Cleavage and polyadenylation specificity factor 30: An RNA-binding zinc-finger protein with an unexpected 2Fe–2S cluster

Geoffrey D. Shimberg[a,1], Jamie L. Michalek[a,1], Abdulafeez A. Oluyadi[a], Andria V. Rodrigues[b], Beth E. Zucconi[c], Heather M. Neu[a], Shanchari Ghosh[a], Kanisha Sureschandra[a], Gerald M. Wilson[c], Timothy L. Stemmler[b], and Sarah L. J. Michel[a,2]

[a]Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD 21201; [b]Department of Pharmaceutical Sciences, Wayne State University, Detroit, MI 48201; and [c]Department of Biochemistry and Molecular Biology, School of Medicine, University of Maryland, Baltimore, MD 21201

Edited by Jacqueline K. Barton, California Institute of Technology, Pasadena, CA, and approved March 1, 2016 (received for review September 8, 2015)

Cleavage and polyadenylation specificity factor 30 (CPSF30) is a key protein involved in pre-mRNA processing. CPSF30 contains five Cys_3His domains (annotated as "zinc-finger" domains). Using inductively coupled plasma mass spectrometry, X-ray absorption spectroscopy, and UV-visible spectroscopy, we report that CPSF30 is isolated with iron, in addition to zinc. Iron is present in CPSF30 as a 2Fe–2S cluster and uses one of the Cys_3His domains; 2Fe–2S clusters with a Cys_3His ligand set are rare and notably have also been identified in MitoNEET, a protein that was also annotated as a zinc finger. These findings support a role for iron in some zinc-finger proteins. Using electrophoretic mobility shift assays and fluorescence anisotropy, we report that CPSF30 selectively recognizes the AU-rich hexamer (AAUAAA) sequence present in pre-mRNA, providing the first molecular-based evidence to our knowledge for CPSF30/RNA binding. Removal of zinc, or both zinc and iron, abrogates binding, whereas removal of just iron significantly lessens binding. From these data we propose a model for RNA recognition that involves a metal-dependent cooperative binding mechanism.

zinc | iron–sulfur | RNA | protein | binding

Zinc-finger proteins (ZFs) are a large class of proteins that use zinc as structural cofactors (1–4). ZFs perform a variety of functions ranging from the modulation of gene expression through specific interactions with DNA or RNA to the control of signaling pathways via protein–protein interactions. The general feature that defines a ZF protein is the presence of one or more domains that contain a combination of four cysteine and/or histidine residues that serve as ligands for zinc. When zinc binds to these ligands, the domain adopts the structure necessary for function (1–4).

ZFs are typically identified by the presence of cysteine and histidine residues in regular repeats and are categorized into classes based upon the number of cysteine and histidine residues, and the spacing between the residues (1, 2). At least 14 distinct classes of ZFs have been identified to date. ZFs are highly abundant, with more than 3% of the proteins in the human genome annotated as ZFs, based upon their sequences (1, 5–9). In some cases, there are considerable in vitro and in vivo data that support the annotation of proteins as ZFs, whereas in other cases the only evidence that a protein is a ZF comes from its amino acid sequence. The best-studied class of ZFs comprises the "classical" ZFs. This class is composed of ZFs that contain a Cys_2His_2 domain ($CysX_{2-5}CysX_{12-13}HisX_{3-5}His$). Classical ZFs adopt an alpha-helical/antiparallel beta-sheet structure when zinc is coordinated and bind DNA in a sequence-specific manner (2, 4). The remaining classes of ZFs are collectively called "nonclassical" ZFs (1). One class of nonclassical ZFs is the Cys_3His class ($CysX_{7-9}CysX_{4-6}CysX_3His$). The first protein of this class to be identified was tristetraprolin, which contains two Cys_3His domains and regulates cytokine mRNAs via a specific ZF domain/RNA binding interaction (1). With the publication of genome sequences this domain has been found in a myriad of proteins. The National Center for Biotechnology Information (NCBI) conserved domain architecture tool identifies 404 distinct proteins (both hypothetical and experimentally validated) that contain this domain, and humans contain at least 60 (Fig. 1). As a class, these proteins are predicted to be involved in RNA regulation; however, the function(s) of most of these proteins have not yet been established (1, 2, 10, 11).

One important Cys_3His ZF protein is cleavage and polyadenylation specificity factor 30 (CPSF30) (2, 12). CPSF30 contains five Cys_3His domains. CPSF30 is part of a complex of proteins, collectively called CPSF, that are involved in the polyadenylation step of pre-mRNA processing (16). The other members of CPSF are CPSF160, CPSF73, CPSF100, Fip1, and Wdr33 (16). Polyadenylation is a 3′ end maturation step that all eukaryotic mRNAs (except histones) undergo (Fig. 2) (12). It involves endonucleolytic cleavage of the pre-mRNA followed by addition of a polyadenosine tail. Polyadenylation occurs at a specific region of the pre-mRNA called the polyadenylation cleavage site (PAS). The PAS consists of an upstream element with the conserved sequence AAUAAA (also called the AU-hexamer), a stretch of bases where cleavage occurs, after which a conserved GU-rich or U-rich sequence is present (usually between 40–60 nt after the cleavage site) (12, 13). CPSF73 is the endonuclease that cleaves the RNA; the roles of the other CPSF proteins are less clear (12, 13). Initially, CPSF160 was identified as the protein within the CPSF complex that recognizes the AU-hexamer (14–16); however, two recent studies using cell-based methods found that CPSF160 does not play this role (17, 18).

Significance

Cleavage and polyadenylation specificity factor 30 (CPSF30) is part of a complex of proteins, collectively called CPSF, that direct pre-mRNA processing. CSPF30 is classified as a "zinc-finger" protein because it contains five repeats of three cysteine and one histidine residue (CCCH) within its amino acid sequence, which are sequences known to bind zinc in other proteins. We report that CPSF30 contains an unexpected 2Fe–2S cluster, at one of the CCCH domains, in addition to zinc. We also demonstrate that CPSF30 selectively recognizes the AU-rich "hexamer" sequence present in all pre-mRNA via a cooperative, metal-dependent binding mechanism. These findings identify a functional role for CPSF30 in pre-mRNA recognition and identify a previously unidentified Fe–S cluster in this zinc-finger protein.

Author contributions: G.M.W., T.L.S., and S.L.J.M. designed research; G.D.S., J.L.M., A.A.O., A.V.R., B.E.Z., H.M.N., S.G., and K.S. performed research; G.D.S., J.L.M., A.A.O., A.V.R., B.E.Z., H.M.N., G.M.W., T.L.S., and S.L.J.M. analyzed data; and S.L.J.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]G.D.S. and J.L.M. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: smichel@rx.umaryland.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1517620113/-/DCSupplemental.

# of CCCH ZF Domains	Consensus Sequence	Frequency in Human Genome	Examples	Overall Function
ZF1	ZF1: C_{x8}-C_{x5}-C_{x3}-H	29	Putative ATP-dependent RNA helicase, Leukocyte receptor cluster member 9	RNA helicase
ZF1 ZF2	ZF1: C_{x8}-C_{x5}-C_{x3}-H ZF2: C_{x9}-C_{x5}-C_{x3}-H	9	tRNA-dihydrouridinesynthase, Tristetraprolin	tRNA synthesis, RNA splicing interactions, post-translational regulation
ZF1 ZF2 ZF3	ZF1: C_{x7}-C_{x5}-C_{x3}-H ZF2: $C_{x7.00}$-$C_{x4.86}$-$C_{x3.00}$-H ZF3: $C_{x7.86}$-$C_{x4.86}$-$C_{x3.00}$-H	7	E3 ubiquitin-protein ligase makorin-3	mRNA transport, poly(A) RNA binding
ZF1 ZF2 ZF3 ZF4	ZF1: $C_{x7.00}$-$C_{x5.11}$-$C_{x3.11}$-H ZF2: $C_{x7.22}$-$C_{x5.22}$-$C_{x3.00}$-H ZF3: $C_{x8.33}$-$C_{x5.33}$-$C_{x3.00}$-H ZF4: $C_{x7.78}$-$C_{x4.56}$-$C_{x3.00}$-H	9	Muscleblind-like protein 1, E3 ubiquitin-protein ligase	Alternative Splicing Regulation, anti-viral RNA degradation
ZF1 ZF2 ZF3 ZF4 ZF5	ZF1: $C_{x7.17}$-$C_{x5.50}$-$C_{x3.00}$-H ZF2: $C_{x7.83}$-$C_{x5.33}$-$C_{x3.17}$-H ZF3: $C_{x6.83}$-$C_{x4.33}$-$C_{x3.00}$-H ZF4: $C_{x7.67}$-$C_{x4.83}$-$C_{x3.00}$-H ZF5: $C_{x7.00}$-$C_{x4.83}$-$C_{x3.00}$-H	6	Cleavage and polyadenylation specificity factor subunit 4, Poly [ADP-ribose] polymerase 12	Pre-mRNA Regulation

Fig. 1. Survey of the CCCH domain containing proteins in *H. sapiens*.

Instead, CPSF30 and Wdr33 were identified as the proteins involved in AU-hexamer recognition (17, 18). These findings are intriguing in light of evidence that the H1N1 human influenza virus protein NS1A targets CPSF30 to obstruct cellular mRNA processing (19–21), suggesting that the link between NS1A and cellular mRNA processing is RNA recognition by CPSF30.

Given these cell-based results that CPSF30 is involved in recognition of the AU-hexamer of pre-mRNA along with our emerging understanding that CCCH-type ZFs are a general ZF motif involved in AU-rich RNA sequence recognition, we sought to determine whether CPFS30 directly recognizes the pre-mRNA AU-hexamer sequence via its CCCH domains by isolating CPSF30 and examining its RNA binding properties at the molecular level. CPSF30 contains five CCCH domains, and our hypothesis was that CPSF30 would bind five zinc ions at these CCCH domains and selectively recognize the AU-rich hexamer of pre-mRNA. To our surprise, CPSF30 was a reddish-colored protein upon isolation and purification, which suggested the presence of an iron cofactor. Here, we report that CPSF30 contains a 2Fe–2S site, with a CCCH ligand set, in addition to zinc. We also report that CPSF30 selectively recognizes the polyadenylation hexamer (AAUAAA) of pre-mRNA in a cooperative and metal-dependent manner. These findings are discussed in the context of CCCH "zinc" domains, iron, and recognition of AU-rich RNA sequences.

Results and Discussion

CCCH ZF Proteins. CPSF30 belongs to a class of proteins that are annotated as ZFs in genome databases. To determine how frequently this "CCCH ZF domain" occurs in eukaryotes a search using UniProt was performed (22). This resulted in 516 reviewed proteins and 25,610 total proteins when all organisms were considered. To provide further context, the search was narrowed to include only "CCCH ZF proteins" found in *Homo sapiens.* This led to the identification of 60 reviewed and 222 total proteins. The 60 reviewed proteins were then grouped based upon the number of CCCH domains present and a consensus sequence for each domain within the context of the number of domains was determined (Fig. 1) The proteins with CCCH ZF domains had between one and five domains, with the proteins with one domain being the most abundant (30 of the 66 proteins had only one domain). The organization of the domains (i.e., spacing between cysteine and histidine ligands) was generally invariant.

CPSF30 Contains Iron. A series of CPSF30 constructs that contained just the five CCCH ZF domains were prepared recombinantly and expressed under standard conditions. Unexpectedly, all of the constructs turned red upon protein overexpression, suggesting that iron was coordinating to the protein in addition to or in lieu of zinc (Fig. S1). The most soluble construct, which contained a maltose-binding fusion tag (MBP), was purified and used for subsequent studies. The UV-visible spectrum of the purified CPSF30 exhibited peaks at 420, 456, and 583 nm, which are indicative of iron (particularly iron–sulfur clusters), in addition to the expected peaks around 220–280 for protein backbone and aromatic amino acid peaks around 280 nm (Fig. 3) (23). Inductively coupled plasma mass spectrometry (ICP-MS) of CPSF30 was performed to measure the metal content. In a given preparation, both zinc and iron were observed, on average 3.78 ± 0.02 zinc and 0.51 ± 0.01 iron per protein. A ferrozine assay independently confirmed the presence of iron (24). The metal content of MBP alone was also measured by ICP-MS and a ferrozine assay. No metal was found to be present in these samples, ruling out MBP as a site for metal binding. The oligomerization state of CPSF30 was measured by size-exclusion chromatography and CPSF30 was found to be primarily a monomer (Fig. S2). Efforts to prepare just iron-loaded, zinc-loaded, and apo-CPSF30 were made. Incubation with EDTA produced the iron-loaded protein (0.29 ± 0.13 zinc and 0.45 ± 0.08 iron) and incubation with *o*-phenanthroline or dipyridyl in the presence of DTT produced apo-CPSF30 (0.11 ± 0.08 zinc and 0.10 ± 0.05 iron with *o*-phenanthroline and 0.10 ± 0.08 zinc and 0.15 ± 0.05 iron with dipyridyl). The zinc-loaded protein was obtained by overexpressing the protein in iron-deplete minimal media that was supplemented with zinc upon induction (3.71 ± 0.2 zinc and 0.07 ± 0.003 iron). Together, these studies revealed that

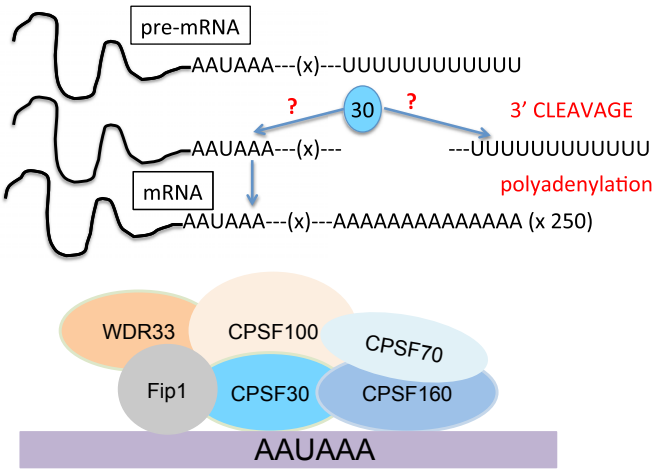

Fig. 2. Cartoon of pre-mRNA processing, with possible roles of CPSF30 highlighted in blue.

CPSF30's CCCH ZF domains have both zinc and iron cofactors, of which zinc is more readily chelated.

X-Ray Absorption Spectroscopy of the Iron Sites. X-ray absorption spectroscopy (XAS) was used to characterize the protein-bound Fe coordination geometry and ligand environment. The X-ray absorption near edge structure (XANES) portion of the XAS spectrum for Fe bound to CPSF30 is shown in Fig. 4*A*. A Fe pre-edge 1s-3d electronic transition peak, observed at ~7,112.4 eV with a corresponding peak area of 23.7×10^{-2} eV^2, is consistent with a four-coordinate tetrahedral ferric complex (25, 26). The Fe edge energy, determined from the first inflection point of the rising edge, occurs at 7,120 eV and is again consistent with a Fe(III) species characteristic of an oxidized 2Fe(III)–2S cluster (27). The Fe extended X-ray absorption fine structure (EXAFS) (Fig. 4*B*) was best simulated in the nearest-neighbor environment with *ca*. three S ligands at an average bond length of 2.26 Å and *ca*. one O/N ligand at 2.03 Å (Table S1). Long-range scattering includes an Fe–Fe vector at a bond length of 2.67 Å. In addition, multiple Fe–C interactions were observed at 3.17, 3.44, and 3.95 Å. These bond lengths represent average values obtained from two independent samples and are consistent with reported values for oxidized 2Fe–2S clusters bound by three Cys and one His residues (28). Thus, we speculate that iron is bound to CPSF30 as a 2Fe–2S cluster with a CCCH ligand set.

XAS of the Zinc Sites. XAS was also used to characterize protein-bound Zn coordination geometry and metrical parameters. The Zn edge inflection energy at 9,662.5 eV is consistent with Zn(II) (Fig. 5*A*). The postedge region of the spectrum shows features at both 9,665.7 eV and 9,671 eV, consistent with independent sulfur and oxygen/nitrogen ligand environments published for ZnN_1S_3 peptides (29, 30), suggesting similar coordination environments. The Zn EXAFS region (Fig. 5*B*) was best fit in the nearest-neighbor ligand environment with *ca*. one O/N ligand at 2.01 Å and *ca*. 2.5 S ligands at 2.31 Å (Table S1). The Fourier transform of the Zn EXAFS for CPSF30 is consistent with a ZnN_1S_3 peptide system published previously (30). Long-range carbon scattering was also observed at 3.07, 3.26, 3.43, and 4.03 Å. These data are consistent with Zn coordinated to CPSF30 in a tetrahedral CCCH ligand environment.

RNA Binding Studies: CPSF30 Binds to the AU-Rich Hexamer of α-Synuclein Pre-mRNA in a Cooperative Manner. CPSF30 is part of a complex of proteins, collectively called CPSF, that regulate pre-mRNA processing. CPSF30 has been proposed to be involved in recognition of the PAS present in pre-mRNA, because the signal contains an AU-rich sequence, which is a favored recognition signal for certain CCCH-type ZF proteins (1, 2). Recently, two laboratories simultaneously reported studies to identify the role of CPSF30 in a cellular setting. In one set of studies, Shi and coworkers (18) immunoprecipitated the entire CPSF complex then performed cross-linking with a pre-mRNA (viral SLV-4) that included the AU-hexamer and digested the complex to identify the specific protein or proteins that bound to the pre-mRNA. From these studies, along with subsequent iCLIP studies, CPSF30 was identified as one of two proteins that directly interact with the pre-mRNA at the AU-hexamer (the other is a newly identified protein, Wrd33). Schönemann et al. (17) reported that they could reconstitute CPSF160, CPSF30, hfip1, and WDR33 as a complex that bound to pre-mRNA (via a filter binding assay) and cross-linking/immunoprecipitation studies implicated CPSF30 and WDR33 as the proteins that directly bind to pre-mRNA. Together, these studies suggest that CPSF30 directly binds to pre-RNA. We sought to evaluate this interaction at the molecular level by measuring the affinity of isolated CPFS30 for a pre-mRNA target.

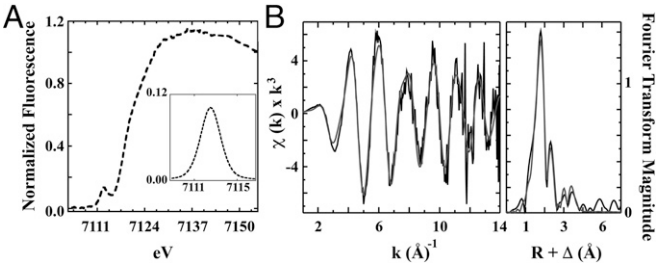

Fig. 4. (*A*) Fe XANES for CPSF30. (*Inset*) An expansion of the individual 1s-3d transition peak for CPSF30 Fe site. (*B*) Iron EXAFS and Fourier transform of EXAFS data for CPSF30 Fe site. (*Left*) Raw EXAFS data displayed in black and best fit in gray. (*Right*) Corresponding Fourier transform plot of raw EXAFS data in black and best fit in gray.

A 38-nt segment of the human alpha-synuclein transcript (NCBI reference NM_001146055.1) was used to study this interaction (31). In the sequence, the AU-rich hexamer (AAUAAA) is centrally located within the 38-nt RNA, where it is flanked by 16 nt on the 5′ and 3′ ends (from now on referred to as αSyn_{38}) (Table 1). Shorter RNA oligomers were also evaluated: αSyn_{30} and αSyn_{24}. αSyn_{30} and αSyn_{24} maintain the polyadenylation signal in the center of the oligomer and are both derived from αSyn_{38}, with either 12 or 9 nt flanking the polyadenylation signal, respectively (Table 1). A fragment of rabbit betaglobin (Rβ) mRNA was used as a negative control RNA oligomer because it does not contain the polyadenylation hexamer and is not enriched in adenosine and uridine nucleotides (Table 1).

Electrophoretic mobility shift assays (EMSA) were performed with the four αSyn oligomers and the Rβ control. CPSF30 formed complexes with all of the αSyn oligomers interrogated, as evidenced by shifting of the RNA oligomers (Fig. 6*A* and Fig. S3*A*), whereas no binding was observed for CPSF30 with Rβ or with the MBP tag with any of the RNA oligomers (Fig. S3*B*).

Fluorescence anisotropy (FA) was performed to determine the binding affinity of CPSF30 for the αSyn_{38}, αSyn_{30}, and αSyn_{24} oligomers. Binding was observed for all three αSyn RNA targets, but not for Rβ. Anisotropy values, corrected for the change in quantum yield, were plotted versus the concentration of CPSF30 protein titrated. The data were fit to two models: a 1:1 binding model and a cooperative binding model (32). The cooperative binding model gave the best fit, with $[P]_{1/2}$ values of 93.5 ± 2.7 nM, 115.0 ± 3.6 nM, and 143.8 ± 3.8 nM, with Hill coefficients of 1.67 ± 0.07, 1.63 ± 0.08, and 1.58 ± 0.07 for αSyn_{38}, αSyn_{30}, and αSyn_{24}, respectively (Fig. 6*B*). Binding was not observed when these RNA targets were titrated with either apo-CPSF30 or iron-loaded (zinc-deplete) CPSF30. The zinc-loaded (iron-deplete) CPSF30 exhibited significantly weakened affinity compared with the iron- and zinc-loaded CPSF30 ($[P]_{1/2}$ 570.5 ± 23.4) nM with αSyn_{38}. No binding was observed for MBP alone with any of the RNA targets.

To determine whether the AU-rich sequence is the site of RNA binding, titrations of CPSF30 with three altered RNA sequences were performed (Table 1). The sequences all lacked the AU-hexamer. In two cases the AU-hexamer was replaced with polyC (CCCCCC) or polyU (UUUUUU), and in the third the GU-rich sequence (UGUUUU) near the polyuridine site that has been proposed as an alternative target sequence for CPSF30 (33).

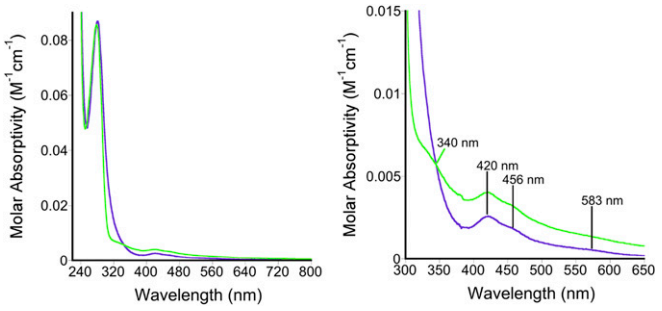

Fig. 3. Optical spectrum of CPSF30 (purple) and "apo"-CPSF30 (blue, after addition of EDTA) in 20 mM Tris, 100 mM NaCl, pH 8. (*Left*) full spectrum. (*Right*) Close-up between 300–650 nm.

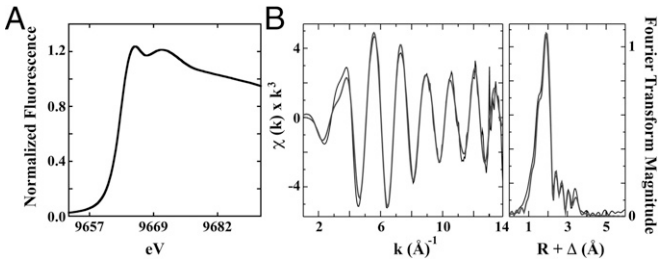

Fig. 5. (*A*) Zn XANES for CPSF30. (*B*) Zinc EXAFS and Fourier transform of EXAFS data for CPSF30 Zn site. (*Left*) Raw EXAFS data displayed in black and best fit in gray. (*Right*) Corresponding Fourier transform plot of raw EXAFS data in black and best fit in gray.

CPSF30 showed no binding to any of these sequences (in the Fe/Zn, Fe-only, Zn-only, or apo forms) (Fig. 6*C*).

Identification of Iron and Zinc Domains. The five CCCH domains of CPSF30 are highly homologous (Fig. S4*A*), and from sequence comparison it is not apparent which domain is loaded with the Fe–S cluster and which is loaded with zinc. A series of mutants in which each domain was modified at the coordinating cysteine and histidine ligands (CCCH to AAAA named ΔZF1, ΔZF2, ΔZF3, ΔZF4, and ΔZF5) were prepared to identify where iron and zinc bind. The mutant proteins were overexpressed and purified following the identical conditions used for WT-CPSF30. All of the mutants were reddish in color, and analysis by ICP-MS (Fig. S4*B*) revealed that they all were loaded with iron, but lost between one and two equivalents of Zn, with ΔZF2 losing the most Zn (2.4 equivalents) and ΔZF1 the least (1.3 equivalents). The affinities of these mutants for RNA (αSyn_{24}) was subsequently measured (Fig. S5) ΔZF4 retained comparable binding affinity to WT CPSF30, ΔZF1, ΔZF3, and ΔZF5 exhibited two- to threefold weaker affinities, and the ΔZF2 mutant did not bind to αSyn_{24} RNA with any appreciable affinity. Like WT-CPSF30, none of the mutants exhibited any affinity for polyC RNA. Taken together, these data support a model in which the sites of Fe and Zn binding are flexible, with iron loading occurring first. The data also suggest that ZF2 is critical for tight RNA binding, whereas the other ZFs seem to be less important.

Model of CPSF30/RNA Binding. The data for CPSF30/RNA binding were best fit to a cooperative binding model with a minimum stoichiometry of 2 CPSF30:1 RNA and an average Hill coefficient of 1.63 ± 0.07 (Fig. 6 *B* and *C*). CPSF30 alone is a monomer, as evidenced by gel filtration chromatography data (Fig. S2), yet exhibits positive cooperativity upon RNA recognition, perhaps indicating RNA-induced protein dimerization. Dimerization of proteins that bind to RNA is not unprecedented; for example, the Cys_3His ZF ZAP is a monomer in solution that then dimerizes upon binding to the ZAP-responsive RNA element (34). Other RNA-binding proteins that recognize their targets in a cooperative manner are the human zipcode-binding protein IMP-1, HIV-1 Rev protein, hordeiviral γb protein, human La protein, and tomato bushy stunt virus p33 protein (35–38). There are also other AU-rich RNA-binding proteins such as AUF-1 (p42), Hsp70, and HuR, which bind to their respective RNA partners in a cooperative manner (39–41). Moreover, a two-domain construct of CPSF30 has been crystallized bound to the NS1 influenza A virus protein, where it is present as a dimer (21).

Conclusions

Several important conclusions can be drawn from the work presented here: (*i*) CPSF30 is the protein within the CPSF complex that directly recognizes PAS RNA; (*ii*) CPSF30 contains an unexpected 2Fe–2S cluster, in addition to zinc; (*iii*) CPSF30 is always loaded with iron first, then zinc; and (*iv*) the sites of iron and zinc binding are flexible but high-affinity RNA binding requires that one Fe site and at least two Zn sites be occupied. The CPSF complex directs mRNA 3′ maturation via a mechanism of RNA recognition, cleavage, and polyadenylation (Fig. 2); however, the roles of the proteins that make up the CPSF complex are not clearly defined. Studies of the complex in a cellular setting have provided tantalizing support for CPSF30 as the protein that binds to the PAS signal. The work described here provides direct evidence that CPSF30 selectively recognizes the AU-rich hexamer of pre-RNA via its CCCH domains. Unexpectedly, the work also shows that CPSF30 contains a 2Fe–2S site with a CCCH ligand set, in addition to its predicted zinc sites.

These 2Fe-2S sites with CCCH ligand sets are rare; 2Fe-2S sites were first identified in the 1960s, with two principal types identified: ferredoxin (1962), which uses a CCCC ligand set, and rieske type (1964), which use a CCHH ligand set. However, it was not until the late 2000s that a 2Fe–2S site with a CCCH ligand set was identified (23). This site was found in a protein called mitoNEET, a mitochondrial protein that is a target of the type-2 diabetes drug pioglitazone (23). Remarkably, like CPSF30, mitoNEET was annotated as a ZF protein based on the presence of a CCCH domain (with spacing of $CXCX_9CX_3H$ compared with $CX_{7-9}CX_{4-5}CX_3H$ for CPSF30 and its homologs) and turned red upon protein expression and purification (42). MitoNEET contains a singular 2Fe–2S cluster with a CCCH ligand set and no zinc sites. The two homologs of MitoNEET, Miner1 (or NAF1) and Miner2, also contain between one and two 2Fe–2S clusters bound to CCCH sites (43–45). This suggests that the annotation of a protein as a ZF protein simply based upon its amino acid sequence (i.e., repeats of cysteine and histidine residues) may not always be correct and care must be taken in defining a ZF protein based just upon amino acid sequence. There is also evidence for a CCCH ligated 2Fe–2S cluster in proteins involved in iron–sulfur cluster assembly in *Escherichia coli* IscR and IscU (46, 47). In addition, in yeast, the Grx3/4/Fra2 signaling proteins have been shown to interact via a 2Fe–2S cluster that has a CCHX ligand set (48).

The biological significance of the 2Fe–2S cluster identified in CPSF30 is not yet known. Fe–S sites have been shown to play roles in electron transfer, oxygen sensing, iron sensing, substrate activation, and catalysis (49). Similarly, in mitoNEET there is not yet a definitive role for the 2Fe–2S site; however, there is evidence that it may be involved in trafficking iron via a redox sensing mechanism (50–54). Redox sensing may also be important for CPSF30: The CPSF30 yeast homolog, YtH1, responds to hypoxic stress (loss of O_2) by shuttling to the cytoplasm from the

Table 1. RNA oligomers tested

RNA oligomer	Length, nt	Sequence (5′–3′)
α-Syn_{38}	38	CCCAUCUCACUUUAAU**AAUAAA**AAUCAUGCUUAUAAGC
α-Syn_{30}	30	UCUCACUUUAAU**AAUAAA**AAUCAUGCUUAU
α-Syn_{24}	24	CACUUUAAU**AAUAAA**AAUCAUGCU
GU-$rich_{24}$	24	CCAGAAGUG**UGUUUU**GGUAUGCAC
Poly U_{24}	24	UUUUUUUUU**UUUUUU**UUUUUUUUU
Poly C_{24}	24	CACUUUAAU**CCCCCC**AAUCAUGCU
Rβ_{31}	31	UGGCCAAUGCCCUGGCUCACAAAUACCACUG

The AU hexamer region is shown in bold.

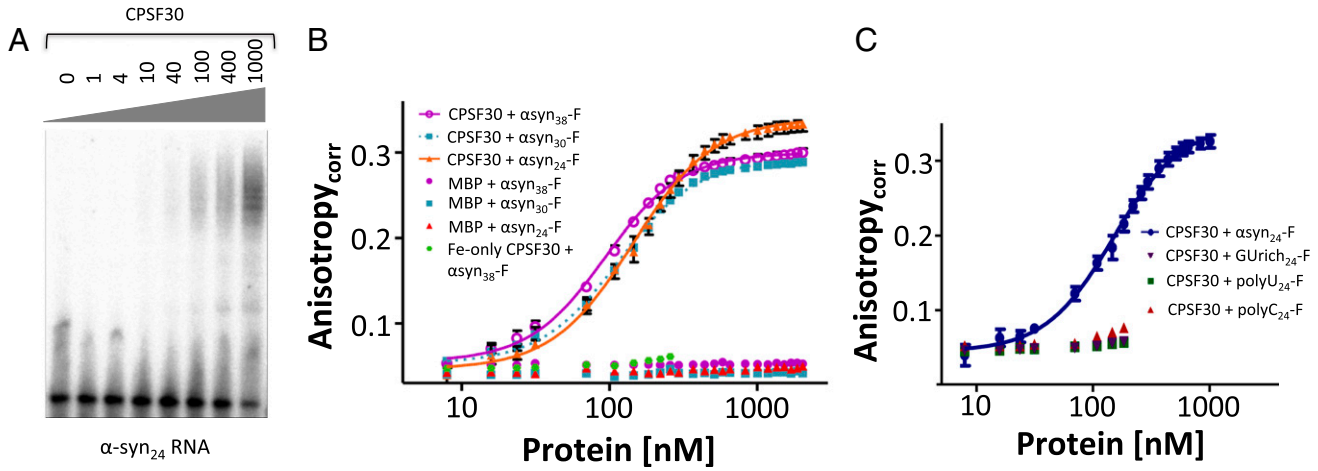

Fig. 6. (*A*) EMSA of α-syn$_{38}$ with CPSF30 (concentration: 0–1,000 nM). (*B*) FA monitored binding of CPSF30 with αSyn$_{38}$-Fl (open magenta circles), αSyn$_{30}$-Fl (teal squares), and αSyn$_{24}$-Fl (orange triangles). A titration of EDTA generated "apo"-CPSF30-5FE with αSyn$_{38}$-Fl (closed green circles) and MBP with αSyn$_{38}$-Fl (closed magenta circles), αSyn$_{30}$-Fl (light blue squares), and αSyn$_{24}$-Fl (red triangles). (*C*) FA monitored binding of CPSF30 with αSyn$_{24}$-Fl (blue circles), polyU$_{24}$ (green squares), GUrich$_{24}$ (purple triangles), and polyC$_{24}$ (red triangles). Data are fit to a cooperative binding equilibrium.

nucleus (55), and the Fe–S cluster may facilitate the protein localization (50–54). Additionally, the *Arabidopsis* homolog of CPSF30, which contains just three CCCH domains, has been shown to be involved in redox signaling, via its cysteine ligands (56). Another possible role for the 2Fe–2S cluster of CPSF30 is to regulate oligonucleotide binding a redox-dependent manner; Fe–S clusters are emerging as key cofactors in a range of regulatory proteins. In some proteins for which the Fe–S cluster plays a regulatory role, such as the base excision repair proteins, the Fe–S cluster's DNA binding affinity is dependent on the oxidation state of the Fe–S cluster and modulated by DNA charge transfer (57–62). In others, such as Aft2 (which also contains a structural zinc site), it is the presence or absence of the Fe–S cluster that drives DNA binding (63).

Taken together, our experimental data for CPSF30 reveal that CPSF30 contains both a 2Fe–2S site and a zinc site. Both sites are important for sequence-specific RNA binding, with the Zn site playing a larger role. CPSF30 selectively recognizes and binds to the polyadenylation AU-rich hexamer of α-synuclein pre-mRNA in a cooperative manner with high affinity. The location of the zinc and iron sites within the 5CCCH domains of CPSF30 seem to be flexible; however, iron is loaded before zinc. CPSF30 is a target of the human influenza virus, NS1A, which binds F2 and F3 of CPSF30 to obstruct cellular pre-mRNA processing, and we speculate that these two CCCH domains may be involved in direct pre-mRNA processing.

Materials and Methods

CPSF30 (33–170) was cloned into pMAL-c5e, expressed in BL21-DE3 cells at 37 °C for 3 h and purified via amylose and SP Sepharose chromatography. Metal content was measured via ICP-MS and oligomerization state by size-exclusion chromatography (Supadex 200 100/300 gel). Metals ions were removed by addition of EDTA, dipyridyl, or phenanthroline. XAS was used to identify metal ion sites and EMSA and FA to characterize RNA binding. Detailed experimental protocols are described in *Supporting Information*.

ACKNOWLEDGMENTS. This work was supported by National Science Foundation Grant CHE1306208 (to S.L.J.M.), American Heart Association/Friedreich's Ataxia Research Alliance Grant 12PRE11720005 (to A.V.R.), American Heart Association Grant 11PRE6900008 (to B.E.Z.), NIH Grants CA102428 (to G.M.W.) and DK068139 (to T.L.S.), and NIH training Grant T32GM066706-13 (to G.D.S.). Portions of this research were carried out at the Stanford Synchrotron Radiation Lightsource, a directorate of the Stanford Linear Accelerator Center, National Accelerator Laboratory, and an Office of Science User Facility operated for the US Department of Energy Office of Science by Stanford University. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by NIH–National Institute of General Medical Sciences (including Grant P41GM103393). Use of the National Synchrotron Light Source, Brookhaven National Laboratory, was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract DE-AC02-98CH10886.

1. Lee SJ, Michel SL (2014) Structural metal sites in nonclassical zinc finger proteins involved in transcriptional and translational regulation. *Acc Chem Res* 47(8):2643–2650.
2. Michalek JL, Besold AN, Michel SL (2011) Cysteine and histidine shuffling: Mixing and matching cysteine and histidine residues in zinc finger proteins to afford different folds and function. *Dalton Trans* 40(47):12619–12632.
3. Maret W (2012) New perspectives of zinc coordination environments in proteins. *J Inorg Biochem* 111:110–116.
4. Jantz D, Amann BT, Gatto GJ, Jr, Berg JM (2004) The design of functional DNA-binding proteins based on zinc finger domains. *Chem Rev* 104(2):789–799.
5. Bertini I, Decaria L, Rosato A (2010) The annotation of full zinc proteomes. *J Biol Inorg Chem* 15(7):1071–1078.
6. Laity JH, Lee BM, Wright PE (2001) Zinc finger proteins: New insights into structural and functional diversity. *Curr Opin Struct Biol* 11(1):39–46.
7. Andreini C, Banci L, Bertini I, Rosato A (2006) Zinc through the three domains of life. *J Proteome Res* 5(11):3173–3178.
8. Andreini C, Banci L, Bertini I, Rosato A (2006) Counting the zinc-proteins encoded in the human genome. *J Proteome Res* 5(1):196–201.
9. Decaria L, Bertini I, Williams RJ (2010) Zinc proteomes, phylogenetics and evolution. *Metallomics* 2(10):706–709.
10. Blackshear PJ, Perera L (2014) Phylogenetic distribution and evolution of the linked RNA-binding and NOT1-binding domains in the tristetraprolin family of tandem CCCH zinc finger proteins. *J Interferon Cytokine Res* 34(4):297–306.
11. Yang C, Huang S, Wang X, Gu Y (2015) Emerging roles of CCCH-type zinc finger proteins in destabilizing mRNA encoding inflammatory factors and regulating immune responses. *Crit Rev Eukaryot Gene Expr* 25(1):77–89.
12. Shi Y, Manley JL (2015) The end of the message: Multiple protein-RNA interactions define the mRNA polyadenylation site. *Genes Dev* 29(9):889–897.
13. Millevoi S, Vagner S (2010) Molecular mechanisms of eukaryotic pre-mRNA 3′ end processing regulation. *Nucleic Acids Res* 38(9):2757–2774.
14. Moore CL, Chen J, Whoriskey J (1988) Two proteins crosslinked to RNA containing the adenovirus L3 poly(A) site require the AAUAAA sequence for binding. *EMBO J* 7(10):3159–3169.
15. Gilmartin GM, Fleming ES, Oetjen J, Graveley BR (1995) CPSF recognition of an HIV-1 mRNA 3′-processing enhancer: multiple sequence contacts involved in poly(A) site definition. *Genes Dev* 9(1):72–83.
16. Keller W, Bienroth S, Lang KM, Christofori G (1991) Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3′ processing signal AAUAAA. *EMBO J* 10(13):4241–4249.
17. Schönemann L, et al. (2014) Reconstitution of CPSF active in polyadenylation: Recognition of the polyadenylation signal by WDR33. *Genes Dev* 28(21):2381–2393.
18. Chan SL, et al. (2014) CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3′ processing. *Genes Dev* 28(21):2370–2380.
19. Kuo RL, Krug RM (2009) Influenza a virus polymerase is an integral component of the CPSF30-NS1A protein complex in infected cells. *J Virol* 83(4):1611–1616.

20. Twu KY, Noah DL, Rao P, Kuo RL, Krug RM (2006) The CPSF30 binding site on the NS1A protein of influenza A virus is a potential antiviral target. *J Virol* 80(8): 3957–3965.
21. Das K, et al. (2008) Structural basis for suppression of a host antiviral response by influenza A virus. *Proc Natl Acad Sci USA* 105(35):13093–13098.
22. C U (2015) UniProt: A hub for protein information. *Nucleic Acids Res* 42(database issue):D204–D212.
23. Tamir S, et al. (2015) Structure-function analysis of NEET proteins uncovers their role as key regulators of iron and ROS homeostasis in health and disease. *Biochim Biophys Acta* 1853(6):1294–1315.
24. Riemer J, Hoepken HH, Czerwinska H, Robinson SR, Dringen R (2004) Colorimetric ferrozine-based assay for the quantitation of iron in cultured cells. *Anal Biochem* 331(2):370–375.
25. Schneider DJ, Roe AL, Mayer RJ, Que L, Jr (1984) Evidence for synergistic anion binding to iron in ovotransferrin complexes from resonance Raman and extended X-ray absorption fine structure analysis. *J Biol Chem* 259(15):9699–9703.
26. Westre TE, et al. (1997) A multiplet analysis of Fe K-Edge 1s → 3d pre-edge features of iron complexes. *J Am Chem Soc* 119(27):6297–6314.
27. Tsang HT, Batie CJ, Ballou DP, Penner-Hahn JE (1989) X-ray absorption spectroscopy of the [2Fe-2S] Rieske cluster in Pseudomonas cepacia phthalate dioxygenase. Determination of core dimensions and iron ligation. *Biochemistry* 28(18):7233–7240.
28. Sazinsky MH, et al. (2007) Characterization and structure of a Zn2+ and [2Fe-2S]-containing copper chaperone from Archaeoglobus fulgidus. *J Biol Chem* 282(35): 25950–25959.
29. Herbst RW, et al. (2010) Communication between the zinc and nickel sites in dimeric HypA: Metal recognition and pH sensing. *J Am Chem Soc* 132(30):10338–10351.
30. Clark-Baldwin K, et al. (1998) The limitations of X-ray absorption spectroscopy for determining the structure of zinc sites in proteins. When is a tetrathiolate not a tetrathiolate? *J Am Chem Soc* 120(33):8401–8409.
31. Rhinn H, et al. (2012) Alternative α-synuclein transcript usage as a convergent mechanism in Parkinson's disease pathology. *Nat Commun* 3:1084.
32. Wilson GM (2005) RNA folding and RNA-protein binding analyzed by fluorescence anisotropy and resonance engery transfer. *Reviews in Fluorescence*, eds Geddes CD, Lakowicz JR (Springer, New York), Vol 2, pp 223–243.
33. Barabino SM, Hübner W, Jenny A, Minvielle-Sebastia L, Keller W (1997) The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins. *Genes Dev* 11(13):1703–1716.
34. Chen S, et al. (2012) Structure of N-terminal domain of ZAP indicates how a zinc-finger protein recognizes complex RNA. *Nat Struct Mol Biol* 19(4):430–435.
35. Nielsen J, Kristensen MA, Willemoës M, Nielsen FC, Christiansen J (2004) Sequential dimerization of human zipcode-binding protein IMP1 on RNA: A cooperative mechanism providing RNP stability. *Nucleic Acids Res* 32(14):4368–4376.
36. Daugherty MD, Liu B, Frankel AD (2010) Structural basis for cooperative RNA binding and export complex assembly by HIV Rev. *Nat Struct Mol Biol* 17(11):1337–1342.
37. Rajendran KS, Nagy PD (2003) Characterization of the RNA-binding domains in the replicase proteins of tomato bushy stunt virus. *J Virol* 77(17):9244–9258.
38. Huelga SC, et al. (2012) Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Reports* 1(2):167–178.
39. Zucconi BE, et al. (2010) Alternatively expressed domains of AU-rich element RNA-binding protein 1 (AUF1) regulate RNA-binding affinity, RNA-induced protein oligomerization, and the local conformation of bound RNA ligands. *J Biol Chem* 285(50): 39127–39139.
40. Fialcowitz-White EJ, et al. (2007) Specific protein domains mediate cooperative assembly of HuR oligomers on AU-rich mRNA-destabilizing sequences. *J Biol Chem* 282(29):20948–20959.
41. Wilson GM, Sutphen K, Bolikal S, Chuang KY, Brewer G (2001) Thermodynamics and kinetics of Hsp70 association with A + U-rich mRNA-destabilizing sequences. *J Biol Chem* 276(48):44450–44456.
42. Wiley SE, Murphy AN, Ross SA, van der Geer P, Dixon JE (2007) MitoNEET is an iron-containing outer mitochondrial membrane protein that regulates oxidative capacity. *Proc Natl Acad Sci USA* 104(13):5318–5323.
43. Conlan AR, et al. (2009) Crystal structure of Miner1: The redox-active 2Fe-2S protein causative in Wolfram Syndrome 2. *J Mol Biol* 392(1):143–153.
44. Lin J, Zhang L, Lai S, Ye K (2011) Structure and molecular evolution of CDGSH iron-sulfur domains. *PLoS One* 6(9):e24790.
45. Baxter EL, Jennings PA, Onuchic JN (2012) Strand swapping regulates the iron-sulfur cluster in the diabetes drug target mitoNEET. *Proc Natl Acad Sci USA* 109(6):1955–1960.
46. Blanc B, Gerez C, Ollagnier de Choudens S (2015) Assembly of Fe/S proteins in bacterial systems: Biochemistry of the bacterial ISC system. *Biochim Biophys Acta* 1853(6): 1436–1447.
47. Kim JH, Bothe JR, Alderson TR, Markley JL (2015) Tangled web of interactions among proteins involved in iron-sulfur cluster assembly as unraveled by NMR, SAXS, chemical crosslinking, and functional studies. *Biochim Biophys Acta* 1853(6):1416–1428.
48. Li H, et al. (2009) The yeast iron regulatory proteins Grx3/4 and Fra2 form heterodimeric complexes containing a [2Fe-2S] cluster with cysteinyl and histidyl ligation. *Biochemistry* 48(40):9569–9581.
49. Rouault TA (2015) Iron-sulfur proteins hiding in plain sight. *Nat Chem Biol* 11(7): 442–445.
50. Bak DW, Elliott SJ (2014) Alternative FeS cluster ligands: Tuning redox potentials and chemistry. *Curr Opin Chem Biol* 19:50–58.
51. Netz DJ, Mascarenhas J, Stehling O, Pierik AJ, Lill R (2014) Maturation of cytosolic and nuclear iron-sulfur proteins. *Trends Cell Biol* 24(5):303–312.
52. Landry AP, Cheng Z, Ding H (2015) Reduction of mitochondrial protein mitoNEET [2Fe-2S] clusters by human glutathione reductase. *Free Radic Biol Med* 81:119–127.
53. Saouma CT, Pinney MM, Mayer JM (2014) Electron transfer and proton-coupled electron transfer reactivity and self-exchange of synthetic [2Fe-2S] complexes: Models for Rieske and mitoNEET clusters. *Inorg Chem* 53(6):3153–3161.
54. Landry AP, Ding H (2014) Redox control of human mitochondrial outer membrane protein MitoNEET [2Fe-2S] clusters by biological thiols and hydrogen peroxide. *J Biol Chem* 289(7):4307–4315.
55. Dastidar RG, et al. (2012) The nuclear localization of SWI/SNF proteins is subjected to oxygen regulation. *Cell Biosci* 2(1):30.
56. Chakrabarti M, Hunt AG (2015) CPSF30 at the interface of alternative polyadenylation and cellular signaling in plants. *Biomolecules* 5(2):1151–1168.
57. Grodick MA, Segal HM, Zwang TJ, Barton JK (2014) DNA-mediated signaling by proteins with 4Fe-4S clusters is necessary for genomic integrity. *J Am Chem Soc* 136(17):6470–6478.
58. Pheeney CG, Arnold AR, Grodick MA, Barton JK (2013) Multiplexed electrochemistry of DNA-bound metalloproteins. *J Am Chem Soc* 135(32):11869–11878.
59. Romano CA, Sontz PA, Barton JK (2011) Mutants of the base excision repair glycosylase, endonuclease III: DNA charge transport as a first step in lesion detection. *Biochemistry* 50(27):6133–6145.
60. Lee PE, Demple B, Barton JK (2009) DNA-mediated redox signaling for transcriptional activation of SoxR. *Proc Natl Acad Sci USA* 106(32):13164–13168.
61. Boal AK, et al. (2005) DNA-bound redox activity of DNA repair glycosylases containing [4Fe-4S] clusters. *Biochemistry* 44(23):8397–8407.
62. Boon EM, Livingston AL, Chmiel NH, David SS, Barton JK (2003) DNA-mediated charge transport for DNA repair. *Proc Natl Acad Sci USA* 100(22):12543–12547.
63. Poor CB, et al. (2014) Molecular mechanism and structure of the Saccharomyces cerevisiae iron regulator Aft2. *Proc Natl Acad Sci USA* 111(11):4043–4048.
64. Cook JD, et al. (2010) Molecular details of the yeast frataxin-Isu1 interaction during mitochondrial Fe-S cluster assembly. *Biochemistry* 49(40):8756–8765.
65. Randall CR, et al. (1995) X-ray absorption pre-edge studies of high-spin iron(II) complexes. *Inorg Chem* 34(5):1036–1039.
66. Bencze KZ, Kondapalli KC, Stemmler TL (2007) X-ray absorption spectroscopy. *Applications of Physical Methods to Inorganic and Bioinorganic Chemistry* (Wiley, Chichester, UK), pp 513–528.
67. Wang B, et al. (2003) Structure and ubiquitin interactions of the conserved zinc finger domain of Npl4. *J Biol Chem* 278(22):20225–20234.
68. Cotelesage JJ, Pushie MJ, Grochulski P, Pickering IJ, George GN (2012) Metalloprotein active site structure determination: Synergy between X-ray absorption spectroscopy and X-ray crystallography. *J Inorg Biochem* 115:127–137.