Structural basis of RNA binding by leucine zipper GCN4

Yaroslav Nikolaev^{1*} and Konstantin Pervushin^{1,2*}

¹Biozentrum of University Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

²School of Biological Sciences, Nanyang Technological University, Singapore 637551, Singapore

Received 18 August 2011; Accepted 19 February 2012

DOI: 10.1002/pro.2051

Published online 28 February 2012 proteinscience.org

Abstract: Recently, we showed that leucine zipper (LZ) motifs of basic leucine zipper (bZIP) transcription factors GCN4 and c-Jun are capable of catalyzing degradation of RNA (Nikolaev et al., PLoS ONE 2010; 5:e10765). This observation is intriguing given the tight regulation of RNA turnover control and the antiquity of bZIP transcription factors. To support further mechanistic studies, herein, we elucidated RNA binding interface of the GCN4 leucine zipper motif from yeast. Solution NMR experiments showed that the LZ-RNA interaction interface is located in the first two heptads of LZ moiety, and that only the dimeric (coiled coil) LZ conformation is capable of binding RNA. Site-directed mutagenesis of the LZ-GCN4 RNA binding interface showed that substrate binding is facilitated by lysine and arginine side chains, and that at least one nucleophilic residue is located in proximity to the RNA phosphate backbone. Further studies in the context of full-length bZIP factors are envisaged to address the biological relevance of LZ RNase activity.

Keywords: leucine zipper; GCN4; bZIP; coiled coil; ribonuclease; NMR; protein-RNA interaction

Introduction

In the recent years, it has become increasingly apparent that individual proteins may exhibit multiple distinct functions within cells.¹ Functional combinations such as metabolic enzyme–growth factor, proteasome component–replication factor and many others are known (reviewed in Ref. 2). New unexpected cases such as DNA-binding/transcription regulation activity of cytoplasmic proteins³ continue to arise.

Proteins featuring a bZIP motif are typically DNA-binding transcription factors which evolved as key regulators in a variety of processes, ranging from cell metabolism to tissue differentiation.⁴ These factors are widespread among eukarya, with only human genome encoding 53 proteins with unique

bZIP motifs.⁵ The leucine zipper (LZ) portion of bZIP proteins enables their dimerization, thereby playing essential structural role in establishment of combinatorial transcription networks.^{4,6}

Recently, we have shown that beyond oligomerization function LZ motifs from factors GCN4 (yeast) and c-Jun (human) exhibit ribonuclease activity.⁷ Catalytic activity was preserved within the full-length c-Jun, implying possible biological relevance of this activity. To facilitate mechanistic and *in vivo* studies of LZ-RNase phenomenon, here, we delineate structural details of RNA binding by the GCN4 LZ motif.

Solution NMR experiments show that binding interface is located in the first two *N*-terminal heptads of LZ moiety and only dimeric (coiled coil) LZ conformation is capable of binding RNA. Mutational analysis of the interaction site indicates that substrate binding is facilitated by cationic side chains, and that at least one nucleophilic sidechain is located near the phosphate backbone. Based on the topology of interactions, we hypothesize that catalytic properties may become manifested upon bZIP factor binding to its target DNA consensus.

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Swiss National Science Foundation; Grant number: 3100A0-118381.

*Correspondence to: Konstantin Pervushin, School of Biological Sciences, Nanyang Technological University, Singapore. E-mail: kpervushin@ntu.edu.sg or Yaroslav Nikolaev, E-mail: yaroslav@troty.com

Results

Catalytic conformation of LZ-GCN4

In the solution, dimeric (coiled coil) leucine zippers are at equilibrium with monomeric (x-form, or partially ordered monomeric) conformations.^{8,9} To elucidate which of the two conformations is responsible for the RNase activity, we sampled LZ-GCN4•RNA interactions under conditions where both dimeric and monomeric conformations are comparably populated (15 μ M LZ-GCN4, pH 4.0). In the acidic environment, the presence of two distinct sets of resonances establishes that the dimeric and monomeric conformations are in the slow conformational exchange on the NMR chemical shift timescale (Supporting Information Fig. S1). Concerted attenuation of the coiled coil resonances upon addition of the RNA substrate, together with the chemical shift perturbations at a number of residues, points to the presence of two consecutive equilibria—(i) an equilibrium between free coiled coil and its RNA-bound form and (ii) an equilibrium between the RNA-bound form and high-molecular weight aggregates. The first equilibrium is manifested in small shifts of coiled coil resonances with concomitant line broadening while the second equilibrium results in the attenuation of coiled coil resonances in the presence of RNA substrate without any detectable resonance shifts and line broadening (Supporting Information Fig. S1B). This conclusion is supported by the fact that further increase of the RNA concentration to 15–20 μ M leads to visible precipitation of the LZ-RNA complexes, revealing this background aggregation process. The linearly coupled reactions (Supporting Information Fig. S1B inset) explain the fact that resonances of the monomeric species are not attenuated in the presence of the substrate due to the absence of the direct contact with RNA. On the other side, the residues Q2, K30, and E34 remain visible on the background of the back-drop of resonances intensities in protein/RNA aggregates due to their reduced transverse relaxation rates. Similar aggregation behavior was observed in the spectra of LZ constructs interacting with the shorter RNA substrate (see Fig. 3 and Discussion of RNA binding of LZ-GCN4 mutants).

RNA binding site of LZ-GCN4

To address the specificity of complex formation and map the residues forming the RNA binding site, LZ-GCN4 interactions with 18-mer RNA were sampled in a series of [¹H-¹⁵N]-HSQC experiments under conditions of the catalytic activity assays (20 mM HEPES-D₁₈, 85 mM KCl, pH 7.1, 37°C). Additional experiments were performed at 12°C to evaluate the involvement of four N-terminal residues (L1-M4) in RNA binding. Resonances of these residues were not observed in [¹H-¹⁵N]-HSQC spectra at 37°C due to rapid exchange of their amide protons with water.

Chemical shift perturbations revealed the proximity of the RNA binding site to the N-terminus of the coiled coil structure, spanning two heptad repeats and involving five charged (R3, K5, E8, D9, K10, E12) and three polar (Q2, M4, Q6) residues [Fig. 1(A,C)]. The RNA substrate was hydrolyzed over time, resulting in a return of the perturbed resonances to their original spectral positions (Supporting Information Fig. S2).

Analysis of chemical shift perturbation data employing a two-state (LZ + RNA $\rightarrow$ [LZ•RNA]) complex formation model yielded an average LZ-GCN4•RNA equilibrium dissociation constant $K_d = 51$ μ M [Fig. 1(B)]. The fitting was performed using a 1:1 LZ:RNA stoichiometry, due to the presence of two identical binding sites within the LZ dimer. Individual per-residue dissociation constants varied from 41 ± 5 to 68 ± 14 μ M. Adjusted to the ambient temperature (25°C) the average dissociation constant is 61 μ M.

Absence of observable imino-proton signals in the 1D RNA spectrum under both neutral and acidic conditions (Supporting Information Fig. S3A-B) suggested that 18-mer RNA lacks a defined structural conformation in the solution. Imino resonances also did not appear upon addition of the protein (Supporting Information Fig. S3C), indicating that no dominant structure is formed upon RNA binding to LZ and/or the binding primarily occurs via the phosphate backbone (LZ residues do not interact with nucleobase moieties of RNA). At the same time, some of the sugar and/or base proton resonances of free RNA were perturbed in the presence of LZ, pointing to protein-RNA interactions and/or structural changes in the RNA. We expected that a shorter RNA could reduce the range of binding modes and lead to a single predominant LZ-RNA structural ensemble. Therefore a shorter 5-mer sequence (GCAGG) was designed taking into account nucleotide-cleaving preferences exhibited by LZ-GCN4 in the previous study (higher cleavage frequency after pyrimidines). As shown by LC-MS, LZ-GCN4 cleaved 5-mer RNA with a rate comparable to that of the 18-mer substrate [$k_2 = 119$ (5-mer) vs. 47 (18-mer) $M^{-1} \text{ min}^{-1}$] (Supporting Information Fig. S4). As expected 5-mer was cleaved only at a single position (GC-AGG) (Supporting Information Fig. S4), implying a single register of LZ-RNA complex formation. Nevertheless, no imino-proton resonances were observed in the presence of LZ peptides (Supporting Information Fig. S5), supporting the notion that RNA binding occurs primarily via the phosphate backbone. Under the conditions employed we were unable to obtain any intermolecular NOE data on either 5-mer or 18-mer LZ•RNA complexes, presumably due to the fast exchange kinetics between the free and the bound states of leucine zipper. Low affinity of LZ-RNA complexes and absence of intermolecular NOEs, suggests that further optimization

Figure 1. RNA-binding by LZ-GCN4. (A) ^1H - ^{15}N -HSQC spectra of LZ-GCN4 at peptide:RNA concentrations of 62:0 μM (blue), 62:35 μM (violet), 62:140 μM (magenta), and 62:421 μM (red). Reaction performed in 40 mM HEPES- D_{18} , 85 mM KCl, pH 7.1, spectra recorded at 12°C . Residue assignments of perturbed resonances are indicated. (B) Equilibrium dissociation constants (K_d) of LZ-RNA complex obtained by fitting H-N chemical shift perturbation data to the two-state bimolecular association rate equation. (C) Plot of maximum chemical shift perturbations observed at different residues of LZ-GCN4. Residues involved in RNA binding are shown in color. Coiled coil heptad positions of residues are indicated. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

of the RNA sequence and/or extension of protein construct towards full-length GCN4 is required to stabilize the LZ-RNA complex and determine its structure. Meanwhile, mutations in LZ-GCN4 were used to verify the mapped interaction interface.

Mutational analysis of LZ-GCN4 RNA binding interface

To corroborate the NMR-mapped RNA-interaction interface and establish a ground for characterization of the catalytic mechanism, six LZ-GCN4 mutants were produced. These mutations replaced the charged residues in the first two heptad repeats with alanine. In addition, a S16A mutant was created as S16 has shown minimal chemical shift perturbation in the presence of RNA in our current NMR analysis, and mutation of this residue lead to reduced catalytic activity of synthetic peptides in our previous study.⁷ A E12Q mutant was created as a control of mutations affecting structural stability.

CD analysis demonstrated that all mutant proteins retained their predominantly α -helical secondary structure [Fig. 2(A)], with only K5A and K10A exhibiting a slight decrease in the overall α -helical content, from 71.6% in the wild type to 52.5 and 59.6%, respectively [Fig. 2(A) inset]. R3A and E8A showed marginal decrease in the helicity (to 64.8 and 61% α -helix, respectively), while other mutants were indistinguishable from the wild-type protein.

Temperature-induced melting (T_m) curves, measured as the fractional ellipticity at 222 nm [Fig. 2(B)]

over the range 15 – 89°C , showed that K5A and E8A mutants exhibit an approximate 10°C decrease in mean T_m while others retained or even slightly increased their stability compared to the wild-type protein (e.g. E12A, E12Q, and S16A exhibited 5 – 7°C increase in T_m).

A comparison of [^1H - ^{15}N]-HSQC NMR spectra of the LZ-GCN4 wild type and mutants confirmed that all constructs exist as stable dimeric complexes under the conditions of ribonuclease assays [Fig. 2(E–H)]. An analysis of the heteronuclear NOE (Supporting Information Fig. S6), transverse relaxation (Supporting Information Fig. S7), and transverse relaxation dispersion (Supporting Information Fig. S8) experiments showed that overall peptide backbone dynamics was largely unaffected in the corresponding ns-ps and ms timescales.

Transverse relaxation data of the K10A mutant revealed increased relaxation rates for several stretches along the sequence (Supporting Information Fig. S7). This effect might be attributed to the disruption of the electrostatic network involving residues K10, E14, K18, and E23 which cover two central heptads of the coiled coil structure and include three intramolecular and two intermolecular ionic interactions [Fig. 2(C)].

Chemical shift perturbation data shows that the effect of sidechain replacement propagates across the network of ionic interactions stabilizing the coiled coil dimer. This is especially prominent in the case of R3A and K10A mutants, where alterations in the local environment are correlated with the boundaries of

Figure 2. Structural characterization of LZ-GCN4 mutants. (A) Equilibrium CD spectra of 30 μM peptides measured at 37°C in 10 mM PO₄, 85 mM KCl, pH 7.1. The average helical content is indicated in the inset. (B) Normalized temperature denaturation curves measured as ellipticity of peptide sample at 222 nm. Midpoint of melting transition is indicated in the inset. (C) Topology of four intra-helix and two inter-helix ionic interactions stabilizing LZ-GCN4 dimer. (D) ^1H and ^{15}N chemical shift perturbations observed upon mutation of particular residues within LZ-GCN4 sequence. Positions of ionic interactions within first two heptads is indicated by dotted lines under the CSP data. (E–H) [^1H , ^{15}N]-HSQC spectra of wild-type LZ-GCN4 and mutants, measured in the conditions of RNase assays (50 μM peptide, 20 mM HEPES, 85 mM KCl, pH 7.1, 37°C). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

K5-E8 and K10-E13-K18 salt-bridge assemblies [Fig. 2(C,D)]. Thus, NMR and CD data demonstrated that the effect of mutations was confined to the first two heptads of the LZ-GCN4 structure minimally interfering with stability of the coiled coil dimer.

To investigate the effect of mutations on the LZ-GCN4 RNA binding activity, a five-nucleotide RNA substrate (GCAGG) was employed. This molecule provided a simpler binding model than the original 18-mer substrate, and showed an increased affinity to LZ-GCN4 (18 μM vs. 50 μM). A NMR titration analysis showed that, as in the case of longer RNA substrate, the 5-mer GCAGG binds to the two *N*-terminal heptads of wild-type LZ-GCN4 [Fig. 3(A)]. The substitution of cationic sidechains (R3, K5, K10) with alanine reduced the RNA binding affinity, while the replacement of the anionic sidechains (E8, D9, E12) or serine (S16) increased the affinity [Fig. (3)]. The latter effect was especially prominent in the case of the E8A mutant. This peptide displayed the strongest chemical shift perturbations and formed aggregated complexes with RNA at concentrations as low as 100

μM [Fig. 3(B)]. In contrast, peptides with the substituted cationic sidechains (R3, K5, K10) remained visible in the NMR spectra even at high RNA concentrations [Fig. 3(B)— ^1H projections] due to the lower affinity to RNA. As in the case of LZ aggregates with 18-mer RNA, the C-terminal residues of LZ-GCN4 (V32–R35) in the complex with RNA retained narrow line-width and were visible in the NMR spectra. This is indicative of high local flexibility of the peptide chain, thus confirming that the C-terminus is not involved in the RNA complex formation.

Preservation of LZ-GCN4 RNA binding site within bZIP family

Based on the binding properties of LZ-GCN4 mutants, the 328 members of bZIP family (PROSITE: ID BZIP; AC PS50217) were analyzed for the conservation of the RNA binding site. Given reduced RNA binding activity upon removal of the cationic R3, K5, and K10 residues and increased affinity of the E8A mutant, three minimal active site patterns were identified.

These feature one anionic and two cationic residues, yielding a minimal catalytic site consisting of

catalytic base, catalytic acid and auxiliary phosphate-stabilization centers. In the case of LZ-GCN4, these patterns correspond to K5-E8-K10 (I), R3-E8-K10 (II), and R3-K5-E8 (III) residue combinations, respectively. Considering the importance of proper positioning of the catalytic sidechains relative to the coiled coil 3D topology, an additional leucine residue was introduced into the PROSITE pattern, anchoring the site alignment with respect to the *d*-position of the heptad repeat.

The search employing the described patterns yielded 52, 29, and 89 unique hits among 328 members of bZIP family, with 30, 24, and 17 having correct positioning of residues within the LZ moiety (Supporting Information Fig. S9). From all hits, 58 are nonredundant with 21 stemming from nonorthologous proteins and nine belonging to human bZIP factors.

Discussion

Catalytic degradation of RNA is an essential part of RNA turnover and decay, a major mechanism for regulation of gene expression, quality control of RNA biogenesis and antiviral defences.^{10–13} Therefore, if manifested *in vivo*, ribonuclease activity of leucine zipper motifs may have significant biological implications due to their antiquity and prevalence in transcription factors. To enable mechanistic and *in vivo* studies of LZ RNase activity, we characterized structural details of RNA binding by LZ motif of factor GCN4.

Active conformation

Our first aim was to elucidate which of the two LZ conformations (coiled coil dimer or partially structured monomer)⁸ are involved in RNA binding. NMR data showed that the RNA substrate interacts only with the dimeric conformation of the protein (Supporting Information Fig. S1). From biological standpoint, coupling of catalysis to a particular oligomeric state may provide a mechanism for sensing the concentration of the catalyst.¹⁴ In the case of oligomerizing transcription factors, this may provide an association of the enzymatic function not only with the concentration of polypeptides but rather with a specific phase of the transcription activation process likely manifested by oligomerization upon reaching the threshold concentrations of the protein factor.

RNA binding site

For initial experiments, an 18-mer RNA substrate (identical to the one used in our previous study⁷) was employed in order to characterize the LZ affinity towards RNA. NMR data showed that RNA binding occurs within two *N*-terminal heptad repeats of LZ-GCN4 (Fig. 1). The LZ binding interface involves several cationic residues with spacing allowing for the RNA stabilization through three sequential

In addition to cationic residues, LZ RNA-binding interface features a central anionic residue (Glu8) which may act as a catalytic base (nucleophile) upon initiation of catalysis. Together with neighboring Asp9 and Glu12 this forms an anionic cluster [Fig. 4(B)] potentially enhancing the nucleophilic properties of Glu8. Under the conditions employed, we were unable to obtain reliable intermolecular NOE data on LZ•RNA complex formation due to the fast exchange kinetics between free and RNA complexed states of leucine zipper. Therefore, mutations in LZ-GCN4 were used to verify the mapped interaction interface.

Chemical shift perturbations observed upon formation of the LZ-GCN4-RNA complex suggested that six charged and three polar residues may contribute to the catalysis (R3, K5, E8, D9, K10, E12 and Q2, M4, Q6, Fig. 1). CD analysis of the mutants demonstrated that all peptides retained predominantly α -helical structure [Fig. 3(A,B)], with weak destabilization in the K5A, K10A, and E8A mutants and stabilization in the E12A, E12Q, and S16A mutants. Notably, K5A and E8A mutations showed identical

Due to a nanomolar dimerization K_d^{18} and preservation of intermolecular salt bridges [K17-E22, E24-K29, Fig. 2(C)], all mutant LZ-GCN4 constructs were expected to be dominated by the coiled coil conformation under conditions of the ribonuclease assays. Indeed, at 37°C, pH 7.1 and 50 μ M protein concentration, only the coiled coil conformation was observed in the [^1H - ^{15}N]-HSQC NMR spectra [Fig. 2(E-H)]. The heteronuclear ^{15}N -(^1H)-NOE (Supporting Information Fig. S6) and transverse relaxation (Supporting Information Fig. S7) data suggested that the introduced mutations had only marginal effect on the overall integrity of LZ.

To determine the residues involved in RNA binding, NMR titration experiments were performed employing ^{15}N -labeled LZ peptides and 5-mer RNA (GCAGG). Due to transient aggregation of LZ only dissociation constant of the K5A mutant was directly determined from the fits of perturbation data, while other constants were determined relative to the K5A and thus provide only the conditional K_d values. Perturbation data shows that lysine and arginine mutants have decreased affinity to the RNA substrate (Fig. 3) indicating that those residues are likely involved in the electrostatic stabilization of

the LZ-RNA complex via the RNA phosphate backbone. Increased chemical shift perturbations observed upon replacement of negatively charged residues (E8A, D9A, E12A) suggested that elimination of negative charges increases the affinity of LZ to RNA due to the reduction of electrostatic repulsion. The S16A mutant has also shown increased CSPs compared to the wild-type protein, however, in contrast to anionic residue mutations, this did not lead to observable precipitation at 180 μ M RNA. This suggests that Ser16 is indeed located in vicinity of the LZ-RNA binding interface, or may indirectly affect RNA binding through modulation of the K10-E13-K18 ionic bridge [Fig. 2(C)]. In conclusion, the results of mutational analysis corroborated the LZ-GCN4-RNA binding interface mapped by NMR.

NMR experiments revealed a strong aggregation of LZ•RNA complexes under a variety of conditions (e.g. Fig. 3, Supporting Information Fig. S1). The effect was especially prominent in the case of the E8A mutant, which formed observable aggregates already at 100 μ M RNA concentration. Aggregation lead to notable variability in the LZ catalytic rates, precluding reliable quantification of kinetic constants and attribution of individual amino acid functional roles in the catalysis. Strong dependence of LZ-catalyst reaction rates on the aggregation/oligomerization state of the molecule was also previously shown in the LZ-aminoacyl transferase system.¹⁹ We envision that tendency to aggregation shall disappear in the context of a full-length bZIP protein, a possibility which we plan to address in the follow-up study.

Putative reaction mechanism

Based on the chemical shift perturbation data, catalytic degradation of RNA by LZ-GCN4 may follow either acid-base mechanism of transesterification which is characteristic for many ribonucleases,^{20–22} or base-induced mechanism characteristic for cationic peptides and oligoamines.^{23,24} The metal ion-dependent phosphoryl transfer mechanism²⁵ can be ruled out due to the independence of LZ catalytic activity on divalent ions.⁷

In the case of acid-base catalysis, the active site of LZ-GCN4 may be comprised of E8, D9, or E12 as a catalytic base and R3, K5 or K10 as a catalytic acid. In this scenario Glu/Asp sidechain would deprotonate 2'OH of the ribose phosphate, resulting in a 2'O[−] nucleophile attack on the positively charged phosphorous atom producing a pentavalent transition state followed by protonation of the 5' leaving group by K5, K10 side-chains, or the R3-activated water molecule [Fig. 4(C)].

For the base-induced catalysis, a nucleophilic attack of the 2'O[−] alkoxide on the phosphorous atom can be facilitated by the lysine-induced proton extraction from 2'OH. One of the remaining cationic

sidechains would then enable protonation of the leaving 5'OH group [Fig. 4(D)]. Both acid-base and base-induced reactions yield a 2',3'-cyclic phosphate, consistent with our experimental observations.

Our previous study suggested that the acid-base mechanism may be more favorable, due to the kinetic effect of the E12A and E12Q mutations on the LZ-GCN4 catalytic activity.⁷ The new data supports this notion showing that the replacement of anionic side-chains (E8A, D9A, E12A) increases the RNA binding affinity in LZ-GCN4. The strongest effect is observed upon the E8A mutation, indicating that this sidechain is the closest to the RNA phosphate backbone and therefore has the highest prospect to act as a catalytic base.

RNA binding site preservation

To analyze occurrence of the potential RNA binding site within the bZIP family, three PROSITE patterns were designed reflecting the minimal catalytic site assembly. Analysis of 328 documented bZIP proteins employing these patterns yields 21 matches stemming from nonorthologous proteins (Supporting Information Fig. S9). Nine among nonorthologous hits are from human proteins ATF1, ATF4, CREB1, CREM, CEBPB, CEBPE, JUN, JUND, and JUNB, thus suggesting that 17% of all 53 human bZIP factors may theoretically possess ribonuclease activity. Importantly, the PROSITE search matches the same *N*-terminal heptads in LZ-cJun, a motif which in our previous study⁷ exhibited RNase activity similar to LZ-GCN4. This suggests that the RNA binding interface is preserved between the yeast GCN4 and human c-Jun proteins.

One caveat in such an analysis is that the charged residues of the RNA binding site can simultaneously contribute to the electrostatic interactions stabilizing LZ structure.^{17,26} Hence, evolutionary preservation of these sidechains may also involve a structural component, which cannot be accurately implemented in the pattern-based search. A follow-up study is envisaged to experimentally evaluate the possibility of RNA binding and catalytic activity preservation in other bZIP motifs.

Possible biological role

The fact of RNA binding being attained by the dimeric form of LZs suggests a possible rationale behind bZIP catalytic activity. From the folding standpoint, dimerization and DNA-binding of bZIP transcription factors may follow either the “dimer pathway” (dimerization occurring prior to DNA binding) or the “monomer pathway” (dimerization occurring upon sequential binding of two monomers to the target DNA). It was shown that the monomer pathway is more relevant *in vivo*, especially for proteins with marginal LZ stability.^{27–29} Combined with our experimental data showing that RNA binding is

attained only through the dimeric form of LZs, this tempts us to hypothesize that catalytic activity of bZIP proteins *in vivo* will primarily be associated with the DNA-bound form of the TF. While in other cellular contexts bZIP motifs may have little or no activity due to the prevalence of the monomer form. This RNase activity of bZIPs may establish a negative feedback loop for the transcription activation process limiting the transcript levels of the activated target genes or reducing the levels of transcriptional noise.³⁰ Here we should emphasize that the question of biological relevance of LZ RNase activity remains open, and a comprehensive mechanistic analysis of LZ-mediated RNA degradation is required to aid potential cell-level studies.

Materials and Methods

Substrate RNA

Substrate RNA was purchased from IBA GmbH (Goettingen, Germany) and Microsynth AG (Balgach, Switzerland). RNA concentrations were determined from absorbance at 260 nm, using molar extinction coefficients of 58 and 179 $\text{mM}^{-1} \text{cm}^{-1}$ for 5-mer GCAGG and 18-mer RNA respectively. 18-mer RNA was identical to the one used in our previous study (GGUCUGCGAAUACCAGG⁷). The water used for samples and buffers was ribonuclease-free (commercial RNase/DNase-free water or ultrafiltered water assayed for the lack of RNA hydrolytic activity).

Recombinant LZ-GCN4

Recombinant LZ-GCN4 was cloned, expressed and purified as described previously.⁸ To achieve sufficient purity peptides were subjected to two rounds of RP-HPLC purification (on preparative 5- μm C8 and then analytical 3- μm C18 columns), and were shown to be homogenous by amino acid analysis, HPLC, MS, and NMR. Peptide concentrations were determined by measuring Tyr absorbance at 280 nm and/or from amino acid analysis. Throughout the manuscript residue numbering of LZ-GCN4 starts from the first native LZ residue ignoring a non-native N-terminal Gly remnant from the expression vector. The sequence of 36-residue LZ-GCN4 was G LQRMKQLEDK¹⁰ VEELLSKNYH²⁰ LENEVARLKK³⁰ LVGER.

Mutant constructs

Mutant constructs of LZ-GCN4 were expressed as GB1 fusions using pGEV2 vector.³¹ Mutagenesis was performed employing standard QuickChange[®] (Stratagene, La Jolla) procedures. Proteins were expressed using BL21 cells in $^{15}\text{NH}_4\text{Cl}$ M9 medium; purified on IgG Sepharose (GE Healthcare, Uppsala) according to manufacturer's protocol; dialyzed against 50 mM Tris, 100 mM NaCl, 5 mM DTT, 5 mM CaCl_2 , pH 8.0; cleaved from fusion by 10 U/mg

Factor Xa (Qiagen, Basel) for 20–40 h at room temperature and double purified by RP-HPLC on 5- μm preparative C8 and 3- μm analytical C18 columns. After lyophilization the purity of the peptides was assessed by ESI-MS and NMR.

CD measurements

CD measurements were performed on a 62ADS AVIV spectropolarimeter (AVIV, Lakewood, NJ US) on 30 μM peptides in 10 mM NaPO_4 , 85 mM KCl, pH 7.1. A 1-mm path length cell was used and two scans were collected with 4 s average for wavelengths between 195 and 260 nm. Melting curves were recorded at 222 nm in the 15–89°C temperature range with a 2°C step, 1°C/min heating rate and 1 min equilibration. The alpha-helical content was estimated from the molar ellipticity at 222 nm using equation [% helix] = $(-[\theta]_{222 \text{ nm}} + 3000)/39,000$.³²

NMR spectroscopy

Measurements were performed on a Bruker AVANCE 600 MHz (Biozentrum, University of Basel and LPC, ETH Zurich), 800 MHz (Biozentrum, University of Basel), or 900 MHz (BMRZ BPC, University of Frankfurt) spectrometers equipped with *z*-axis gradient triple resonance cryoprobes. Data were processed with XWINNMR 3.5, TopSpin 2.1 (Bruker Biospin, Fallanden, Switzerland) or NMRPipe³³ and analyzed using CARA (cara.nmr.ch).³⁴

Chemical shift assignments

Chemical shift assignments of LZ-GCN4 dimeric (coiled coil) and monomeric (x-form) conformations were reported earlier (BMRB 17511).⁸ Resonance assignments of the LZ-GCN4 mutants were obtained employing 3D ^{15}N -NOESY spectra.

^{15}N -Relaxation rates

^{15}N -Relaxation rates were measured using a 600 MHz spectrometer with TXI probe with 210 μM peptide samples in 20 mM HEPES- D_{18} , 85 mM KCl, 3% D_2O , pH 7.1 at 15°C. Heteronuclear steady-state ^{15}N - $\{^1\text{H}\}$ -NOE, R2 and R2 CPMG data were recorded with 40–48 scans. Transverse relaxation rates and corresponding errors were calculated employing CurveFit (A. G. Palmer, Columbia University, www.palmer.hs.columbia.edu) based on six relaxation delay values (8, 16, 32, 64, 96, and 128 ms).

RNA titration

NMR titration experiments at neutral pH (Figs. 1, 4 and Supporting Information Fig. S2) were performed in 20–40 mM HEPES- D_{18} (Spectra Stable Isotopes). In experiments at acidic pH 50 mM, acetate- D_3 was used (pH 4.0, Supporting Information Fig. S1). All samples contained 85 mM KCl and 3% D_2O . In some experiments temperatures were employed to retain visibility of the N-terminal residues (Fig. 1). Sample

and measurement conditions for NMR titration experiments are summarized in the Supporting Information Table S10.

Chemical shift perturbations were calculated as $\Delta\delta = \sqrt{((\Delta\delta_H)^2 + (0.2 \times \Delta\delta_N)^2)/2}$.

Equilibrium dissociation constants were determined by correlating chemical shift perturbations with total RNA concentration using a two-state interaction model (assuming 1:1 stoichiometry, due to the presence of two symmetrical RNA binding sites):

$$\Delta\delta = \Delta\delta_{\max} \times A - \sqrt{A^2 - 4 \times [\text{RNA}] \times [\text{LZ}]} / (2 \times [\text{RNA}])$$

$$A = K_d + [\text{RNA}] + [\text{LZ}]$$

taking into account the change in total peptide concentration ([LZ]) as a codependent variable. For LZ-GCN4 mutants interacting with 5-mer RNA, only K5A exhibited saturation binding kinetics devoid of aggregation effects. Dissociation constants of all other mutants were obtained by normalization of $K_{d,K5A}$ using the mean of relative chemical shift perturbation of residues within the first two heptad repeats ($K_d = K_{d,K5A} \times \overline{\Delta\delta_i/\Delta\delta_{K5A}}$; $i = \text{Q2:S16}$) at 50 μM RNA concentration (when aggregation effects are minimal). The reported errors (SD) reflect the variation of the relative CSP [$\sigma(\Delta\delta_i/\Delta\delta_{K5A})$] for the residues in first two heptads.

PROSITE searches were performed employing database release 20.66.³⁵

Acknowledgments

The authors thank Prof. Grzesiek (Biozentrum, University of Basel) and EU-NMR BMRZ facility (University of Frankfurt, EU-NMR, Contract # RII3-026145) for the vested NMR spectrometer time. They also acknowledge Dr. Xue (FMI, Basel) and Dr. Jelesarov (University Zurich) for the cDNAs of the c-Jun and GCN4.

References

- Jeffery CJ (1999) Moonlighting proteins. *Trends Biochem Sci* 24:8–11.
- Jeffery CJ (2009) Moonlighting proteins—an update. *Mol Biosys* 5:345–350.
- Hu S, Xie Z, Onishi A, Yu X, Jiang L, Lin J, Rho HS, Woodard C, Wang H, Jeong JS, Long S, He X, Wade H, Blackshaw S, Qian J, Zhu H (2009) Profiling the human protein-DNA interactome reveals ERK2 as a transcriptional repressor of interferon signaling. *Cell* 139:610–622.
- Deppmann CD, Alvania RS, Taparowsky EJ (2006) Cross-species annotation of basic leucine zipper factor interactions: insight into the evolution of closed interaction networks. *Mol Biol Evol* 23:1480–1492.
- Tupler R, Perini G, Green MR (2001) Expressing the human genome. *Nature* 409:832–833.
- Vinson C, Acharya A, Taparowsky EJ (2006) Deciphering B-ZIP transcription factor interactions in vitro and in vivo. *Biochim Biophys Acta* 1759:4–12.
- Nikolaev Y, Deillon C, Hoffmann SR, Bigler L, Friess S, Zenobi R, Pervushin K, Hunziker P, Gutte B (2010) The leucine zipper domains of the transcription factors GCN4 and c-Jun have ribonuclease activity. *PLoS ONE* 5:e10765.
- Nikolaev Y, Pervushin K (2007) NMR spin state exchange spectroscopy reveals equilibrium of two distinct conformations of leucine zipper GCN4 in solution. *J Am Chem Soc* 129:6461–6469.
- Schwalbe M, Dutta K, Libich DS, Venugopal H, Claridge JK, Gell DA, Mackay JP, Edwards PJ, Pascal SM (2010) Two-state conformational equilibrium in the Par-4 leucine zipper domain. *Proteins* 78:2433–2449.
- Newbury SF (2006) Control of mRNA stability in eukaryotes. *Biochem Soc Trans* 34:30–34.
- Mata J, Marguerat S, Bahler J (2005) Post-transcriptional control of gene expression: a genome-wide perspective. *Trends Biochem Sci* 30:506–514.
- Meyer S, Temme C, Wahle E (2004) Messenger RNA turnover in eukaryotes: pathways and enzymes. *Crit Rev Biochem Mol Biol* 39:197–216.
- Wilusz CJ, Wilusz J (2004) Bringing the role of mRNA decay in the control of gene expression into focus. *Trends Genet* 20:491–497.
- Renatus M, Stennicke HR, Scott FL, Liddington RC, Salvesen GS (2001) Dimer formation drives the activation of the cell death protease caspase 9. *Proc Natl Acad Sci U S A* 98:14250–14255.
- Meisner WK, Sosnick TR (2004) Fast folding of a helical protein initiated by the collision of unstructured chains. *Proc Natl Acad Sci U S A* 101:13478–13482.
- Camacho CJ, Kimura SR, DeLisi C, Vajda S (2000) Kinetics of desolvation-mediated protein-protein binding. *Biophys J* 78:1094–1105.
- Nikolaev Y, Pervushin K (in press) Rethinking Leucine Zipper—a ubiquitous signal transduction motif. *Nat Precedings*. Available at: <http://hdl.handle.net/10101/npre.2009.3271.1>.
- Zitzewitz JA, Bilsel O, Luo J, Jones BE, Matthews CR (1995) Probing the folding mechanism of a leucine zipper peptide by stopped-flow circular dichroism spectroscopy. *Biochemistry* 34:12812–12819.
- Leman LJ, Weinberger DA, Huang ZZ, Wilcoxon KM, Ghadiri MR (2007) Functional and mechanistic analyses of biomimetic aminoacyl transfer reactions in de novo designed coiled coil peptides via rational active site engineering. *J Am Chem Soc* 129:2959–2966.
- Findlay D, Herries DG, Mathias AP, Rabin BR, Ross CA (1961) The active site and mechanism of action of bovine pancreatic ribonuclease. *Nature* 190:781–784.
- Sorrentino S (1998) Human extracellular ribonucleases: multiplicity, molecular diversity and catalytic properties of the major RNase types. *Cell Mol Life Sci* 54:785–794.
- Koroleva LS, Serpukrylova IY, Vlassov VV, Silnikov VN (2007) Design and synthesis of metal-free artificial ribonucleases. *Protein Pept Lett* 14:151–163.
- Barbier B, Brack A (1992) Conformation-controlled hydrolysis of polyribonucleotides by sequential basic polypeptides. *J Am Chem Soc* 114:3511–3515.
- Yoshinari K, Yamazaki K, Komiyama M (1991) Oligoamines as simple and efficient catalysts for RNA hydrolysis. *J Am Chem Soc* 113:5899–5901.
- Scott WG (1999) RNA structure, metal ions, and catalysis. *Curr Opin Chem Biol* 3:705–709.
- Steinmetz MO, Jelesarov I, Matousek WM, Honnappa S, Jahnke W, Missimer JH, Frank S, Alexandrescu AT, Kammerer RA (2007) Molecular basis of coiled-coil formation. *Proc Natl Acad Sci U S A* 104:7062–7067.

27. Kohler JJ, Schepartz A (2001) Kinetic studies of Fos.-Jun.DNA complex formation: DNA binding prior to dimerization. *Biochemistry* 40:130–142.
28. Cranz S, Berger C, Baici A, Jelesarov I, Bosshard HR (2004) Monomeric and dimeric bZIP transcription factor GCN4 bind at the same rate to their target DNA site. *Biochemistry* 43:718–727.
29. Seldeen KL, McDonald CB, Deegan BJ, Farooq A (2008) Evidence that the bZIP domains of the Jun transcription factor bind to DNA as monomers prior to folding and homodimerization. *Arch Biochem Biophys* 480: 75–84.
30. Bluthgen N (2010) Transcriptional feedbacks in mammalian signal transduction pathways facilitate rapid and reliable protein induction. *Molecular bioSystems* 6: 1277–1284.
31. Koenig BW, Rogowski M, Louis JM (2003) A rapid method to attain isotope labeled small soluble peptides for NMR studies. *J Biomol NMR* 26:193–202.
32. Woody RW (1995) Circular dichroism. *Methods Enzymol* 246:34–71.
33. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293.
34. Keller RLJ.2004. *The Computer Aided Resonance Assignment Tutorial*, Germany: CANTINA Verlag.
35. Sigrist CJ, Cerutti L, de Castro E, Langendijk-Genevaux PS, Bulliard V, Bairoch A, Hulo N (2010) PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res* 38: D161–D166.