and fusion proteins eluted with buffer (50 mM Tris-HCl, 10 mM reduced

(combined  $^1\text{H}$  and  $^{15}\text{N}$  shifts) are marked with arrows. (C) Enlargement of the region corresponding to the Asn and Gln  $\text{NH}_2$  correlations (NH cross-peaks are shown in pale colors for clarity). Side chains showing significant perturbations are marked with arrows.

See also [Figures S9](#) and [S10](#).



**Figure 8. Schematic Representation of the Polyadenosine RNA Binding Mode of Nab2p**

The molecular model was constructed from the structures of Nab2p TZF<sub>12</sub>, TZF<sub>34</sub> (present study), and Nab2p Zf5–7 (PDB: 2LHN; Brockmann et al., 2012). Segments connecting these three structures were built up with Pymol in extended conformation. The RNA A<sub>25</sub>, which represents the molecular target of Nab2p Zf1–7 (see references in text), was constructed using a similar approach and is represented in the same scale as the protein to allow comparison of their sizes. See also Figure S13.

glutathione [pH 8.0]). Elutes were dialyzed (20 mM Tris-HCl [pH 8.0]) and simultaneously cleaved with 1 ml of home-made TEV protease (~1 mg/ml) overnight at 4°C. Upon complete digestion (confirmed by SDS-PAGE), samples were loaded into two 5 ml SP and mono Q columns (GE Healthcare) arranged in series. Major “impurities,” GST tag and TEV protease bound to the cation-exchange column (first in series), while Nab2p constructs bound to the mono Q column (second in series), with the exception of Nab2p Zf3–4, which did not bind to either of the columns and emerged in the flowthrough. Proteins were eluted with a linear salt gradient (0–1 M NaCl) in a buffer containing 20 mM Tris-HCl (pH 8.0) and 1 mM DTT. Finally, samples were purified by gel filtration chromatography (Superdex 75, GE Healthcare). Sample purity and homogeneity were determined by SDS-PAGE and mass spectrometry.

#### NMR Spectroscopy

Nab2p samples were prepared in 20 mM potassium phosphate (pH 6.5) aqueous buffer (10% D<sub>2</sub>O) at concentrations of 100–500 μM (10 μM DSS for internal reference). NMR experiments were recorded at 25°C in cryoprobe-equipped Bruker AV600 and AV800 spectrometers. Spectral assignment of Nab2p Zf1–2 and Zf3–4 constructs was performed with double-labeled (<sup>15</sup>N/<sup>13</sup>C) samples using standard triple resonance 3D data sets (HNCO, HNCA, HN(CO)CA, CBCA(CO)NH, and CBCANH; Sattler et al., 1999). Nab2p Zf1–4 spectrum was assigned by comparison with Zf1–2 and Zf3–4 and confirmed with an HNCA, CBCA(CO)NH pair. Side-chain assignment of Nab2p Zf1–2 and Zf3–4 constructs was completed in 3D HCCH-TOCSY experiments (Kay et al., 1993), and the HNHA experiment (Vuister and Bax, 1993) was recorded to obtain <sup>3</sup>J<sub>HN-Hα</sub> restraints. NOE-derived distance restraints for Nab2p Zf1–2 and Zf3–4 were extracted from 2D NOESY experiments (80 ms mixing time) in water and deuterium oxide solvents. Finally, <sup>15</sup>N {<sup>1</sup>H} NOE was computed from the signal ratios of two experiments recorded interleaved with and without saturation in a 600 MHz spectrometer (Farrow et al., 1995). The reference experiment was run with a recycling delay of 10 s, while in the experiment with saturation a train of 120° proton pulses separated by 500 ms was applied during this period. All NMR spectra were processed with NMRPipe (Delaglio et al., 1995) and Topspin (Bruker) and analyzed with CcpNmr Analysis (Vranken et al., 2005) software.

#### Structure Calculation

Nab2p TZF<sub>12</sub> and TZF<sub>34</sub> structures were calculated from NOE-derived distance constraints and  $\phi$  angle restraints (derived from <sup>3</sup>J<sub>HN-Hα</sub>; Table 1). The coordination of the zinc center was determined during preliminary structure calculations with the program CYANA 2.1 (Güntert et al., 1997) using only experimental restraints, and was corroborated spectroscopically (Figures S4–S5). Next, explicit zinc atoms (two per tandem) were included in the definition of the system by linking them to the end of the polypeptide chain. A new set of distance restraints (Hammarström et al., 1996) was included to maintain the tetrahedral geometry of the zinc centers (Supplemental Experimental Procedures) and structures were calculated with CYANA 2.1. Initial coordinates of 50 conformers were randomly generated and subjected to the restrained simu-

lated annealing protocol. The 20 lowest target function conformers with no distance violations greater than 0.2 Å and no angle violations greater than 5.0° were selected for further refinement. In

this last step, the structures were subjected to restrained molecular dynamics simulation with AMBER 9.0 (Case et al., 2006) at 0 K using implicit water solvent (igb = 2). The tetrahedral geometry of the zinc site was maintained by a set of restraints (see Supplemental Experimental Procedures). The chirality of the zinc centers (as defined by Berg, 1988) was determined using a matrix-based algorithm (Cieplak and Wisniewski, 2001). The final structures of Nab2p TZF<sub>12</sub> (PDB: 3ZJ1) and Nab2p TZF<sub>34</sub> (PDB: 3ZJ2) were analyzed and represented using the molecular graphic packages MOLMOL (Koradi et al., 1996) and PyMOL v0.98 (DeLano Scientific).

Peptides corresponding to the Nab2p zinc fingers were chemically synthesized (Caslo Laboratory). The apo-peptides were dissolved at a concentration of 1 mM in 20 mM potassium phosphate buffer (pH 6.5) and titrated with equimolar amounts of ZnCl<sub>2</sub>. Changes in the molecules were monitored by 1D proton NMR. Two-dimensional NOESY and TOCSY spectra were acquired for zinc fingers exhibiting nonaggregating behavior and assigned using classic NMR protocols. The NMR structures of Nab2p Zf2 and Zf3 peptides were calculated from NOE-derived distance restraints (Table S1). The coordination of the histidines relative to the zinc was established in preliminary calculation runs in CYANA. The 20 structures of NMR ensembles were generated with this program using the same protocol as used for the tandems and later subjected to energy minimization with AMBER (statistics in Table S1).

#### NMR Titrations

Nab2p Zf1–4 <sup>15</sup>N and <sup>15</sup>N/<sup>13</sup>C samples at concentrations of 100–250 μM were titrated with 1 mM stock solutions of A<sub>10</sub> and A<sub>12</sub> RNA oligonucleotides (Integrated DNA Technologies); both components were prepared in the same buffer conditions (20 mM potassium phosphate [pH 6.5], 10% D<sub>2</sub>O). Changes in the <sup>1</sup>H-<sup>15</sup>N HSQC spectrum of the protein were monitored up to protein:RNA ratios of approximately 1:1.5. Three-dimensional HNCA and CBCA(CO)NH experiments were performed using a Bruker AV800 NMR spectrometer for the protein/RNA complex formed with the double-labeled sample and used to partially reassign the NMR spectra of the RNA-bound protein.

#### Fluorescence Anisotropy Binding Titrations and Competition Assays

Fluorescence anisotropy-based binding titrations were conducted using a PC1 photon-counting steady-state ISS spectrofluorometer set at 20°C. For the measurements, quartz cuvettes of 3 × 3 mm path length (Starna or Hellma) were used. Excitation and emission wavelengths of 495 and 520 nm, respectively, were employed. Each sample was repeatedly measured and the anisotropy values reported represent the average of the last 8–10 values obtained after reaching equilibrium. The concentration of 3'-end fluorescein-labeled RNA (Integrated DNA Technologies) in the titrations was always 22 nM and the buffer used consisted of 25 mM Tris HCl, 100 mM NaCl, 0.1 mM DTT (pH 8.1). Competition experiments were conducted under equivalent conditions after adding 20 μM of unlabeled RNA. No changes were observed in the fluorescence intensity measured for the labeled RNA during the course of the titrations. The binding isotherms correspond to the mean of three independent experiments.

Analysis of the binding isotherms was performed using BIOEQS software (Rosales and Royer, 2008; Royer et al., 1990), which determines the free energies of formation of the postulated complexes from their individual elements by means of a numerical solver engine. The errors in the free energies retrieved were calculated by rigorous confidence limit testing at the 67% level. In this way, uncertainties arising from parameter correlations were taken into account. A model based on a single (1:1) RNA/protein complex was fit to the isotherms recovered from titrations of fluorescently labeled RNA with protein. For the competition titrations, two (1:1) complexes were postulated, consisting of protein and labeled or unlabeled RNA, respectively. In the analysis, the free energy of formation of the protein/labeled RNA complex and its anisotropy were fixed to the values obtained in the absence of competitor RNA.

## ACCESSION NUMBERS

The PDB accession numbers for the structures reported in this paper are 3ZJ1 and 3ZJ2. The BMRB accession numbers for the NMR chemical shifts are 18955 and 18956.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, 13 figures, five tables, and two 3D molecular models and can be found with this article online at <http://dx.doi.org/10.1016/j.str.2013.07.019>.

## ACKNOWLEDGMENTS

We wish to thank Diana Velázquez for technical assistance. This work was supported by a CSIC Special Intramural project (200880I088), a MINECO project (CTQ2011-26665), and a “Comunidad Autónoma de Madrid” project (CCG07-CSIC GEN2333). S.M.L. holds an undergraduate contract from the “Consejería de Educación de la Comunidad Autónoma de Madrid” and the European Social Fund. Y.M. held a predoctoral “Junta de Ampliación de Estudios (JAE)” contract from the CSIC.

Received: January 18, 2013

Revised: July 22, 2013

Accepted: July 23, 2013

Published: August 29, 2013

## REFERENCES

- Anderson, J.T., Wilson, S.M., Datar, K.V., and Swanson, M.S. (1993). NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability. *Mol. Cell. Biol.* 13, 2730–2741.
- Berg, J.M. (1988). Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. *Proc. Natl. Acad. Sci. USA* 85, 99–102.
- Brockmann, C., Soucek, S., Kuhlmann, S.I., Mills-Lujan, K., Kelly, S.M., Yang, J.C., Iglesias, N., Stutz, F., Corbett, A.H., Neuhaus, D., and Stewart, M. (2012). Structural basis for polyadenosine-RNA binding by Nab2 Zn fingers and its function in mRNA nuclear export. *Structure* 20, 1007–1018.
- Case, D.A., Darden, T.A., Cheatham, T.E.I., Simmerling, C.L., Wang, J., Duke, R.E., Luo, R., Merz, K.M., Pearlman, D.A., Crowley, M., et al. (2006). AMBER 9.0 (San Francisco: University of California, San Francisco).
- Cieplak, T., and Wisniewski, J.L. (2001). A new effective algorithm for the unambiguous identification of the stereochemical characteristics of compounds during their registration in databases. *Molecules* 6, 915–926.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277–293.
- Deo, R.C., Bonanno, J.B., Sonenberg, N., and Burley, S.K. (1999). Recognition of polyadenylated RNA by the poly(A)-binding protein. *Cell* 98, 835–845.
- Dheur, S., Nykamp, K.R., Viphacone, N., Swanson, M.S., and Minvielle-Sebastia, L. (2005). Yeast mRNA Poly(A) tail length control can be reconstituted in vitro in the absence of Pab1p-dependent Poly(A) nuclease activity. *J. Biol. Chem.* 280, 24532–24538.
- Eckmann, C.R., Rammelt, C., and Wahle, E. (2011). Control of poly(A) tail length. *WIREs RNA* 2, 348–361.
- Farrow, N.A., Zhang, O., Szabo, A., Torchia, D.A., and Kay, L.E. (1995). Spectral density function mapping using <sup>15</sup>N relaxation data exclusively. *J. Biomol. NMR* 6, 153–162.
- Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* 273, 283–298.
- Hammarström, A., Berndt, K.D., Sillard, R., Adermann, K., and Otting, G. (1996). Solution structure of a naturally-occurring zinc-peptide complex demonstrates that the N-terminal zinc-binding module of the Lasp-1 LIM domain is an independent folding unit. *Biochemistry* 35, 12723–12732.
- Hector, R.E., Nykamp, K.R., Dheur, S., Anderson, J.T., Non, P.J., Urbinati, C.R., Wilson, S.M., Minvielle-Sebastia, L., and Swanson, M.S. (2002). Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. *EMBO J.* 21, 1800–1810.
- Hudson, B.P., Martinez-Yamout, M.A., Dyson, H.J., and Wright, P.E. (2004). Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. *Nat. Struct. Mol. Biol.* 11, 257–264.
- Kay, L.E., Xu, G.Y., Singer, A.U., Muhandiram, D.R., and Forman-Kay, J.D. (1993). A gradient-enhanced HCCH-TOCSY experiment for recording side-chain <sup>1</sup>H and <sup>13</sup>C correlations in H<sub>2</sub>O samples of proteins. *J. Magn. Reson. B.* 101, 333–337.
- Kelly, S.M., Pabit, S.A., Kitchen, C.M., Guo, P., Marfatia, K.A., Murphy, T.J., Corbett, A.H., and Berland, K.M. (2007). Recognition of polyadenosine RNA by zinc finger proteins. *Proc. Natl. Acad. Sci. USA* 104, 12306–12311.
- Kelly, S.M., Leung, S.W., Apponi, L.H., Bramley, A.M., Tran, E.J., Chekanova, J.A., Wente, S.R., and Corbett, A.H. (2010). Recognition of polyadenosine RNA by the zinc finger domain of nuclear poly(A) RNA-binding protein 2 (Nab2) is required for correct mRNA 3'-end formation. *J. Biol. Chem.* 285, 26022–26032.
- Koradi, R., Billeter, M., and Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* 14, 51–55.
- Kornhaber, G.J., Snyder, D., Moseley, H.N., and Montelione, G.T. (2006). Identification of zinc-ligated cysteine residues based on <sup>13</sup>Cα and <sup>13</sup>Cβ chemical shift data. *J. Biomol. NMR* 34, 259–269.
- Kühn, U., and Wahle, E. (2004). Structure and function of poly(A) binding proteins. *Biochim. Biophys. Acta* 1678, 67–84.
- Kühn, U., Gündel, M., Knoth, A., Kerwitz, Y., Rüdél, S., and Wahle, E. (2009). Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. *J. Biol. Chem.* 284, 22803–22814.
- Lebreton, A., and Séraphin, B. (2008). Exosome-mediated quality control: substrate recruitment and molecular activity. *Biochim. Biophys. Acta* 1779, 558–565.
- Legge, G.B., Martinez-Yamout, M.A., Hambly, D.M., Trinh, T., Lee, B.M., Dyson, H.J., and Wright, P.E. (2004). ZZ domain of CBP: an unusual zinc finger fold in a protein interaction module. *J. Mol. Biol.* 343, 1081–1093.
- Lemay, J.F., D'Amours, A., Lemieux, C., Lackner, D.H., St-Sauveur, V.G., Bähler, J., and Bachand, F. (2010). The nuclear poly(A)-binding protein interacts with the exosome to promote synthesis of noncoding small nucleolar RNAs. *Mol. Cell* 37, 34–45.
- Marfatia, K.A., Crafton, E.B., Green, D.M., and Corbett, A.H. (2003). Domain analysis of the *Saccharomyces cerevisiae* heterogeneous nuclear ribonucleoprotein, Nab2p. Dissecting the requirements for Nab2p-facilitated poly(A) RNA export. *J. Biol. Chem.* 278, 6731–6740.
- Neidhardt, F.C., Bloch, P.L., and Smith, D.F. (1974). Culture medium for enterobacteria. *J. Bacteriol.* 119, 736–747.
- Proudfoot, N., and O'Sullivan, J. (2002). Polyadenylation: a tail of two complexes. *Curr. Biol.* 12, R855–R857.



- Rosales, T., and Royer, C.A. (2008). A graphical user interface for BIOEQS: a program for simulating and analyzing complex biomolecular interactions. *Anal. Biochem.* **381**, 270–272.
- Royer, C.A., Smith, W.R., and Beechem, J.M. (1990). Analysis of binding in macromolecular complexes: a generalized numerical approach. *Anal. Biochem.* **191**, 287–294.
- Sattler, M., Schleucher, J., and Griesinger, C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 93–158.
- Schmid, M., Poulsen, M.B., Olszewski, P., Pelechano, V., Saguez, C., Gupta, I., Steinmetz, L.M., Moore, C., and Jensen, T.H. (2012). Rrp6p controls mRNA poly(A) tail length and its decoration with poly(A) binding proteins. *Mol. Cell* **47**, 267–280.
- Soucek, S., Corbett, A.H., and Fasken, M.B. (2012). The long and the short of it: the role of the zinc finger polyadenosine RNA binding protein, Nab2, in control of poly(A) tail length. *Biochim. Biophys. Acta* **1819**, 546–554.
- Teplova, M., and Patel, D.J. (2008). Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1. *Nat. Struct. Mol. Biol.* **15**, 1343–1351.
- Tran, E.J., Zhou, Y., Corbett, A.H., and Wentz, S.R. (2007). The DEAD-box protein Dbp5 controls mRNA export by triggering specific RNA:protein remodeling events. *Mol. Cell* **28**, 850–859.
- Viphakone, N., Voisinet-Hakil, F., and Minvielle-Sebastia, L. (2008). Molecular dissection of mRNA poly(A) tail length control in yeast. *Nucleic Acids Res.* **36**, 2418–2433.
- Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696.
- Vuister, G.W., and Bax, A. (1993). Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNH.alpha.) coupling constants in <sup>15</sup>N-enriched proteins. *J. Am. Chem. Soc.* **115**, 7772–7777.
- Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids* (New York: John Wiley & Sons).
- Zhou, P., Tian, F., Lv, F., and Shang, Z. (2009). Geometric characteristics of hydrogen bonds involving sulfur atoms in proteins. *Proteins* **76**, 151–163.