Published in final edited form as:

J Mol Biol. 2012 September 21; 422(3): 347–365. doi:10.1016/j.jmb.2012.05.037.

The Brace for a Growing Scaffold: Mss116 Protein Promotes RNA Folding by Stabilizing an Early Assembly Intermediate

Olga Fedorova and Anna Marie Pyle*

Howard Hughes Medical Institute and Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06520, USA

Abstract

The ai5 γ group II intron requires a protein cofactor to facilitate native folding in the cell. Yeast protein Mss116 greatly accelerates intron folding under near-physiological conditions both *in vivo* and *in vitro*. Although the effect of Mss116 on the kinetics of ai5 γ ribozyme folding and catalysis has been extensively studied, the precise structural role and interaction sites of Mss116 have been elusive. Using Nucleotide Analog Interference Mapping to study the folding of splicing precursor constructs, we have identified specific intron functional groups that participate in Mss116-facilitated folding and we have determined their role in the folding mechanism. The data indicate that Mss116 stabilizes an early, obligate folding intermediate within intron domain 1, thereby laying the foundation for productive folding to the native state. In addition, the data reveal an important role for the IBS2 exon sequence and for the terminus of domain 6, during the folding of self-splicing group IIB intron constructs.

Keywords

ribozyme; RNA folding; DEAD-box protein; NAIM

Introduction

The biological function of large RNAs is contingent upon their ability to fold into a native conformation inside the cell. Multiple studies converge on two main folding landscapes for large RNAs: smooth, which is a direct transition from unfolded to the native state,¹⁻⁵ and rugged, which involves one or more misfolded off-pathway intermediates ('kinetic traps').⁶⁻⁹ Although *in vitro* folding of some large RNAs is fast,^{10,11} the majority of them fold slowly. Slow folding of RNAs may be caused by instability of obligate on-pathway intermediates,^{4,12-14} by high contact order,¹⁵⁻¹⁷ or by misfolding into kinetic traps.⁶⁻⁹ In the cell, it is thought that these folding problems are mitigated by protein cofactors, which use a variety of mechanisms to guide proper RNA folding *in vivo*.¹⁸⁻²⁰ Some proteins facilitate the process by destabilizing misfolded intermediates. For example, RNA helicases^{21,22} can unwind inhibitory RNA duplexes in an active mechanical process that requires ATP

hydrolysis,^{19,23} while RNA chaperones (e.g., StpA) transiently bind and weaken misfolded structures independent of ATP.^{18,20,24} Other proteins promote RNA folding in the opposite manner, by stabilizing RNA structure.^{13,14} For example, some proteins function as RNA annealing enzymes that stimulate the formation of RNA secondary structures,^{18,25,26} while others stabilize transient on-pathway intermediates²⁷⁻³⁰ or trap the native tertiary structure.³¹

Multiple studies suggest that ribozymes derived from the yeast mitochondrial group II intron ai5 γ fold slowly *in vitro* as they proceed directly to the native state via several on-pathway intermediate states.³⁻⁵ We have recently shown that folding of ai5 γ ribozymes is under kinetic control under near-physiological conditions and that there is a high energy barrier between unfolded and compact intermediate states.³² This process takes 24–26 h,³² which explains the need for protein cofactors for ai5 γ intron folding *in vivo*. Yeast mitochondrial protein Mss116 stimulates folding and splicing of the ai5 γ intron both *in vitro* and *in vivo*.^{28-30,33-38} Two different models have been proposed to describe the role of Mss116 during intron folding. The first mechanism suggests that the protein facilitates intron folding by unwinding undesirable secondary structural elements that interfere with proper assembly to the native state.^{34,36-38} This is a reasonable hypothesis given that Mss116 is a DEAD-box protein, containing conserved ATPase motifs that typify a subset of RNA helicase enzymes.³⁹⁻⁴¹ The hypothesis is supported by the ability of two related proteins (CYT-19 and Ded-1) to selectively destabilize misfolded structures within the *Tetrahymena* ribozyme (derived from an unrelated group I intron in a protist), thus allowing this RNA to correctly refold.^{38,42}

The second mechanism suggests that Mss116 stimulates ai5 γ intron folding by binding to the RNA and stabilizing weak folding intermediates.^{28-30,35} It has been shown that, although folding of the ai5 γ intron is slow, its folding pathway is smooth and devoid of misfolded intermediates.^{3,4,17} Direct monitoring of Mss116-stimulated ai5 γ folding using both single-molecule methods and native gel electrophoresis shows that the protein promotes the formation of an obligate, compact intermediate early in the folding pathway and that this process is ATP independent.^{30,35} Subsequent ATP hydrolysis is required to recycle the protein and to release protein molecules that are trapped on the RNA.^{30,43} These data are inconsistent with a requirement for un-winding activity of Mss116 during the formation of the compact state.

Both of the models for Mss116 function employ folding or activity assays to support their arguments. However, the identity of intronic regions affected by the protein remains unknown. In order to directly determine which RNA structural elements are influenced by the protein, and to ascertain their precise mechanistic role, we used Nucleotide Analog Interference Mapping (NAIM)⁴⁴⁻⁴⁶ coupled to a functional assay for Mss116-stimulated ai5 γ intron folding. This approach shows that, while Mss116 influences the behavior of nucleotides throughout the intron, it has dramatic effects on a small set of nucleotides located in a three-way junction at the center of intron domain 1 (D1). Of particular significance to the actual mechanism of folding, we observe that nucleotide analogs known to destabilize helical structure display large interference effects in the presence of Mss116, while nucleotide analogs that stabilize RNA duplexes show pronounced enhancements at the same positions within the D1 substructure. These results show that Mss116 stimulates ai5 γ

intron folding by stabilizing a specific junction within intron D1, thereby contributing to the formation of the D1 scaffold that is required for subsequent rapid stages of intron assembly.

Results

Experimental design for NAIM analysis of Mss116-stimulated compaction

The reaction conditions used in this work differ from previous NAIM studies of ai5 γ folding, which employed high concentrations of monovalent and divalent cations (100 mM MgCl₂ and 0.5 M KCl) at elevated temperature (42 °C).¹² In this work, we conducted NAIM studies of Mss116-dependent intron folding under near-physiological conditions (8 mM MgCl₂ and 100 mM KCl, 30 °C). Since these changes in reaction condition are expected to affect the pattern of NAIM effects observed in the absence of the protein, it was necessary to reexamine NAIM effects on compaction of the RNA alone.

As shown previously, Mss116 converts the slow-folding population of intron molecules into a fast-folding population and increases the folding rate constant, suggesting that this protein affects the mechanism of the ai5 γ RNA folding.^{30,35,36} In order to probe atoms specifically involved in the rate-limiting step of folding, we limited the RNA incubation time under folding conditions. This approach pinpoints residues important for early stages of the folding pathway, as described previously.¹²

During the NAIM experiments, RNA was transcribed in the presence of a nucleotide analog phosphorothioate, resulting in a pool of RNA molecules containing one modified nucleotide per molecule⁴⁷ (Fig. 1a). The RNA pool was then labeled with ³²P either at its 3'- or 5'-end and incubated under near-physiological conditions until only 15–20% of the compact population is formed. The unfolded population was then separated from the compact population by native gel electrophoresis and both populations were treated with iodine, which selectively cleaves phosphorothioate linkages.¹² Cleavage patterns corresponding to the unfolded and compact populations were then analyzed on a sequencing gel. Nucleotides containing modifications that interfere with or facilitate compaction are underrepresented or overrepresented in the compact population, respectively (Fig. 1).

In order to determine whether Mss116 stabilizes or destabilizes RNA duplexes in order to promote ai5 γ intron folding, we selected nucleotide analogs that specifically stabilize or destabilize base pairing (Fig. 1b and c). When an adenosine residue is replaced by a 2,6-diaminopurine (2,6-DAP), the latter stabilizes the duplex by forming an additional hydrogen bond with a complementary uridine residue⁴⁸ (Fig. 1b and c). Conversely, when an adenosine residue is replaced by an N6 methyl adenosine, the modification hinders the formation of one of the two hydrogen bonds with a uridine residue and thus destabilizes base pairing.⁴⁸ The same destabilizing effect is observed when a guanosine is replaced by inosine, which reduces the number of hydrogen bonds with the complementary cytidine from 3 to 2.⁴⁸ We reasoned that, if Mss116 unwinds an RNA duplex to promote successful compaction, then the presence of inosine and N6 methyl adenosine residues in this duplex will be beneficial for protein-facilitated compaction, and we will observe enhancement effects with the respective nucleotide analogs. By contrast, adenosine substitution with 2,6-DAP in the same duplex will result in interference with protein-assisted compaction. If

Mss116 stabilizes an RNA duplex to facilitate compaction, then the NAIM effects will be reversed: interference will be observed for inosine and N6 methyl adenosine analogs and enhancement for 2,6-diaminopurine substitution. In addition to these three analogs, we have used adenosine and guanosine phosphorothioates to account for potential phosphorothioate effects. In parallel experiments, 7-deaza analogs were used to probe the importance of Hoogsteen hydrogen bonding for intron compaction.

Previous NAIM folding studies were conducted on the D135 ribozyme, which contains domains 1, 3, and 5 and short stem-loops replacing domains 2 and 4.^{12,48} However, this construct is less suitable for studies on the influence of Mss116 because 15–20% of D135 rapidly forms the compact state even in the absence of protein, thereby hindering identification of NAIM effects caused specifically by Mss116.³⁰ In order to avoid this problem, we used the SE self-splicing ribozyme construct, which contains the full-length intron and short exons (Supplementary Fig. 1). In the absence of Mss116, this RNA folds slowly and its fast-folding population is very small.³⁰ In order to facilitate NAIM mapping of residues located in the middle of the intron molecule, we also used shortened versions of the SE construct with truncated D4 or D2 helices (SE 2–4, see Supplementary Material).

In order to compare NAIM effects observed with and without Mss116, the selection step (RNA compaction) was timed to harvest the same initial compact population (15–20%) in both cases. This population of compact species was formed after 45 min of incubation in the absence of protein and after only 1 min of incubation in the presence of Mss116 and ATP (Fig. 2). Although Mss116 stimulates intron compaction without ATP,^{30,35} compaction is further enhanced in the presence of ATP.^{30,35,49} We therefore conducted our experiments in the presence of ATP in order to optimize signal and stringently differentiate effects observed in the presence and absence of Mss116. In this way, we optimized amounts of protein-stimulated compact species for iodine sequencing.

Interestingly, slightly different folding products were observed for the full-length and SE 2–4 constructs (Fig. 2). In the absence of the protein, SE 2–4 forms small amounts of an additional extended intermediate species that is eliminated when the protein is added to the reaction. Since behavior of this intermediate is influenced by Mss116, we isolated and characterized it by NAIM, observing small effects in the D1c2 region (Supplementary Fig. 2).

Binding of the 5'-exon is important for folding of the self-splicing RNA

Because previous NAIM studies on group II intron folding were conducted on the D135 ribozyme, which lacks both 5'- and 3'-exons,^{12,48} we were unable to evaluate the role of flanking exons during intron folding. In this study, we employ self-splicing precursor constructs, which enables us to learn about Mss116 and also to obtain much-needed information on the participation of exons during folding. The NAIM approach allows us to assess the role of individual exon functional groups during formation of the compact state.

Interestingly, we found that inosine modifications at positions G-16, G-14, and G-11 in the 5'-exon interfere with intron compaction (Fig. 3, left panel, also see the bar graph). Remarkably, these guanosine positions are located in the IBS2 sequence, which is one of

two regions that base pairs with complementary sequences within intron D1 (the exon-binding or EBS1 and 2 sequences), thereby contributing to recognition of the 5'-splice site and exogenous targets by group II introns.⁵⁰⁻⁵⁴ The unusual aspect of this result is that IBS2 is the recognition element located farthest from the intron, and as a seemingly 'peripheral' element, it was not presumed to play a major role in function. Since the observed inosine substitutions in this region weaken base pairing, these data suggest that the formation of the IBS2-EBS2 duplex contributes intron compaction, stabilization, and global folding.

Surprisingly, no interference or enhancement effects were observed in the EBS1 sequence, suggesting that the formation of the EBS1-IBS1 duplex does not strongly influence intron folding. However, moderate interference of both 7-deaza adenosine and 2,6-DAP were observed at A-8, which is located between the EBS1 and EBS2 regions (Fig. 3, middle and right panels and the bar graph; Fig. 4a and b). These data suggest that the Hoogsteen face and possibly the minor groove of A-8 are involved in tertiary contacts that contribute to formation of the compact state.

Notably, interference effects of the same magnitude were observed, at the same nucleotide positions, both in the presence and in the absence of Mss116 (Fig. 3, bar graph, compare also Fig. 4a and b). This suggests that the 5'-exon plays the same role in the intron folding both in the absence and in the presence of the protein.

Nucleotides in the upstream portion of D1 contribute to folding

According to previously published NAIM results, nucleotides located at the beginning of D1 are not important for D135 compaction under high salt conditions.¹² By contrast, this region appears to be important for folding of the self-splicing RNA at 30 °C and 8 mM MgCl₂. Our data indicate that many functional groups in the first 20 nucleotides of D1 are important for compaction of self-splicing constructs, both in the absence (Figs. 4a and 5) and in the presence (Figs. 4b and 5) of Mss116. For example, moderate (~2-fold) inosine interference is observed at G3 (Fig. 4a and b; Fig. 5a, left panel, and b, green bars). This suggests that the formation of the ϵ - ϵ' contact may contribute to intron compaction. The internal bulge in the 'i' stem (the i-loop) exhibits a cluster of interference effects. For example, there are 7-deaza A (Fig. 4a and b; Fig. 5a, middle panel, and b, blue bars) and 2,6-DAP (Fig. 4a and b; Fig. 5a, right panel, and b, red bars) interferences at A12 and inosine interferences at G10 and G14 (Fig. 4a and b; Fig. 5a, left panel, and b, green bars). Notably, all these effects are located in the z-anchor region, which welds together long helices IC and I(ii) in a side-by-side arrangement that would be expected to considerably reduce the volume and surface area of D1.⁵⁵ These data therefore indicate that the interaction between the i-loop and the z-anchor is beneficial for the intron compaction. Additional interference effects are found in the double-stranded region of the i stem, including 7-deaza A interference at A17 (Fig. 4a and b; Fig. 5a, middle panel, and b, blue bars) and 2,6-DAP effect at A20 (Fig. 4a and b; Fig. 5a, right panel, and b, red bars). Since these residues are in close proximity to the terminus of the c helix,⁵⁶ this raises a possibility of tertiary contacts between i and c helices of the ai5 γ intron, a fact that cannot be verified by comparison with the O.i. group II intron crystal structure since the two introns diverge at this position. However, in ai5 γ , this putative

interaction is likely to stabilize the complex network of tertiary contacts formed by the z-anchor.

As observed for D135,^{12,48} nucleotides located in the five-way junction of the self-splicing RNA exhibit moderate to strong interference effects with multiple nucleotide analogs (Figs. 4 and 5). Interestingly, in addition to previously described effects at A27, A72, A191, and A391, several additional interference effects are found in this region. For example, 2,6-DAP effects are observed at A26 and A71 (Fig. 4a and b; Fig. 5a, right panel, and b, red bars), and 7-deaza A interference is observed at A29 (Fig. 4a and b; Fig. 5a, middle panel, and b, blue bars). This implies that folding of self-splicing constructs is particularly sensitive to the conformation of the five-way junction.

Formation of the κ - ζ element is critical for intron compaction

Previous results indicate that the formation of the κ - ζ element (previously dubbed ‘the folding control element’) is critical for compaction of the D135 ribozyme at 42 °C and high concentrations of potassium and magnesium ions.¹² Our results confirm that this substructure is of major importance for compaction of self-splicing constructs under near-physiological conditions both in the absence and in the presence of Mss116. Most of the intron positions that exhibit interference or enhancement effects upon incorporation of a modified analog are located in this region (Fig. 4a and b). With the exception of a strong phosphorothioate enhancement at A209, the rest of the NAIM effects follow a characteristic pattern for weak duplexes that benefit from additional stabilization: interference effects observed with N6 methyl A at adenosine positions located in helical regions are coupled with enhancement effects observed at the same positions in the presence of 2,6-DAP (Fig. 1c; Fig. 4a and b; Fig. 6a, top, first and second panels from the left). Additionally, interference effects with inosine are observed at guanosine positions located in the same helix (Fig. 1c; Fig. 4a and b; Fig. 6b, left panel). This pattern is particularly pronounced in helical regions **d''**, **d'''**, and **d2a** located in D1 (Figs. 4a and b and 6a and b).

Interestingly, regions **d**, **d'**, **d''**, **d'''**, and **d2a** contain the highest number of NAIM effects that are specifically sensitive to the presence of Mss116 (denoted by arrows on the left of the respective gels in Fig. 6a and b; see also Fig. 6c). Some of the effects, which are weakly observed in the absence of the protein, become ~2-fold stronger in the presence of Mss116 and ATP. For example, there was a 2.2-fold phosphorothioate enhancement at A209 in the absence of the protein and a 5-fold enhancement in the presence of the protein (Fig. 6a, bottom, compare gray and black bars). Enhancements observed at adenosine positions 219 and 220 in the presence of 2,6-DAP increase from 1.7- and 1.8-fold in the absence to 2.6- and 2.8-fold in the presence of Mss116, respectively (Fig. 6a, bottom, compare light and dark red bars at respective nucleotide positions). In the same duplex, interference effects at A218 and A219 in the presence of N6 methyl A increase from 2- and 1.4-fold in the absence to 3- and 1.9-fold in the presence of Mss116 (Fig. 6a, bottom, compare light and dark purple bars at respective nucleotide positions). In addition, N6 methyl A interference at A376, which is located on the opposite strand of the same duplex, increases from 2.1- to 3.4- fold in the presence of the protein (Fig. 6a, bottom, compare light and dark purple bars). A similar NAIM pattern is observed in the helix **d2a** (Fig. 6a and c, also compare Fig. 4a and

b). In this duplex, N6 methyl A interference effects at positions 245, 246, 251, and 250 increase from 1.6- to 1.9-fold in the absence of Mss116 to 2.4- to 4.4-fold in the presence of the protein (Fig. 6a, bottom, compare light and dark purple bars at respective nucleotide positions). Likewise, 2,6-DAP enhancement effects at A246 and A251 increase from 1.7- and 1.9-fold to 2.4- and 3.6-fold, respectively, when Mss116 is present (Fig. 6a, bottom, compare light and dark red bars at respective nucleotide positions). Notably, NAIM effects higher than 4-fold are quite rare, and therefore, the effects reported here are exceptionally strong. Most commonly observed NAIM effects usually range from 1.5- to 3-fold.^{12,57-61} Strong (over 4-fold) effects have only been observed when a modified residue is truly critical for the formation of an important long-range tertiary contact^{57,60} or for chemical catalysis itself.⁶¹ Thus, the large NAIM effects observed in the κ - ζ region in the presence of Mss116 indicate that the stabilization of this region is absolutely essential for protein-facilitated ai5 γ intron folding.

Many of the moderate twofold NAIM effects in the κ - ζ region only appear in the presence of Mss116. This is the case for inosine and N6 methyl A interference effects at G363 and A233, respectively, in the helix **d'''**, as well as for inosine effect at G379 in the helix **d'** (Fig. 6b, left panel; Fig. 6a, second panel from the left, positions denoted by blue arrows). These NAIM effects suggest that destabilization of **d'''** and **d'** helices is detrimental for protein-assisted intron folding. Other NAIM effects in these two duplexes are consistent with this assumption. For example, N6 methyl A interference and 2,6-DAP enhancement effects at A365 and A366 in the helix **d'''** increase by ~1.5-fold in the presence of Mss116, and so does N6 methyl A interferences at A376 in the duplex **d'** (Fig. 6a, bottom, compare light red with dark red and light purple with dark purple bars).

These data are unequivocally inconsistent with a pathway in which Mss116 unwinds or weakens duplexes within the intron during intron folding. Rather, the data clearly show the opposite effect: that stabilization of specific sections within the κ - ζ intronic region, which includes five short helices (**d**, **d'**, **d''**, **d'''**, and **d2a**, Fig. 6c), is required for compaction in the presence of Mss116. These data provide direct structural evidence that Mss116 plays a stabilization role during intron compaction, and they pinpoint the exact location where Mss116 influences local structure. It is important to emphasize that these NAIM experiments involve a randomized *in vitro* selection that allows us to examine the behavior of every nucleotide within the intron, to identify only those nucleotides that are influenced by Mss116, and to ascertain their precise mechanistic role (destabilizing or stabilizing to base pairing, etc). The results are therefore unbiased by any previous models for ai5 γ folding.

Role of other domains during intron compaction

Previously published studies have shown that domains 3 and 5 are not important for compaction of the D135 ribozyme.^{4,12,48} This supported the notion that D1 is the primary scaffold for folding of the entire intron. In order to generalize this concept, it was necessary to test the role of other domains in the context of the intact, self-splicing RNAs under near-physiological conditions, in the presence and absence of Mss116.

By using 3'-end-labeled, truncated SE 2-4 RNA, it was possible to map nucleotides within domains 3, 5, and 6 in addition to the 3'-half of D1. Consistent with the previous studies, we found no interference or enhancement effects in D5, suggesting that D5 is critical for catalysis but not for folding of the ribozyme (data not shown). In contrast with previous findings, several residues in D3 exhibit interference or enhancement effects. Moderate phosphorothioate interference is observed at both A661 and A662 in the highly conserved internal bulge of D3 (Fig. 4a and b; Fig. 7, top, second panel from the left, bottom, gray and black bars). In addition, inosine interference and N6 methyl A enhancement effects are found at G647 and A652, respectively (Fig. 4a and b; Fig. 7, top, third panel from the left, bottom, green bars).

Intriguingly, several interference effects are located in D6, which is thought to be one of the most flexible regions of the intron and therefore was not expected to affect the intron folding (Figs. 4a and b and 8). Previously studied ribozyme constructs of the ai5 γ intron did not contain D6, and thus, its structural role has remained obscure. There are multiple interference effects in the tetraloop that caps D6: A868 and A869 exhibit both 2,6-DAP and N6 methyl A interference and A870 shows N6 methyl A interference only (Figs. 4a and b and 8, top, panels 1–4 from the left, bottom, red and purple bars). In addition, phosphorothioate interference is observed at G866 (Figs. 4a and b and 8, top, panels 5 and 6 from the left, bottom, gray and black bars). Notably, in the full-length intron, the D6 tetraloop is part of the η - η' tertiary contact, which is thought to serve as a conformational switch between two steps of the group II intron splicing.⁶² Since the SE 2-4 RNA contains a truncated D2, it cannot form the η - η' contact. In order to determine if the interference pattern for the D6 tetraloop will change upon formation of the η - η' interaction, we have conducted the same experiment using the full-length SE RNA, which contains the entire D2 and can therefore form the η - η' contact with the tetraloop of D6 (Supplementary Fig. 1). We found that in the full-length SE RNA, the interference pattern for the D6 tetraloop is exactly the same as observed for the truncated SE 2-4 construct (Fig. 8, top, compare panels 1 and 2, 3 and 4, and 5 and 6). Thus, the role of this tetraloop during intron folding is independent of the η - η' interaction, which is therefore unlikely to participate in global intron folding.

The full-length SE RNA was also used to identify functional groups in D4 that contribute to intron compaction. We found that within the entire D4, only three modifications interfere with compaction: inosine effects are observed at G739, G743, and G744 (Fig. 4a and b; Fig. 7, top, first panel from the right, bottom, green bars). All three guanosine positions are located in the same helical region, suggesting that the stability of this basal duplex is beneficial for intron compaction.

In order to identify residues in D2 that contribute to compaction, we used a partially truncated SE RNA, containing a full-length D2 and a truncated D4 (Fig. 4; Supplementary Fig. 1). Similar to D4, only a few weak inosine interference effects were observed at guanosine positions G534 and G535, which are part of the η - η' tertiary contact, and guanosine positions G580 and G581, which are involved in the θ - θ' interaction (Fig. 4a and b; Fig. 7, top, first panel from the left, bottom, green bars). Interestingly, these effects are observed in the absence of Mss116, and they largely disappear when the protein is added (Fig. 7, top, first panel from the left, also compare D2 interference maps in Fig. 4a and b).

This suggests that the stability of the respective helices is important for the intron compaction only in the absence of Mss116. Interestingly, 7-deaza interference at A101, which is part of the θ tetraloop in D1, is observed only in the presence of Mss116 (Supplementary Fig. 2). This suggests that, while formation of the θ - θ' duplex is not critical for Mss116-facilitated intron folding, functional groups on this substructure may facilitate stabilization of the compact state.

Discussion

Mss116 stimulates ai5 γ RNA folding by stabilizing a specific D1 substructure

In order to identify functional groups that are important for Mss116-facilitated folding of a self-splicing intron, we carried out a NAIM study in which compaction of intron ai5 γ was used as a selection step. This compaction event has been shown through native gel electrophoresis,^{4,32} hydroxyl radical and dimethyl sulfate footprinting,^{4,32} and single-molecule methods^{5,35} to reflect the productive native folding of intron D1, which is the largest domain of the intron and the scaffold for docking of all other domains. In addition, D1 compaction represents the rate-limiting, initial event during assembly of the entire intron.^{4,12}

Here, we show that adenosine or guanosine residues that exhibit nucleotide analog interference or enhancement effects in the absence of Mss116 tend to show the same type of effects in the presence of the protein. This suggests that, although Mss116 drastically accelerates the rate of the intron folding, it does not cause fundamental changes in the mechanism of the folding pathway. Furthermore, it suggests that similar functional groups are important for intron folding in both the presence and the absence of protein. Notably, we did not observe NAIM signatures corresponding to any discrete misfolded structures. This is consistent with our previously published results suggesting that the ai5 γ intron folding pathway is smooth and devoid of misfolded intermediates.^{3,4,17}

To determine whether Mss116 assists intron folding by unwinding or by stabilizing RNA duplexes, we chose pairs of nucleotide analogs that cause either stabilization or destabilization of specific base pairings. Thus, the incorporation of these analogs into the intron RNA results in either interference or enhancement effects, depending on the role of the specific base pair during intron compaction⁴⁸ (Fig. 1b and c). We find that, both in the absence and in the presence of Mss116, the majority of adenosine and guanosine residues in duplex regions exhibit a NAIM pattern that is typical for base pairs that require additional stabilization: inosine interference at G positions and N6 methyl A interference and 2,6-DAP enhancement at A positions (Fig. 4).

A few exceptions include 2,6-DAP interference at A20, A26, A69, A71, and A72. Notably, 2,6-DAP interference at A72 substantially decreases in the presence of Mss116 (Supplementary Fig 2, dashed arrow), and the magnitude of the remaining effects remains the same in the presence and in the absence of Mss116 (Supplementary Fig. 2). This suggests that destabilization of these base pairs plays the same role or even a lesser role in the presence of Mss116 than it does in its absence. Furthermore, three of these residues (A26, A71, and A72) form terminal base pairs at the five-way junction, suggesting that a

steric clash caused by an additional amino group may be an alternative explanation for observed interference effects.

While most regions of the intron show the same NAIM effects with and without Mss116, one region in D1 stands apart from the rest of the molecule by exhibiting exceptional sensitivity to the Mss116 protein. This intronic substructure includes the κ - ζ region of D1, spanning helices **d** through **d2a**, and it has been described previously as a ‘folding control element’ (Fig. 4). Within this region, most interference or enhancement effects are 1.5–3 times stronger in the presence of the protein (Fig. 6). These unusually large effects are located primarily in helical regions and they include N6 methyl A interferences coupled with 2,6-DAP enhancements, as well as inosine interferences. This type of NAIM pattern is indicative of a weak duplex that requires additional stabilization in order to form (Fig. 1c). The additional increase in NAIM signal in the presence of Mss116 suggests that stabilization of the folding control element is critical for the protein-stimulated folding of the ai5 γ intron (Fig. 9). Together, the data suggest that Mss116 affects and facilitates intron folding by directly or indirectly stabilizing the folding control element. This stabilization lowers the energy barrier between the unfolded and the compact intermediate states, thus accelerating intron folding in the presence of Mss116 protein (Fig. 9).

While these results cannot address whether Mss116 directly binds the folding control region, other results suggest that this may be one of the possibilities. Recent single-molecule studies show that ATP is not required for protein-facilitated compaction of D135 ribozymes, but it is necessary for the native state formation in the presence of Mss116.³⁵ Multiple studies have demonstrated that ATP is required for Mss116 recycling, which involves dissociation from the product or unproductive binding sites and rebinding to the new substrate.^{30,43} Since the folding control element also serves as the receptor for binding of the catalytic domain 5, docking of D5 to this substructure is required for the formation of an active group II intron ribozyme. Therefore, if Mss116 binds the folding control element, it must dissociate in order for D5 to bind, and this is consistent with the ATP requirement for Mss116 release and native state formation.

There is ample precedent for protein cofactors that stabilize RNA folding intermediates at key stages along a sequential assembly pathway. Indeed, such protein-facilitated events are essential during ribo-somal RNA folding and ribosome assembly.^{13,14,63} While 16S rRNA can adopt a tertiary structure independent of ribosomal proteins,⁶⁴ the process requires unphysiologically high (~ 20 mM) concentrations of Mg^{2+} .¹³ Ribosomal proteins S4 and S17 serve to stabilize specific structural elements within the 16S rRNA: S4 stabilizes the five-way junction, and S17 stabilizes central junctions and the kink-turn motif.¹³ The association of these proteins reduces the magnesium ion requirements for rRNA folding by more than a thousand-fold¹³ and facilitates correct folding of the 16S rRNA under physiological conditions.¹³ Similar effects on stabilization and reduction of magnesium ion requirements are observed during Mss116-stimulated folding of group II introns.^{30,35} The two systems also share similarities when one examines the kinetic role of proteins during folding. For example, the addition of S4 causes the slow-folding five-way junction of 16S rRNA to fold within the first minute of incubation,¹⁴ which is on the same time scale as Mss116-stimulated folding of the ai5 γ intron.³⁰ Despite these similarities, there are

differences in the mechanisms by which ribosomal proteins and Mss116 promote RNA folding. Ribosomal proteins bind the junction they stabilize and stay bound during the course of ribosomal assembly, while Mss116 uses ATP to dissociate from the RNA once it has folded. Nevertheless, taken together, these data suggest that stabilization of an obligate RNA folding intermediate by a protein cofactor is a common mechanism of protein-facilitated RNA folding.

An intron–exon complex forms early in the folding pathway

Most of the previous group II intron folding studies employed the truncated D135 ribozyme, which lacks peripheral domains and exonic sequences. Although these studies provided valuable insight into the folding pathway of the intron core, they did not address the possible role of exons during intron folding. Indeed, a positive role for exons in the folding of any intronic RNA has never been demonstrated. Multiple studies have demonstrated that the D135 or D1356 ribozymes do not need to be folded in the presence of exon substrates in order to efficiently recognize and cleave them.^{17,32} Indeed, a standard ribozyme activity assay involves pre-folding of the ribozyme to the native state followed by the addition of the exonic substrate *in trans* as a short RNA oligonucleotide.⁶⁵ Furthermore, hydroxyl radical footprinting studies have suggested that adding an oligonucleotide substrate *in trans* does not affect the D135 folding pathway.¹⁷ These results led to the assumption that intron folding is completely independent of the exon participation. However, recently published results suggest that exons may influence RNA folding. Single-molecule studies using the D135 ribozyme and a complementary oligonucleotide substrate revealed an increase in native population upon addition of substrate.³⁵ In addition, we observed that *cis*-exons may stabilize the native state of a group II ribozyme.³⁰

This study further implicates the involvement of *cis*-exons during intron folding. We observe unusually strong (3- to 9-fold) inosine interference effects in the 5'-exon. Since our experiment is based on a kinetic rather than a thermodynamic selection, these data suggest that binding of the *cis*-5'-exon contributes to a rate-limiting event during intron compaction. Thus, binding of *cis*-exons is likely to occur early in the folding pathway and not after the native state is formed.

Surprisingly, all the inosine interference effects are located in the IBS2 sequence of the 5'-exon, and no interference effects are observed for the IBS1 region. This suggests that the EBS2–IBS2 helix contributes to intron folding and that formation of the EBS1–IBS1 duplex is not necessary for compaction. This is counterintuitive because EBS2–IBS2 is not believed to be essential for group II intron function. All group II introns form the EBS1–IBS1 interaction, but the EBS2–IBS2 is completely absent in some classes of group II introns, such as the ancient IIC lineage that is structurally well characterized.⁶⁶ However, the EBS2–IBS2 interaction may have evolved in more modern lineages of group II introns (such as IIA and IIB) not only for the purpose of substrate recognition but also to facilitate proper folding to a catalytically active conformation.

One of the regions that exhibited strong NAIM effects in previous studies on compaction of the D135 ribozyme is the long-range kissing loop α – α' , which brings the 5'- and the 3'-halves of D1 into proximity. Surprisingly, the same nucleotide analog substitutions in the α –

α' helix do not interfere with compaction of the splicing constructs, suggesting that this interaction does not have the same critical function during folding of the larger molecule. This raises an interesting possibility that the EBS2–IBS2 helix, which also brings the 5'- and the 3'-halves of D1 closer in space, may play a similar role in the folding pathway of splicing precursor constructs as the α – α' interaction in the D135 ribozyme. Put differently, EBS2–IBS2 formation may have become rate limiting (rather than the α – α' formation) in the context of the self-splicing construct.

The contribution of individual domains to intron compaction

Previous NAIM studies indicated that only D1 is essential for the formation of the compact state by the D135 ribozyme^{12,48} and that D3 and D5 serve a purely catalytic function. Studies of precursor compaction are largely consistent with these previous results, but there are subtle differences that reflect the more complex process of precursor assembly. For example, folding of precursor requires that many domains still adopt their proper secondary structure, as evident from several inosine interference effects in helical stem regions (Fig. 7). Moderate interferences are also observed in D3, as phosphorothioate effects are observed at A661 and A662 within the D3 internal bulge (Fig. 7). These phosphorothioate effects may indicate that the binding of magnesium ions in D3 contributes to overall intron compaction.

Particularly interesting effects are observed in D6, which has never been monitored previously because it was not included in the well-studied D135 constructs. For example, interference effects by 2,6-DAP and N6 methyl A analogs are observed at A868–A870, which are located within the η' tetraloop in D6 (Fig. 8). This is unexpected because the η' tetraloop is believed to be part of a long-range tertiary contact with D2 that mediates conformational changes between the first and the second steps of splicing.⁶² However, the same NAIM pattern within the η' tetraloop is observed for both the SE RNA (which contains full-length D2) and SE 2–4 (which contains a shortened D2 and lacks the receptor for the D6 tetraloop). Thus, the observed NAIM pattern for the D6 tetraloop during RNA folding is independent of the η – η' contact that may form later with D2. This indicates that the D6 tetraloop interacts with a different intronic element early in the folding pathway and that this as-yet uncharacterized interaction may be beneficial for intron compaction. This conclusion is consistent with modeling experiments that localize the position of D6 based on the recent crystal structure of a group IIC intron.^{55,67} In that work, the η – η' interaction locks the D6 branch-point into a position approximately 20 Å from the 5'-splice site, suggesting that the η – η' contact between D6 and D2 can only occur after the first step of splicing (and long after folding).⁶⁸ Thus, the D6 tetraloop appears to interact with a different region of the intron during early stages of intron folding and this same interaction is likely to persist during the first step of splicing.

Conclusions

In this work, NAIM experiments were used to identify functional groups that are important for group II intron precursor folding and to reveal the precise mechanistic role of the Mss116 folding cofactor. These experiments show that Mss116 acts specifically to stabilize a key junction within the heart of intron D1, which acts as a scaffold for assembly of the

entire intron. Proper formation of this junction (the folding control element) sets the stage for all subsequent steps of intron folding. While Mss116 influences folding in a manner similar to that previously reported for the addition of high Mg^{2+} concentrations, it is much more effective at stabilizing the folding control element and it facilitates RNA folding under near-physiological conditions. Importantly, the NAIM method is an unbiased, combinatorial approach that allowed us to examine every nucleotide within the intron and flanking exons and to evaluate whether the protein exerted a stabilizing or destabilizing role at every position. The data clearly show that Mss116 serves to strongly stabilize intronic substructures and the findings are inconsistent with a model in which Mss116 stimulates group II intron folding by acting as a helicase that unwinds misfolded substructures within the intron. In addition to establishing the structural role of Mss116 during folding, these data provide new insights into the folding pathway of intact group II intron precursors. The data reveal unexpected and important roles for the IBS2 region within the 5'-exon and for the terminal tetraloop of intron D6, thereby implicating new tertiary interactions that contribute to intron stability.

Materials and Methods

RNA and protein preparation

The SE, SE 4, and Se 2 4 RNAs were transcribed from plasmids pAS 10, pGD 58, and pGD 59, linearized by HindIII. The RNAs were transcribed in the presence of nucleotide analog phosphorothioates (Glen Research) as previously described.^{47,48} RNA transcripts were purified on a 5% denaturing polyacrylamide gel, followed by 5'- or 3'-end labeling as previously described.⁶⁹ Mss116 was expressed and purified as previously described.²⁸

NAIM in the absence of Mss116

5'-End-labeled RNA (2 nM final concentration) or 3'-end-labeled RNA (20 nM final concentration) was denatured at 95 °C for 1 min in 80 mM Mops-KOH, pH 7.0, and cooled to 30 °C. Then, 100 mM KCl and 8 mM $MgCl_2$ (final concentrations) were added and the resulting mixture was incubated at 30 °C in a final volume of 80 μ l for 40–45 min (until 15–20% of the compact population was formed). Samples were purified on 6% native gels containing 3 mM $MgCl_2$ in 34 mM Tris–66 mM Hepes and 0.1 mM ethylenediaminetetraacetic acid as previously described.^{12,32} Bands corresponding to unfolded, compact, and extended intermediate species were visualized by PhosphoImager (GE Healthcare), eluted from the gel, and subjected to iodine cleavage as previously described.^{12,47} Samples were analyzed on 5%, 8%, or 15% sequencing gels and quantified as previously described⁴⁷ in order to determine nucleotide analog interference or enhancement values.

NAIM in the presence of Mss116

5'-End-labeled RNA (2 nM final concentration) and 3'-end-labeled RNA (20 nM final concentration) were denatured at 95 °C for 1 min in 80 mM Mops-KOH, pH 7.0, and cooled to 30 °C. Then, 100 mM KCl, 8 mM $MgCl_2$, and 20 nM Mss116 (final concentrations) were added and RNA was incubated at 30 °C in a final volume of 80 μ l for 1–3 min (until 15–20% of the compact population was formed). To facilitate native gel purification, we treated

the samples with Proteinase K (8 mg/ml final concentration) for 30 s at 30 °C. Unfolded and compact species were separated on a 6% native gel containing 3 mM MgCl₂ in 34 mM Tris–66 mM Hepes and 0.1 mM ethylenediaminetetraacetic acid as previously described.^{32,48} Bands corresponding to the unfolded and compact species were visualized using a PhosphoImager (GE Healthcare), eluted from the gel, and subjected to iodine cleavage essentially as previously described.^{47,48} Samples were analyzed on 5%, 8%, or 15% sequencing gels and quantified as previously described⁴⁷ in order to determine nucleotide analog interference or enhancement values.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Gaby Drews for technical assistance. Maximillian Bailor and Megan Fitzgerald are thanked for the critical reading of the manuscript. O.F. is a Research Specialist and A.M.P. is an Investigator with Howard Hughes Medical Institute. This work was supported by the National Institutes of Health grant GM50313.

Abbreviations used

NAIM	Nucleotide Analog Interference Mapping
D1	domain 1
2,6-DAP	2,6-diaminopurine.

References

1. Chauhan S, Caliskan G, Briber RM, Perez-Salas U, Rangan P, Thirumalai D, Woodson SA. RNA tertiary interactions mediate native collapse of a bacterial group I ribozyme. *J. Mol. Biol.* 2005; 353:1199–11209. [PubMed: 16214167]
2. Chauhan S, Behrouzi R, Rangan P, Woodson SA. Structural rearrangements linked to global folding pathways of the *Azoarcus* group I ribozyme. *J. Mol. Biol.* 2009; 386:1167–1178. [PubMed: 19154736]
3. Su L, Brenowitz M, Pyle AM. An alternative route for the folding of large RNAs: apparent two-state folding by a group II intron ribozyme. *J. Mol. Biol.* 2003; 334:639–652. [PubMed: 14636593]
4. Su LJ, Waldsich C, Pyle AM. An obligate intermediate along the slow folding pathway of a group II intron ribozyme. *Nucleic Acids Res.* 2005; 33:6674–6687. [PubMed: 16314300]
5. Steiner M, Karunatilaka KS, Sigel RK, Rueda D. Single-molecule studies of group II intron ribozymes. *Proc. Natl Acad. Sci. USA.* 2008; 105:13853–13858. [PubMed: 18772388]
6. Treiber D, Rook M, Zarrinkar P, Williamson JR. Kinetic intermediates trapped by native interactions in RNA folding. *Science.* 1998; 279:1943–1946. [PubMed: 9506945]
7. Treiber D, Williamson JR. Exposing the kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.* 1999; 9:339–345. [PubMed: 10361090]
8. Pan J, Woodson SA. Folding intermediates of a self-splicing RNA: mispairing of the catalytic core. *J. Mol. Biol.* 1998; 280:597–609. [PubMed: 9677291]
9. Pan T, Sosnick TR. Intermediates and kinetic traps in the folding of a large ribozyme revealed by circular dichroism and UV absorbance spectroscopies and catalytic activity. *Nat. Struct. Biol.* 1997; 4:931–938. [PubMed: 9360610]
10. Fang X, Pan T, Sosnick T. Mg²⁺-dependent folding of a large ribozyme without kinetic traps. *Nat. Struct. Biol.* 1999; 6:1091–1095. [PubMed: 10581546]

11. Rangan P, Masquida B, Westhof E, Woodson SA. Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme. *Proc. Natl Acad. Sci. USA*. 2003; 100:1574–1579. [PubMed: 12574513]
12. Waldsich C, Pyle AM. A kinetic intermediate that regulates proper folding of a group II intron RNA. *J. Mol. Biol.* 2008; 375:572–580. [PubMed: 18022197]
13. Ramaswamy P, Woodson SA. Global stabilization of rRNA structure by ribosomal proteins S4, S17 and S20. *J. Mol. Biol.* 2009; 392:666–677. [PubMed: 19616559]
14. Mayerle M, Bellur DL, Woodson SA. Slow formation of stable complexes during coincubation of minimal rRNA and ribosomal protein S4. *J. Mol. Biol.* 2011; 412:453–465. [PubMed: 21821049]
15. Heilman-Miller SL, Woodson SA. Perturbed folding kinetics of circularly permuted RNAs with altered topology. *J. Mol. Biol.* 2003; 328:385–394. [PubMed: 12691747]
16. Sosnick TR. Kinetic barriers and the role of topology in protein and RNA folding. *Protein Sci.* 2008; 17:1308–1318. [PubMed: 18502978]
17. Swisher J, Su LJ, Brenowitz M, Anderson VE, Pyle AM. Productive folding to the native state by a group II intron ribozyme. *J. Mol. Biol.* 2002; 315:297–310. [PubMed: 11786013]
18. Rajkowitsch L, Chen D, Stampfl S, Semrad K, Waldsich C, Mayer O, et al. RNA chaperones. RNA annealers and RNA helicases. *RNA Biol.* 2007; 4:118–130. [PubMed: 18347437]
19. Pan C, Russell R. Roles of DEAD-box proteins in RNA and RNP folding. *RNA Biol.* 2010; 7:667–676. [PubMed: 21045543]
20. Woodson SA. Taming free energy landscapes with RNA chaperones. *RNA Biol.* 2010; 7:677–686. [PubMed: 21045544]
21. Diges CM, Uhlenbeck OC. *Escherichia coli* DbpA is an RNA helicase that requires hairpin 92 of 23S rRNA. *EMBO J.* 2001; 20:5503–5512. [PubMed: 11574482]
22. Karginov FV, Uhlenbeck OC. Interaction of *Escherichia coli* DbpA with 23S rRNA in different functional states of the enzyme. *Nucleic Acids Res.* 2004; 32:3028–3032. [PubMed: 15173385]
23. Henn A, Cao W, Hackney DD, De La Cruz EM. The ATPase cycle mechanism of the DEAD-box rRNA helicase DbpA. *J. Mol. Biol.* 2008; 377:193–205. [PubMed: 18237742]
24. Mayer O, Waldsich C, Grossberger R, Schroeder R. Folding of the td pre-RNA with the help of the RNA chaperone StpA. *Biochem. Soc. Trans.* 2002; 30:1175–1180. [PubMed: 12440999]
25. Yang Q, Jankowsky E. ATP- and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1. *Biochemistry.* 2005; 44:13591–13601. [PubMed: 16216083]
26. Vogel J, Luisi BF. Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* 2011; 9:578–589. [PubMed: 21760622]
27. Bokinsky G, Nivon LG, Liu S, Chai G, Hong M, Weeks KM, Zhuang X. Two distinct binding modes of a protein cofactor with its target RNA. *J. Mol. Biol.* 2006; 361:771–784. [PubMed: 16872630]
28. Solem A, Zingler N, Pyle AM. A DEAD protein that activates intron self-splicing without unwinding RNA. *Mol. Cell.* 2006; 24:611–617. [PubMed: 17188036]
29. Zingler N, Solem A, Pyle AM. Dual roles for the Mss116 cofactor during splicing of the ai5 {gamma} group II intron. *Nucleic Acids Res.* 2010; 38:6602–6609. [PubMed: 20554854]
30. Fedorova O, Solem A, Pyle AM. Protein-facilitated folding of group II intron ribozymes. *J. Mol. Biol.* 2010; 397:799–813. [PubMed: 20138894]
31. Weeks KM. Protein-facilitated RNA folding. *Curr. Opin. Struct. Biol.* 1997; 7:336–342. [PubMed: 9204274]
32. Fedorova O, Waldsich C, Pyle AM. Group II intron folding under near-physiological conditions: collapsing to the near-native state. *J. Mol. Biol.* 2007; 366:1099–1114. [PubMed: 17196976]
33. Huang HR, Rowe CE, Mohr S, Jiang Y, Lambowitz AM, Perlman PS. The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. *Proc. Natl Acad. Sci. USA.* 2005; 102:163–168. [PubMed: 15618406]
34. Halls C, Mohr S, Del Campo M, Yang Q, Jankowsky E, Lambowitz AM. Involvement of DEAD-box proteins in group I and group II intron splicing. biochemical characterization of Mss116p,

- ATP hydrolysis-dependent and -independent mechanisms, and general RNA chaperone activity. *J. Mol. Biol.* 2007; 365:835–855. [PubMed: 17081564]
35. Karunatilaka KS, Solem A, Pyle AM, Rueda D. Single-molecule analysis of Mss 116-mediated group II intron folding. *Nature*. 2010; 467:935–939. [PubMed: 20944626]
36. Potratz JP, Del Campo M, Wolf RZ, Lambowitz AM, Russell R. ATP-dependent roles of the DEAD-box protein Mss 116p in group II intron splicing *in vitro* and *in vivo*. *J. Mol. Biol.* 2011; 411:661–679. [PubMed: 21679717]
37. Del Campo M, Tijerina P, Bhaskaran H, Mohr S, Yang Q, Jankowsky E, et al. Do DEAD-box proteins promote group II intron splicing without unwinding RNA? *Mol. Cell.* 2007; 28:159–166. [PubMed: 17936712]
38. Del Campo M, Mohr S, Jiang Y, Jia H, Jankowsky E, Lambowitz AM. Unwinding by local strand separation is critical for the function of DEAD-box proteins as RNA chaperones. *J. Mol. Biol.* 2009; 389:674–693. [PubMed: 19393667]
39. Linder P, Jankowsky E. From unwinding to clamping—the DEAD box RNA helicase family. *Nat. Rev. Mol. Cell Biol.* 2011; 12:505–516. [PubMed: 21779027]
40. Cordin O, Banroques J, Tanner NK, Linder P. The DEAD-box protein family of RNA helicases. *Gene*. 2006; 367:17–37. [PubMed: 16337753]
41. Linder P. Dead-box proteins: a family affair—active and passive players in RNP-remodeling. *Nucleic Acids Res.* 2006; 34:4168–4180. [PubMed: 16936318]
42. Bhaskaran H, Russell R. Kinetic redistribution of native and misfolded RNAs by a DEAD-box chaperone. *Nature*. 2007; 449:1014–1018. [PubMed: 17960235]
43. Liu F, Putnam A, Jankowsky E. ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. *Proc. Natl Acad. Sci. USA.* 2008; 105:20209–20214. [PubMed: 19088201]
44. Strobel SA, Shetty K. Defining the chemical groups essential for *Tetrahymena* group I intron function by nucleotide analog interference mapping (NAIM). *Proc. Natl Acad. Sci. USA.* 1997; 94:2903–2908. [PubMed: 9096319]
45. Ryder SP, Strobel SA. Nucleotide analog interference mapping. *Methods*. 1999; 18:38–50. [PubMed: 10208815]
46. Ryder SP, Ortoleva-Donnelly L, Kosek AB, Strobel SA. Chemical probing of RNA by nucleotide analog interference mapping. *Methods Enzymol.* 2000; 317:92–109. [PubMed: 10829274]
47. Fedorova, O.; Boudvillain, M.; Kawaoka, J.; Pyle, AM. Nucleotide analog interference mapping and suppression: specific applications in studies of RNA tertiary structure, dynamic helicase mechanism and RNA–protein interactions. In: Hartmann, RK.; Bindereif, A.; Schoen, A.; Westhof, E., editors. *Handbook of RNA Biochemistry*. Vol. 1. Wiley-VCH; Weinheim: 2005. p. 259–293.
48. Waldsich C, Pyle AM. A folding control element for tertiary collapse of a group II intron ribozyme. *Nat. Struct. Mol. Biol.* 2007; 14:37–44. [PubMed: 17143279]
49. Cao W, Coman MM, Ding S, Henn A, Middleton ER, Bradley MJ, et al. Mechanism of Mss 116 ATPase reveals functional diversity of DEAD-box proteins. *J. Mol. Biol.* 2011; 409:399–414. [PubMed: 21501623]
50. Michel F, Umesono K, Ozeki H. Comparative and functional anatomy of group II catalytic introns—a review. *Gene*. 1989; 82:5–30. [PubMed: 2684776]
51. Michel F, Ferat J-L. Structure and activities of group II introns. *Annu. Rev. Biochem.* 1995; 64:435–461. [PubMed: 7574489]
52. Qin PZ, Pyle AM. The architectural organization and mechanistic function of group II intron structural elements. *Curr. Opin. Struct. Biol.* 1998; 8:301–308. [PubMed: 9666325]
53. Pyle, AM.; Lambowitz, AM. Group II introns: ribozymes that splice RNA and invade DNA. In: Gesteland, RF.; Cech, TR.; Atkins, JF., editors. *The RNA World*. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, New York: 2006. p. 469–506.
54. Xiang Q, Qin PZ, Michels WJ, Freeland K, Pyle AM. The sequence-specificity of a group II intron ribozyme: multiple mechanisms for promoting unusually high discrimination against mismatched targets. *Biochemistry*. 1998; 37:3839–3849. [PubMed: 9521704]

55. Toor N, Keating KS, Taylor SD, Pyle AM. Crystal structure of a self-spliced group II intron. *Science*. 2008; 320:77–82. [PubMed: 18388288]
56. Toor N, Keating KS, Fedorova O, Rajashankar K, Wang J, Pyle AM. Tertiary architecture of the *Oceanobacillus iheyensis* group II intron. *RNA*. 2010; 16:57–69. [PubMed: 19952115]
57. Ortoleva-Donnelly L, Sweczak AA, Gutell RR, Strobel SA. The chemical basis of adenosine conservation throughout the *Tetrahymena* ribozyme. *RNA*. 1998; 4:498–519. [PubMed: 9582093]
58. Ryder SP, Strobel SA. Nucleotide analog interference mapping of the hairpin ribozyme: implications for secondary and tertiary structure formation. *J. Mol. Biol.* 1999; 291:295–311. [PubMed: 10438622]
59. Boudvillain M, Pyle AM. Defining functional groups, core structural features and interdomain tertiary contacts essential for group II intron self-splicing: a NAIM analysis. *EMBO J.* 1998; 17:7091–7104. [PubMed: 9843513]
60. Soukup J, Minakawa N, Matsuda A, Strobel SA. Identification of A-minor tertiary interactions within a bacterial group I intron active site by 3-deazaadenosine interference mapping. *Biochemistry*. 2002; 41:10426–10438. [PubMed: 12173929]
61. Suydam IT, Strobel SA. Fluorine substituted adenosines as probes for nucleobase protonation in functional RNAs. *J. Am. Chem. Soc.* 2008; 130:13639–13648. [PubMed: 18803382]
62. Chanfreau G, Jacquier A. An RNA conformational change between the two chemical steps of group II self-splicing. *EMBO J.* 1996; 15:3466–3476. [PubMed: 8670849]
63. Ramaswamy P, Woodson SA. S16 throws a conformational switch during assembly of 30S 5' domain. *Nat. Struct. Mol. Biol.* 2009; 16:438–445. [PubMed: 19343072]
64. Adilakshmi T, Ramaswamy P, Woodson SA. Protein-independent folding pathway of the 16S rRNA 5' domain. *J. Mol. Biol.* 2005; 351:508–519. [PubMed: 16023137]
65. Fedorova O, Su LJ, Pyle AM. Group II introns: highly specific endonucleases with modular structures and diverse catalytic functions. *Methods*. 2002; 28:323–335. [PubMed: 12431436]
66. Toor N, Hausner G, Zimmerly S. Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA*. 2001; 7:1142–1152. [PubMed: 11497432]
67. Li CF, Costa M, Michel F. Linking the branchpoint helix to a newly found receptor allows lariat formation by a group II intron. *EMBO J.* 2011; 30:3040–3051. [PubMed: 21712813]
68. Pyle AM. The tertiary structure of group II introns: implications for biological function and evolution. *Crit. Rev. Biochem. Mol. Biol.* 2010; 45:215–232. [PubMed: 20446804]
69. Pyle AM, Green JB. Building a kinetic framework for group II intron ribozyme activity: quantitation of interdomain binding and reaction rate. *Biochemistry*. 1994; 33:2716–2725. [PubMed: 8117737]

Fig. 1.

(a) A general schematic of the NAIM experiment, in which the selection step is RNA folding in the presence or in the absence of Mss116. (b) Nucleotide analog phosphorothioates used in the NAIM experiment to identify functional groups important for protein-assisted group II intron folding. (c) A schematic of expected NAIM effects consistent with the destabilizing or stabilizing function of the protein.

Fig. 2.

Representative native gels showing the formation of the compact species (NAIM selection step) in the presence and in the absence of Mss116 by the SE RNA (a) and SE $\Delta 2\Delta 4$ RNA (b). Both RNAs were transcribed in the presence of 2,6-DAP.

Fig. 3.

(Top) Representative sequencing gels showing functional groups in the 5'-exon, which interfere with compaction. On each gel, lane U corresponds to unfolded species, I corresponds to extended intermediate species, and C corresponds to compact species. Symbols on the left of each gel indicate nucleotide positions that interfere with the formation of the compact (red) or intermediate (green) species. (Bottom) A bar graph showing average value and standard deviation for interference effects observed in the 5'-exon upon formation of the compact and intermediate states in the absence of Mss116 and compact state in the presence of Mss116.

Fig. 4.

Functional groups important for compaction of the SE 2-4 RNA in the absence (a) and in the presence (b) Mss116 and ATP under near-physiological conditions. Insets, functional groups in full-length domains 2 and 4, which important for compaction of SE 4 and SE RNAs, respectively, in the absence (a) and in the presence (b) of Mss116 and ATP.

Fig. 5.

(a) Representative sequencing gels showing functional groups in the 5'-end of D1, which interfere with intron compaction. On each gel, lane U corresponds to unfolded species, I corresponds to extended intermediate species, and C corresponds to compact species. Symbols on the left of each gel indicate nucleotide positions that interfere with the formation of the compact (red) or intermediate (green) species. Respective substructures in D1 are denoted in gray on the left. (b) A bar graph showing average value and standard deviation for interference effects observed in the 5'-end of D1 upon formation of the compact state in the absence and in the presence of Mss116.

Fig. 6.

(Top) Representative sequencing gels showing functional groups interfering with compaction at adenosine (a) and guanosine (b) positions in the κ - ζ element. On each gel, lane U corresponds to unfolded species, I corresponds to extended intermediate species, and C corresponds to compact species. Symbols on the left of each gel indicate nucleotide positions that interfere with (filled symbols) or enhance (open symbols) the formation of the compact (red) or intermediate (green) species. Black arrows indicate positions where NAIM effects are increased in the presence of Mss116. Positions where NAIM effects appear only in the presence of Mss116 are denoted by blue arrows. Respective substructures in D1 are denoted in gray on the left. Bar graphs on the bottom (a) and on the right (b) show average value and standard deviation for interference and enhancement effects observed in the κ - ζ element upon formation of the compact state in the absence and in the presence of Mss116. (c) A schematic of functional groups exhibiting increased interference or enhancement effects in the presence of Mss116.

Fig. 7.

(Top) Representative sequencing gels showing functional groups in D2, D3, and D4, which affect the intron compaction. On each gel, lane U corresponds to unfolded species, I corresponds to extended intermediate species, and C corresponds to compact species. Symbols on the left of each gel indicate nucleotide positions that interfere with (filled symbols) or enhance (open symbols) the formation of the compact (red) or intermediate (green) species. (Bottom) A bar graph showing average value and standard deviation for interference effects observed in the D2, D3, and D4 upon formation of the compact state in the absence and in the presence of Mss116.

Fig. 8.

Comparison of NAIM effects in D6 in the full-length (SE) and truncated (SE $\Delta 2-4$) RNAs. (Top) Representative sequencing gels showing functional groups in D6, which affect the intron compaction. On each gel, lane U corresponds to unfolded species, I corresponds to extended intermediate species, and C corresponds to compact species. Symbols on the left of each gel indicate nucleotide positions that interfere with the formation of the compact species. (Bottom) A bar graph showing average value and standard deviation for interference effects observed in D6 upon formation of the compact state in the absence and in the presence of Mss116 in the full-length and truncated self-splicing intron RNAs.

Fig. 9.

(a) Scheme illustrating the mechanism of Mss116-promoted folding of the ai5 γ group II intron. (b) Scheme of a hypothetical free energy diagram for the ai5 γ intron folding in the presence of Mss116.