nature
structural &
molecular biology

Evolution from DNA to RNA recognition by the bI3 LAGLIDADG maturase

Antonella Longo[1,2,3], Christopher W Leonard[1,3], Gurminder S Bassi[1,3], Daniel Berndt[1], Joseph M Krahn[2], Traci M Tanaka Hall[2] & Kevin M Weeks[1]

LAGLIDADG endonucleases bind across adjacent major grooves via a saddle-shaped surface and catalyze DNA cleavage. Some LAGLIDADG proteins, called maturases, facilitate splicing by group I introns, raising the issue of how a DNA-binding protein and an RNA have evolved to function together. In this report, crystallographic analysis shows that the global architecture of the bI3 maturase is unchanged from its DNA-binding homologs; in contrast, the endonuclease active site, dispensable for splicing facilitation, is efficiently compromised by a lysine residue replacing essential catalytic groups. Biochemical experiments show that the maturase binds a peripheral RNA domain 50 Å from the splicing active site, exemplifying long-distance structural communication in a ribonucleoprotein complex. The bI3 maturase nucleic acid recognition saddle interacts at the RNA minor groove; thus, evolution from DNA to RNA function has been mediated by a switch from major to minor groove interaction.

Nature often adapts one structural scaffold for a new purpose. One persistent example of this process is the co-opting of existing proteins to function as cofactors in group I intron RNA splicing reactions. Group I introns are widespread in biology and fold into a well-defined tertiary structure[1–4]. A few prominent group I introns self splice either under roughly physiological conditions or at elevated Mg^{2+} concentrations[1]. However, like cellular RNAs in general, most group I introns seem to require the assistance of protein cofactors to fold stably to an active conformation[5–8]. Protein cofactors recruited by group I introns include functioning tRNA synthetases[8,9], DNA homing endonucleases[10–12], homologs of DNA junction resolvases[7] and proteins with no known alternate function[13,14].

The recruitment of proteins that function with RNA substrates as group I intron splicing factors suggests that some group I intron RNAs may contain or have evolved structures similar to the original substrates of these enzymes. Alternatively, some nucleic acid–binding proteins seem to have added auxiliary domains or interaction surfaces specific for binding group I intron substrates[9,15,16]; these cases suggest that the protein has evolved to accommodate the new function. The recruitment of DNA homing endonuclease and DNA junction–resolving proteins as RNA splicing factors raises the questions of how these proteins have evolved to bind group I intron RNAs and whether the introns have evolved RNA sites with global features similar to those of double-stranded or branched DNA.

The largest class of DNA-binding protein that has been co-opted to stabilize RNA folding is the LAGLIDADG family of homing endonucleases[17,18], named for a loosely conserved sequence motif. LAGLIDADG homing endonucleases are encoded by mobile genetic elements and catalyze DNA cleavage at specific sites lacking the endonuclease coding region[18]. LAGLIDADG proteins are either homodimers or monomeric pseudo dimers. The two copies of the LAGLIDADG motif form, in part, either the dimer interaction surface or the interdomain interaction surface. A subset of LAGLIDADG proteins facilitate splicing by group I introns and are called maturases[10]. Group I intron maturase proteins are translated from an open reading frame within the intron that is inserted at a peripheral position that does not disrupt RNA folding, and they are typically monomeric[17]. The sequences encoding LAGLIDADG maturases are thought to have evolved from sequences encoding homing endonucleases that were then inserted into group I introns[19]. Once the open reading frame has been inserted, the maturase, the intron RNA or both then adapt to form a stable ribonucleoprotein (RNP) complex.

In a distinctive elaboration of the maturase paradigm, the third RNA intron of the *Saccharomyces cerevisiae* cytochrome *b* gene (the bI3 group I intron) requires two proteins in order to fold into a stable, catalytically active conformation: the intron-encoded bI3 LAGLIDADG maturase and the nuclear-encoded Mrs1 protein[7,20]. The active 420-kDa bI3 RNP consists of the intron RNA, a bI3 maturase monomer and two Mrs1 dimers. In *S. cerevisiae*, the bI3 maturase no longer functions as a DNA endonuclease, now apparently facilitating only the splicing of bI3 intron RNA[7].

To define the pathway of structural evolution that has transformed a DNA-binding protein into an RNA cofactor and, coincidently, a self-splicing RNA into an obligatory RNP enzyme, we determined the crystal structure of the LAGLIDADG domain of the bI3 maturase. We then defined the RNA recognition site for this maturase using

[1]Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290, USA. [2]Laboratory of Structural Biology, National Institute of Environmental Health Sciences, US National Institutes of Health, Research Triangle Park, North Carolina 27709, USA. [3]These authors contributed equally to this work. Correspondence should be addressed to K.M.W. (weeks@unc.edu) or T.M.T.H. (hall4@niehs.nih.gov).

Published online 21 August 2005; doi:10.1038/nsmb976

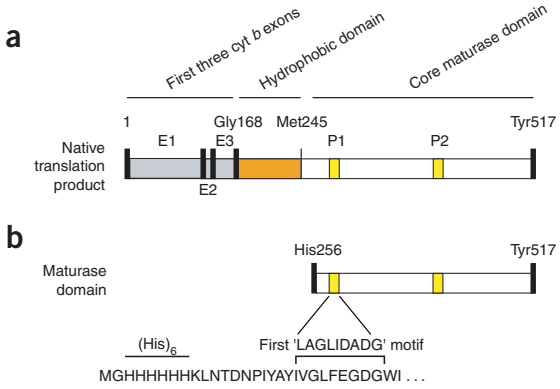


Figure 1 Translation of the bI3 maturase holoprotein and core maturase domain. (**a**) Diagram of bI3 intron. The bI3 maturase is translated from the bI3 intron and is in frame with the first three cytochrome *b* exons (E1, E2, E3; gray). The intron-encoded portion spans hydrophobic (orange) and core maturase (white) domains. Yellow, LAGLIDADG motifs (P1 and P2). (**b**) Diagram of the core maturase domain expressed for structural studies.

chemical probing experiments. The global architecture of the bI3 maturase remains unchanged from its DNA-cleaving homologs, and the maturase binds at a distal, nonconserved structure in the P5-P4-P6 domain of the intron. Functional evolution by the bI3 LAGLIDADG maturase has been accompanied by a switch from nucleic acid recognition at the DNA major groove to recognition at the RNA minor groove.

RESULTS

Structure determination

The bI3 maturase is translated in frame with three upstream cytochrome *b* exons[21] and with a hydrophobic domain[7] encoded by the intron (**Fig. 1a**). The core LAGLIDADG-motif maturase domain, in concert with the Mrs1 protein cofactor, is sufficient to facilitate splicing of the bI3 intron *in vitro*[7]. We therefore determined the three-dimensional structure of a 262-residue C-terminal fragment (residues His256–Tyr517) spanning the region required for RNA-based catalysis, expressed with an N-terminal $(His)_6$ affinity tag (**Fig. 1b**). There are two maturase molecules in the asymmetric unit. Although we did not use noncrystallographic symmetry restraints, the A and B molecules superimpose with an r.m.s. deviation of 0.38 Å over 256 Cα positions. Our analysis focuses on the nearly complete and better-ordered model of the A molecule.

Maturase architecture

The bI3 maturase has a mixed α-helix–β-sheet topology similar to that of other LAGLIDADG proteins whose structures are known (**Fig. 2a**)[18,22] and strongly resembles a saddle (**Fig. 2b**). The structure is formed from similar N- and C-terminal domains related by pseudo two-fold symmetry (**Fig. 2a**). LAGLIDADG motifs in the first α-helix (α1) of each domain (**Fig. 2**, yellow) form a tightly packed inter-domain interface on one side and participate in the hydrophobic core of each domain on the other side (**Fig. 2c**). The four core α helices (α1–α4) in each bI3 maturase domain pack against a four-stranded anti-parallel β-sheet (**Fig. 2a**, green) to form the large saddle-shaped putative nucleic acid–binding surface (**Fig. 2b**).

Residues downstream of α4 fold differently in the two domains. In the N-terminal domain, a long linker runs parallel to α4, bends to form a short α-helix and then connects to the C-terminal domain (**Fig. 2a**, magenta). In the C-terminal domain, residues following α4 form a short α-helix (α5) that makes hydrophobic interactions with α3.

Conserved interactions in the LAGLIDADG motifs include those between small side chains that allow tight packing of the helices and extend to the residues just before the LAGLIDADG sequences (at the –1 positions), Tyr267 and Trp404. Tyr267 and Trp404 each form a bridging hydrogen bond with the sixth LAGLIDADG residue of the opposite helix (Glu410 and Glu273, respectively; **Fig. 2c**). This cross-domain bridging interaction provides a structural explanation for the conservation of an acidic residue at position 6 in the LAGLIDADG motif. The (−1)-6 bridging interaction, involving acidic and hydrogen-bonding aromatic residues, is observed in a large subset of LAGLIDADG motif proteins, including I-CreI, PI-PfuI, PI-SceI and I-AniI[17]. Inspection of a previous alignment[17] shows that one or both of the (−1)-6 interactions can be replaced by a hydrophobic interaction if the residues at positions −1 and 6 are coordinately changed to hydrophobic groups. For example, I-DmoI[23] contains two examples of the hydrophobic version of the cross-domain interaction.

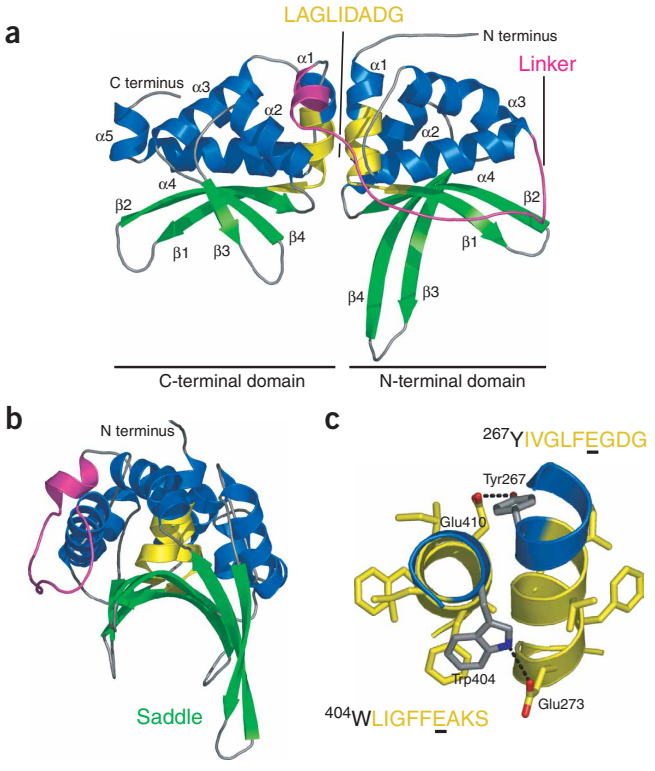


Figure 2 Overview of the bI3 maturase structure. (**a**,**b**) Two orientations of the maturase structure related by an ~180° rotation about the vertical axis. Blue, the α-helical structural core; yellow, LAGLIDADG motifs; green, the four antiparallel β-strands in each domain comprising the saddle. Secondary structure elements are numbered for each domain as for the I-CreI homodimer[25], starting with the LAGLIDADG helices as α1. The N- and C-terminal domains have almost identical architectures (r.m.s. deviation of 1.68 Å over 91 Cα atoms in the core α-helical and central β-strand regions). (**c**) LAGLIDADG motif and the (–1)-6 interaction. Black dashed lines indicate the proposed conserved cross-domain interaction between a hydrogen-bonding aromatic residue at position –1 (gray) and a conserved aspartic or glutamic acid residue at position 6.

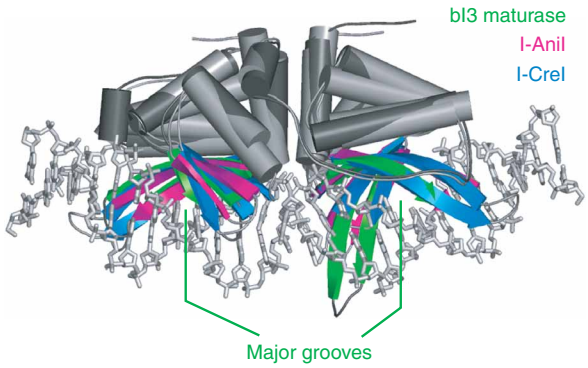


Figure 3 Conserved shape complementarity between LAGLIDADG proteins and consecutive DNA major grooves. The structures of the bI3 maturase, I-AniI[15] and I-CreI[25] are superimposed and β-sheet structures are colored as noted. Gray, core α-helices; dark gray, α-helices for the bI3 maturase; stick representation of the DNA structure is from the I-AniI structure. Overall r.m.s. deviations: for the bI3 maturase and I-AniI, 1.01 Å over 237 Cα positions; for the bI3 maturase and I-CreI, 1.61 Å over 187 Cα positions.

A conserved nucleic acid–binding surface

The bI3 maturase structure represents the first structure of a LAGLIDADG protein that functions as an RNA-splicing factor but has lost its endonuclease function[7]. Despite the substantially different present biological function of the bI3 maturase, the core protein architecture is identical to that of LAGLIDADG proteins that function to cleave DNA substrates (**Fig. 3**).

The bI3 maturase structure is very similar to that of the homologous I-AniI maturase-endonuclease, even though the I-AniI structure was determined with a cleaved DNA product[15]. The major differences in backbone structures for the bI3 maturase and I-AniI are in the positions of the third and fourth β-strands in the C-terminal domain and in the longer lengths of the third and fourth β-strands in the N-terminal domain. Structural conservation is also maintained between the bI3 maturase and other LAGLIDADG endonucleases, such as I-CreI, that function as dimers of identical subunits[24,25].

Structural alignments with these DNA-binding homologs suggest that the β-sheet saddle of the RNA-binding bI3 maturase remains, in principle, competent to bind snugly across adjacent DNA major grooves (**Fig. 3**). In addition, 17 of 23 residues that interact with DNA in the I-AniI structure[15] are conserved or similar in the bI3 maturase. We measured binding of the bI3 maturase to the native DNA junction sequence using both direct binding experiments and competition experiments. Equilibrium dissociation constants (K_ds) measured by these two approaches are 19 ± 4 and 8 ± 3 nM, respectively (data not shown), and are similar to the K_d of I-AniI with target DNA (8–13 nM)[15,26]. Thus, the nucleic acid–interacting surface in the bI3 maturase has not been altered markedly to allow recognition of an RNA target.

The (in)active site

LAGLIDADG endonuclease sequences are thought to be parasitic genes[19,27] that self-propagate by encoding proteins that cleave chromosome sites lacking an endonuclease coding region. Recombinatorial repair using the endonuclease-coding DNA as a template yields a new copy of the endonuclease gene. This process is called homing. Potential costs of homing to the host include biosynthesis of the endonuclease protein and the possibility of compromising an essential gene. These costs are presumably counteracted if a parasitic endonuclease becomes co-opted for another function, such as facilitating splicing by group I intron RNAs.

The bI3 maturase has responded to the dual pressures against homing by evolving so that it no longer functions as a homing endonuclease[7]. For proteins active as DNA endonucleases, the penultimate residues in each of the two LAGLIDADG motifs are acidic residues that together bind three catalytic metal ions[22,28]. The absence of DNA endonuclease activity in the bI3 maturase is readily explained by the substitution of Lys412 for aspartic acid (Asp′) at the eighth residue of the second LAGLIDADG motif (**Fig. 4**), resulting in what we term an (in)active site. Many amino acid mutations can inactivate a LAGLIDADG endonuclease[11,28,29]; but here, the single lysine substitution in the bI3 maturase effectively replaces the catalytically essential aspartyl-ion ligand interaction.

When the bI3 maturase (in)active site is aligned with the active site of I-CreI, Lys412 lies between the positions of two of the three catalytic metal ions in the I-CreI structure and forms a salt bridge with the 'metal-binding' Asp275 from the first LAGLIDADG motif, probably stabilizing this region of the protein (**Fig. 4**, black dashed line). In addition, there are two water molecules in the bI3 maturase (in)active site, one of which occupies a position similar to that of the third catalytic divalent metal ion in functional LAGLIDADG endonucleases. Both water molecules lie within hydrogen bonding distance of conserved groups; together with the interactions observed for Lys412, this observation emphasizes that the (in)active site in the bI3 maturase remains well structured (**Fig. 4**). Thus, the endonuclease activity, not required for an RNA splicing cofactor, is eliminated while the overall protein architecture is preserved.

bI3 maturase binds the P5-P4-P6 group I intron domain

We used two classes of hydroxyl radical footprinting experiments, solvent-based and site-directed, to identify the RNA interaction site for the bI3 maturase. We first used solvent-based hydroxyl radical footprinting to map the global consequences of maturase binding to the bI3 intron RNA (**Fig. 5**). In this technique, the RNA backbone is cleaved at solvent-accessible positions independent of sequence[30].

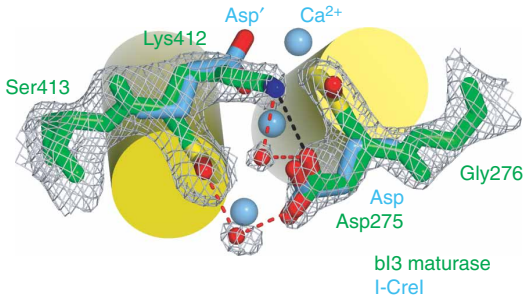


Figure 4 Superposition of the bI3 maturase (in)active site with catalytic residues and divalent ions from the I-CreI DNA endonuclease–DNA substrate complex. View is from the nucleic acid–binding surface and looking down the LAGLIDADG helices (yellow). Green, last two residues of the bI3 maturase LAGLIDADG motifs; light blue, I-CreI[25] aspartic acid residues and three catalytic metal ions (Ca^{2+}); red dashed lines connect atoms located within hydrogen bonding distance (≤3.2 Å); black dashed line indicates an electrostatic interaction between Asp275 and Lys412; red and dark blue indicate oxygen and nitrogen atoms near the (in)active site, respectively; white mesh is a maturase simulated annealing $2|F_o|-|F_c|$ omit electron density map[43] contoured at 1σ that excluded the area shown.

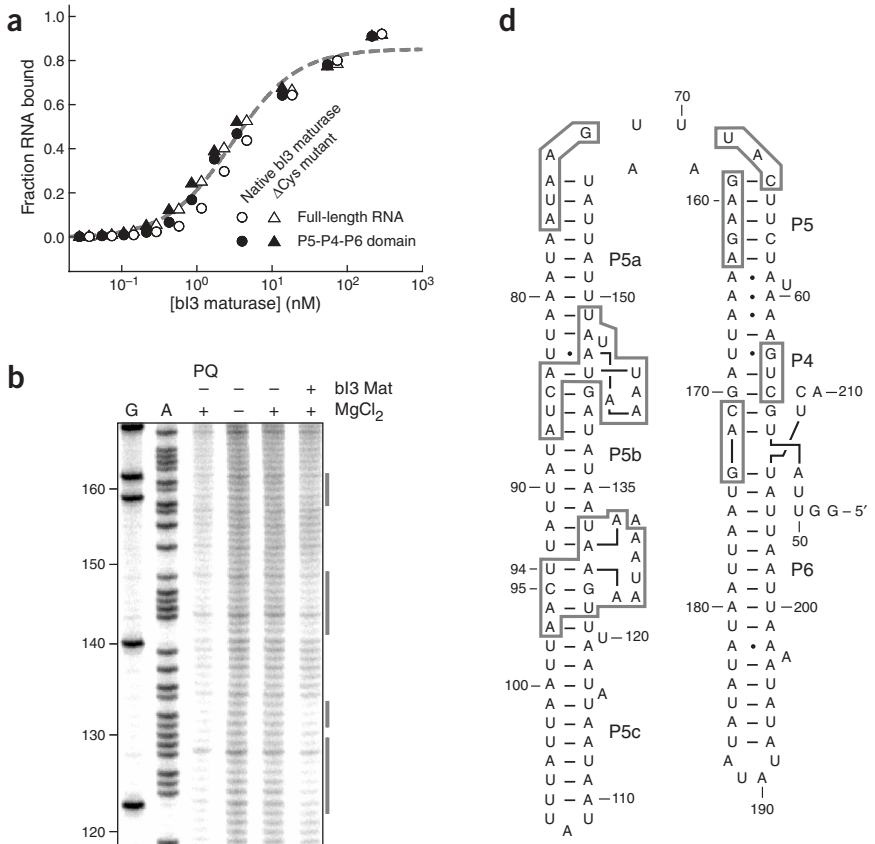


Figure 5 Solvent-based hydroxyl radical footprinting of the bI3 maturase–P5-P4-P6 domain complex. (**a**) RNA binding data showing that the bI3 maturase binds with equivalent affinities to the intact bI3 intron[7] and to the P5-P4-P6 domain. Dashed line shows a best fit to: fraction RNA bound = [maturase]/([maturase] + K_d). K_ds are 2–5 ± 0.5 nM in all cases, as measured in at least three experiments per complex. Open and closed symbols are offset slightly for visual clarity. (**b**) Representative hydroxyl radical footprinting data for bI3 maturase binding to 5′ ^{32}P end-labeled P5-P4-P6 domain RNA. If present (+), the $MgCl_2$ concentration is 20 mM; PQ indicates a prequench control in which the thiourea/urea stop solution was added before cleavage reagents. G and A markers were generated by iodine-mediated cleavage of phosphorothioate-substituted RNAs. Every tenth nucleotide in the P5-P4-P6 domain is marked. (**c**) Quantitative histogram of hydroxyl radical protection upon maturase binding. Protection factors ($I - I_{maturase}$) greater than three-fold above the average background (dashed line) are scored as substantial protections. Asterisks indicate two positions at which background bands precluded quantitative analysis. (**d**) Summary of protection data superimposed on a secondary structure model of the P5-P4-P6 domain RNA. Regions protected from cleavage upon binding by the bI3 maturase are shown with gray bars in **b** and gray boxes in **d**.

Group I introns fold into two approximately coaxially stacked domains[2], the P5-P4-P6 and P9-P7-P3-P8 domains (where each P*x* denotes one of the helices that together comprise a domain). These pack together to form a cleft into which the splice site helix docks and where splicing occurs[1,3,4]. In exploratory experiments with a 541-nucleotide (nt) splicing-competent RNA spanning the entire bI3 intron[7], the bI3 RNA showed little or no protection from hydroxyl radical cleavage in the absence of the maturase and Mrs1 proteins. This was consistent with the observation that these cofactors are required for splicing of this RNA[7,20]. Upon binding by the bI3 maturase, extensive protection from hydroxyl radical cleavage was observed only in the P5-P4-P6 domain and in the P3 helix and J3/4 linker region (data not shown).

Prompted by these results, we examined whether the maturase binds independently to the P5-P4-P6 domain. The bI3 maturase binds indistinguishably to either the entire intron RNA or a 160-nt RNA spanning the P5-P4-P6 domain alone (**Fig. 5a**). We therefore focused on understanding the interaction between the maturase and this 160-nt RNA domain. Hydroxyl radical cleavage experiments using 5′ end-labeled RNA show that the P5-P4-P6 domain is accessible to solvent in both the absence and presence of 20 mM Mg^{2+} (compare – and +$MgCl_2$ lanes with the prequench (PQ) lane in **Fig. 5b**). Upon addition of the bI3 maturase, extensive regions in the P5-P4-P6 domain are protected from solvent-based hydroxyl radical cleavage (compare – and +bI3 maturase lanes in **Fig. 5b**). Regions specifically protected from hydroxyl radical cleavage were determined by integrating[31] band intensities at almost every nucleotide in the RNA and then subtracting the net cleavage intensity observed upon binding by the maturase (**Fig. 5c**). The overall hydroxyl radical protection pattern is essentially identical to that observed for footprinting of the maturase in the context of the complete intron (data not shown). The protection pattern is also very similar to that observed for folding of the homologous *Tetrahymena thermophila* group I intron domain in the conserved P4, P5 and P5a regions[32,33]. Cleavage protection reflects both direct maturase-RNA interactions and also protein-induced RNA-RNA interactions. Thus, the similarity with the protection pattern of the folded *Tetrahymena* intron suggests this part of the protection pattern is due to RNA-RNA interactions induced upon folding the bI3 intron. The bI3 intron protection pattern additionally includes protection of regions in P5b and P5c. We infer that the bI3

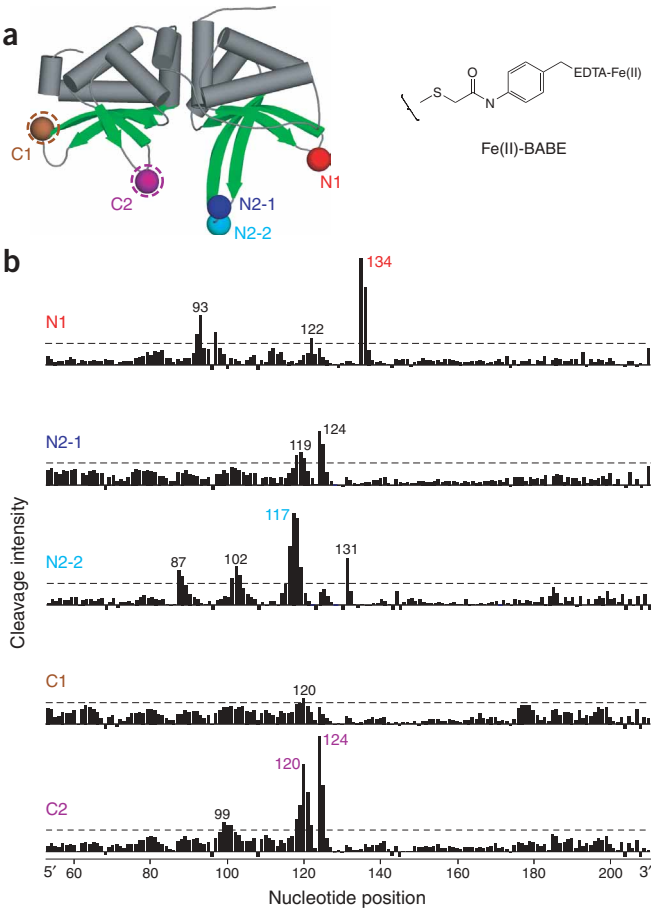


Figure 6 Site-directed hydroxyl radical footprinting of the bI3 maturase–P5-P4-P6 domain complex. (**a**) Sites of unique cysteine substitutions (spheres, left) used to attach the Fe(II)-BABE cleavage agent (right). The C1 and C2 sites (dashed circles) are located in loop structures that have high crystallographic *B* factors (60–100 Å^2) and are relatively flexible. (**b**) Histograms of absolute cleavage intensity minus background for maturase protein variants. Sites of specific medium-to-weak or strong cleavage are numbered; strong cleavages are emphasized in color; dashed lines, threshold at which cleavage intensity is three-fold over the average background (calculated from reactions of the ΔCys protein treated with Fe(II)-BABE).

maturase functions by interacting with and stabilizing the structure of the P5-P4-P6 domain, including the nonconserved P5b and P5c structures (**Fig. 5a**,**d**).

RNA interaction site for the bI3 maturase

We next used site-directed hydroxyl radical cleavage to refine the binding site of the bI3 maturase on the P5-P4-P6 RNA domain, focusing on the interaction of the saddle-shaped nucleic acid–binding surface of the maturase with the RNA. Fe(II)-EDTA groups were tethered to the maturase via linkages with the Fe(II)-BABE reagent[34] at unique solvent-exposed cysteine residues (**Fig. 6a**). Three cysteines were individually introduced into the N-terminal maturase domain (G284C, M322C and G325C; mutants N1, N2-1 and N2-2, respectively) and two into the C-terminal domain (M425C and F459C; mutants C1 and C2). The loops connecting the β-strands in the C-terminal domain have high *B* factors in the crystal structure (**Fig. 6a**, dashed circles), suggesting these regions are more flexible than reagent attachment sites in the N-terminal domain.

Complexes were formed with the Fe(II)-BABE–derivatized single-cysteine maturase proteins or a mock-derivatized maturase protein lacking cysteine residues (ΔCys) and 5′ end-labeled P5-P4-P6 domain RNA. RNA cleavage was induced by adding hydrogen peroxide and ascorbic acid. No specific cleavage was observed for the ΔCys variant, relative to controls in which either the Fe(II)-BABE or the cleavage reagents were omitted (data not shown). Reproducible cleavage patterns were obtained for each of the five derivatized proteins (**Fig. 6b**). Four of the five proteins yielded substantial cleavages in the P5b and P5c helices or in the asymmetric loop that joins these two helices (**Fig. 7a**). Although the C1 protein bound tightly to the RNA, it barely yielded weak RNA cleavage near nucleotide 120.

Architecture of the bI3 maturase–RNA complex

To visualize the bI3 maturase–intron RNA complex, we developed an approximate three-dimensional model using five classes of constraints. (i) The P5-P4-P6, P5a and P5b structures closely resemble those in the homologous[2] *Tetrahymena* group I intron domain; thus, we modeled most of the bI3 RNA by modest structural substitution of the high-resolution structure for this domain[33]. This assumption is supported by the similarity in the solvent-based hydroxyl radical protection pattern of this region of the bI3 RNA to the protection pattern of the *Tetrahymena* intron domain (**Fig. 5d** and refs. 32,33). (ii) Starting with the basic architecture of the *Tetrahymena* domain, we then appended P5c, constraining it to be a helix as indicated by phylogenetic[2,6,7] analysis. Because the site-directed probing data (**Fig. 6b**) implicate helices P5b and P5c as the primary RNA interaction site for the bI3 maturase, the simple assumption of helicity places strong constraints on the protein-RNA interaction. (iii) The P5b and P5c helices are linked by a single-stranded bulge. We therefore allowed these two helices to rotate relative to one another to facilitate meeting other constraints. This rotation is still strongly constrained because helices P5b and P5c are linked at nucleotides 94 and 95. (iv) We used primarily the site-directed probing data (**Fig. 6b**) and secondarily the solvent-based probing data (**Fig. 5c**,**d**) to constrain manual docking of individual P5b and P5c helices onto the maturase. (v) Four well-ordered sulfate molecules lie near the N-terminal domain of the bI3 maturase saddle structure in our crystal structure. (Two were present in both A and B molecules and two were unique to each molecule.) We used these four positions as late-stage modeling constraints by assuming that phosphate groups from the P5b and P5c helices should be near these sulfates. Despite the large number of experimental constraints and the diverse experiments from which they originate, we could accommodate most constraints with a small class of similar structures, differing slightly in their relative rotational orientations along the minor groove of the P5b and P5c helices (**Figs. 7** and **8**).

The site-directed Fe(II)-BABE cleavage data place the N1 site at the minor groove of helix P5b and the N2-1 and N2-2 sites at the minor groove of helix P5c (**Fig. 7a**, top series). Site C2 lies at the minor groove of helix P5c one turn away from where N1 and N2-2 interact (**Fig. 7**, purple). Site C1 lies at the periphery of the P5c stem-loop and yields only a single site of weak cleavage (**Fig. 7a**, C1 panel).

The site-directed cleavage data thus imply that the maturase binds the bI3 RNA primarily via the N-terminal domain at the junction between the P5b and P5c helices in the minor groove. We tested this minor groove binding model independently using 2′-*O*-methyl interference experiments[35] (**Supplementary Fig. 1** online). The 2′-ribose position lies in the RNA minor groove, and converting the 2′-OH to the bulkier 2′-*O*-methyl should interfere with protein recognition

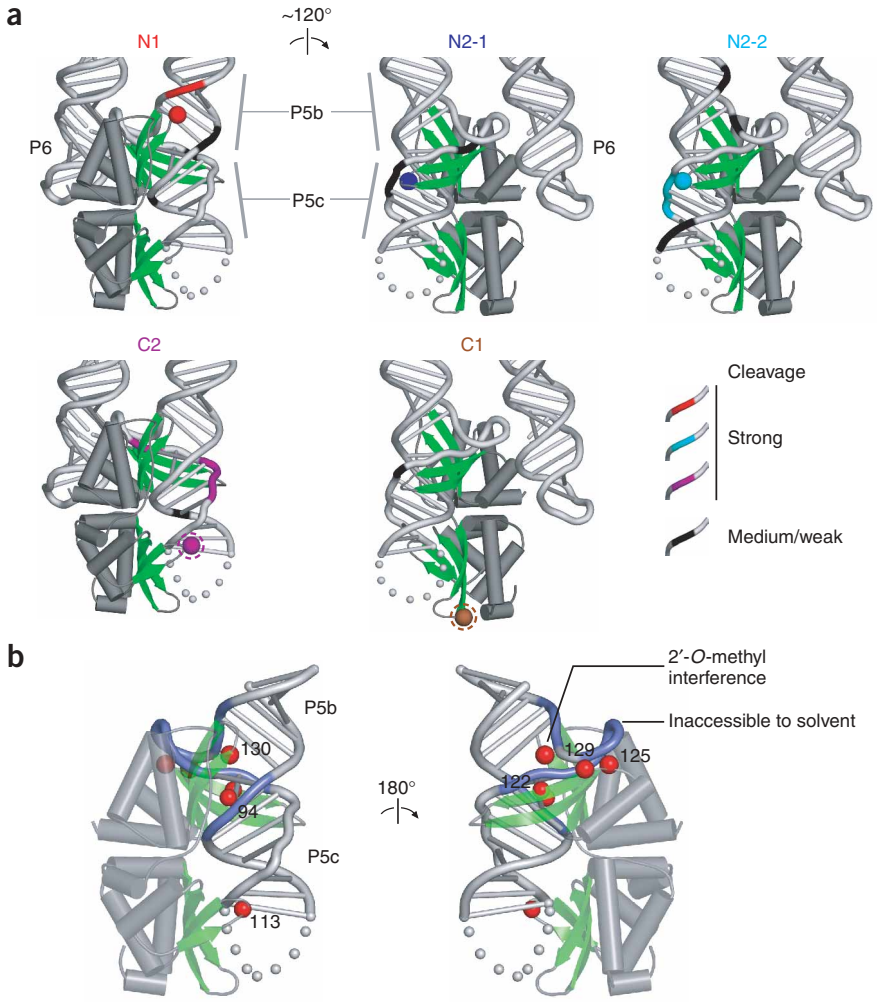


Figure 7 Maturase-RNA interactions. (**a**) Site-directed cleavage patterns superimposed on a three-dimensional model for the bI3 maturase and P5-P4-P6 domain complex. Spheres, sites of derivatization; colored RNA backbones, cleavage sites; green, β-strands of the bI3 maturase. (**b**) Summary of 2′-*O*-methyl interference in the P5b and P5c helices. Maturase is shown slightly transparent for clarity. Red spheres, sites of 2′-*O*-methyl interference (**Supplementary Fig. 1**); blue backbone, solvent-based hydroxyl radical cleavage sites (**Fig. 5c**,**d**).

mediated by minor groove interactions. We evaluated interference at adenosine and uridine nucleotides, which comprise 145 of 161 positions in this A/U-rich intron. When 2′-*O*-methyl phosphorothioate analogs were incorporated into the RNA and compared[35] with the interference attributable to the phosphorothioate substitution alone, interference was observed at 16 adenosine and uridine positions in the bI3 P5-P4-P6 domain RNA (**Supplementary Fig. 1**). Most of these lie in the conserved[2] P4, P5 and P5a regions and reflect close RNA contacts in the folded RNA. However, three positions in P5c and three in the loop linking P5b with P5c also interfere with maturase binding. Five of the interfering 2′-*O*-methyl substitutions form a compact constellation in or near the minor groove at the proposed interface with the N-terminal maturase domain (**Fig. 7b**, red spheres). The remaining interfering position (nucleotide 113) lies approximately under the C2 site in the maturase nucleic acid–binding saddle (**Fig. 7**). The interference data also support the interpretation that solution-phase hydroxyl radical footprinting protection in P5b and P5c primarily reflects direct interactions with the bI3 maturase (**Fig. 7b**).

DISCUSSION

LAGLIDADG proteins function as homing endonucleases by binding and cleaving DNA via intimate interactions across consecutive DNA major grooves. The bI3 maturase has lost the ability to cleave DNA but retains an architecture that is globally indistinguishable from its DNA-cleaving homologs (**Fig. 3**). Instead of global changes, a single substitution in the bI3 maturase amino acid sequence has compromised the endonuclease active site (**Fig. 4**). Thus, evolution of the bI3 maturase from a DNA endonuclease to an RNA cofactor involved macroscopic conservation of structure with focused local amino acid changes.

Because the global structure of the maturase changed so little relative to its DNA-binding homologs (**Fig. 3**), our work indicates that LAGLIDADG-motif proteins, with no further modification, are good candidate binding partners for group I introns and other RNAs. In this view, it is primarily the RNA that evolves a binding site for the maturase binding saddle, although this model does not preclude coevolution by the protein as well.

The narrow major and wide minor grooves in RNA are very different from the wide major and narrow minor grooves in DNA. However, a DNA-binding protein could certainly interact at the RNA major groove if bulges or other helical irregularities locally widened the major groove[36], as observed for many RNA-binding proteins[37]. For example, NF-κB binds tightly to an RNA aptamer at a distorted A-form major groove using the same side chains that normally recognize DNA[38]. Alternately, recognition of RNA by a DNA-binding protein could involve shifting to a binding target at the phosphate backbone, as occurs for the fifth zinc finger of transcription factor IIIA[39,40]: this domain interacts at the major groove of its DNA target, but interacts at the phosphate backbone to bridge the major groove in a helical region of 5S ribosomal RNA.

The bI3 maturase seems to have adopted a different strategy. Site-directed cleavage and 2′-*O*-methyl interference experiments implicate the RNA minor groove as the primary interaction site for the saddle-shaped nucleic acid–binding surface of the maturase (**Fig. 7**). This switch from major to minor groove recognition presumably takes advantage of the fact that it is the minor groove that is broad and most accessible in RNA.

The bI3 maturase is closely related to I-AniI (**Fig. 3**) and the bI3 maturase can, by itself, facilitate splicing by the AnCOB group I intron target of I-AniI (**Supplementary Fig. 2** online). Both proteins have a positively charged area on the surface of the maturase opposite the one that includes the nonconserved linker[15] (**Fig. 2a**,**b**). RNA interactions with this secondary site enhance binding by the I-AniI maturase by a small but notable increment (~1 kcal mol^{-1})[15]. In our model, this surface of the maturase lies adjacent to the P6 helix (**Fig. 8**, brown

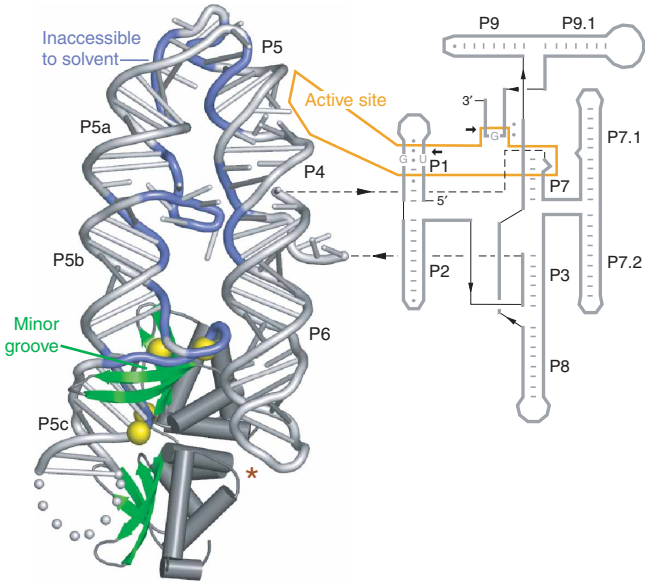


Figure 8 Maturase-facilitated folding of the bI3 intron RNA via action at a distance. The bI3 maturase recognizes a distal structure in a peripheral domain and lies at least 50 Å from the group I intron active site (orange). Green, the maturase β-strands; dark gray, the remainder of the protein; light blue, regions protected from solvent-based hydroxyl radical cleavage upon maturase binding; yellow spheres, sulfate groups visualized crystallographically; brown asterisk, a potential auxiliary RNA-protein interaction site.

asterisk), and the bI3 maturase may use this surface along with the saddle-shaped minor groove–binding surface to clamp the two ends of the U-shaped P5-P4-P6 domain together.

Despite these structural similarities, RNA-binding modes by the bI3 and I-AniI homologs seem to be distinct in that the two maturases use different domains to interact with their cognate RNAs. The N-terminal domain of the bI3 maturase forms more intimate contacts with the RNA than does the C-terminal domain (**Figs. 7** and **8**). In contrast, as a fusion protein with NusA, the C-terminal domain of I-AniI is sufficient to facilitate intron splicing[41]. It remains to be assessed whether the bI3 maturase binds the AnCOB intron in the same mode that it binds the bI3 intron or instead uses its C-terminal domain, as proposed for I-AniI.

Large RNAs rarely function independently in biological systems, but instead almost always require obligate protein cofactors that facilitate, in part, stable folding into a functional structure. In several well-studied examples[5], protein cofactors that facilitate splicing by group I intron RNAs bind near the splicing active site, within roughly one helical diameter, even if they do not directly facilitate transformations at the active site. The bI3 maturase clearly functions differently, binding instead at a distal, nonconserved extension of a peripheral intron domain. The maturase lies roughly 50 Å from the group I intron splicing active site at its closest approach (**Fig. 8**). Given the relative rigidity of the A-form helical building blocks of RNA, protein cofactors may be especially well suited to regulate RNA function by acting at long distances via structural changes in the intervening RNA.

METHODS

Protein expression and purification for crystallography. A cDNA encoding the soluble domain of the bI3 maturase (residues His256–Tyr517) with five additional N-terminal histidine residues forming a $(His)_6$ tag was amplified by PCR from an original bI3 maturase sequence generated by total gene synthesis[7] and cloned into pET19b. The maturase was expressed in *Escherichia coli* strain BL21(DE3) after induction with 0.4 mM isopropyl-β-D-thiogalactoside at 22 °C. Cells were harvested after 6 h and lysed by sonication in 100 mM HEPES (pH 7.6), 0.9 M NaCl, 2 mM 2-mercaptoethanol, 1 mM PMSF and 10% (v/v) glycerol. The maturase was purified by affinity chromatography on a Ni^{2+}-NTA column (Qiagen) and eluted with sonication buffer containing 100 mM imidazole. The protein was dialyzed into 50 mM HEPES (pH 7.6), 0.5 M NaCl, 2 mM DTT and 10% (v/v) glycerol, applied to a heparin affinity column (HiTrap, Amersham-Pharmacia Biotech) and eluted over a linear gradient (0.5–1.2 M NaCl). Fractions containing the maturase were pooled, dialyzed (10 mM HEPES (pH 7.6), 100 mM NaCl and 2 mM DTT) and concentrated (Centricon YM-10 concentrator, Millipore) to ~5 mg ml^{-1}. Selenomethionine (SeMet)-substituted protein was purified from *E. coli* strain B834(DE3) using the same protocol but with 5 mM DTT in the buffers.

Crystallization, data collection and phase determination. Diffraction-quality native or SeMet-substituted protein crystals (~0.4 × 0.3 × 0.05 mm) were obtained at 22 °C using sitting drop vapor diffusion by mixing 4 μl protein with 1 μl 0.1–0.3 mM *N*-tetradecyl- or *N*-tridecyl-β-D-maltoside and 4 μl reservoir solution containing 23% (w/v) PEG-MME 5,000, 180 mM ammonium sulfate and 100 mM MES (pH 5.5) and equilibrating over the reservoir solution. SeMet-Hg derivatives were obtained by soaking SeMet protein crystals in reservoir solution with either 10 mM methyl mercury (II) chloride (CH_3HgCl) overnight or 5 mM mercuric potassium iodide (K_2HgI_4) for 3 d.

For data collection, crystals were moved stepwise to a final cryoprotectant solution containing 30% (w/v) PEG-MME 5,000, 180 mM ammonium sulfate, 100 mM MES (pH 5.5) and 10% (v/v) ethylene glycol, and flash frozen in a stream of nitrogen gas at 100 K. MAD data were collected at three wavelengths (**Table 1**) from a single SeMet protein crystal at the Brookhaven National Laboratory (beamline X9B equipped with a Quantum 4 CCD detector). SeMet-Hg derivative datasets were collected using an RU-H3R rotating anode generator and a Raxis IV detector. Data were integrated and scaled using HKL[42]. Crystals are monoclinic and have two molecules in the asymmetric unit.

Of 24 expected selenium sites, 18 were identified from a weighted anomalous and dispersive Patterson map using CNS[43]. Initial phases and electron density maps were calculated using the SeMet MAD data. Map quality improved with the inclusion of two SeMet-Hg derivative datasets. Difference Fourier maps were used to identify two mercury sites for each of the derivatives. New experimental phases were calculated with MLPHARE[44] using a combination of SeMet MAD and multiple isomorphous replacement (MIR) with anomalous scattering using the SeMet-Hg derivatives. The SeMet MAD and SeMet-Hg derivative data were treated as a special case of MIR[45] with the SeMet MAD remote wavelength dataset used as the native dataset and the SeMet MAD datasets and two SeMet-Hg derivative datasets treated as derivatives. Histogram matching and solvent flattening to improve the electron density maps were carried out using DM[44].

Model building and refinement. Model building was performed with O[46] and XtalView/Xfit[47]. Iterative models were refined against the 2.2-Å SeMet dataset using CNS[43] with individual *B*-factor refinement at later steps. The final model contains almost the entire sequence of molecule A (residues Lys257–Tyr517), residues His256–Lys419 and Met425–Tyr517 of molecule B, eight sulfate molecules and 248 water molecules. Density for the 457–461 loop region in both protein molecules is weak but visible in $2|F_o|-|F_c|$ maps contoured at 0.6–0.8 σ. Inclusion of these high–*B* factor residues results in a lower R_{free} factor. A Ramachandran analysis (MolProbity server)[48] shows that 96.9% of backbone torsion angles are in the favored regions and 100% are in allowed regions; all residues except 262, 399, 456, 457 and 480 in molecule A and 261, 452 and 490 in molecule B refine to favored rotamers[48]. Proteins were superimposed and r.m.s. deviations were calculated with LSQMAN[49]; molecular structure figures were composed with PyMOL (http://www.pymol.org).

bI3 RNA construct, mutant protein expression and reaction conditions. The P5-P4-P6 domain RNA was produced by *in vitro* transcription from a PCR-generated template and purified by denaturing gel electrophoresis. The RNA (**Fig. 5d**) contains two 5′-nontemplated guanosine nucleotides to facilitate

Table 1 Crystallographic data and structure refinement

	SeMet	SeMet MAD			SeMet/MeHgCl	SeMet/K_2HgI_4
Data collection						
Space group	$P2_1$	$P2_1$			$P2_1$	$P2_1$
Cell dimensions						
a, *b*, *c* (Å)	63.1, 80.1, 66.5	63.3, 80.1, 66.6			63.3, 80.2, 66.1	63.2, 80.2, 66.5
α, β, γ (°)	90, 115.6, 90	90, 115.6, 90			90, 115.6, 90	90, 115.7, 90
		Peak	*Inflection*	*Remote*		
Wavelength (Å)	1.54	0.9791	0.9795	0.9714	1.54	1.54
Resolution (Å)	2.2	2.6	2.6	2.6	2.8	2.8
R_{sym}[a]	8.2 (30.3)	10.6 (53.1)	12.4 (48.7)	10.1 (42.6)	10.3 (33.2)	16.3 (42.3)
$I / \sigma I$[a]	11.0 (2.4)	15.1 (3.0)	11.3 (2.1)	14.2 (2.9)	9.3 (2.7)	7.6 (2.3)
Completeness (%)[a]	99.1 (94.8)	98.9 (96.3)	98.2 (87.6)	99.2 (97.7)	98.9 (94.8)	98.8 (93.7)
Redundancy	3.1	3.6	3.7	3.7	3.4	3.2
Refinement						
Resolution (Å)	2.2					
No. reflections	30,022 (2,858)					
R_{work} / R_{free}	20.8 / 26.5					
No. atoms						
Protein	4,294					
Ligand / ion	40					
Water	248					
B-factors						
Protein	38.2					
Ligand / ion	51.1					
Water	38.6					
R.m.s. deviations						
Bond lengths (Å)	0.007					
Bond angles (°)	1.32					

[a]Highest resolution shell is shown in parenthesis.

transcription. RNA footprinting and binding reactions were carried out in 40 mM MOPS (pH 7.7), 80 mM potassium acetate (pH 7.6) and 20 mM $MgCl_2$. RNA was refolded by denaturation in 10 mM Tris (pH 7.5) and 1 mM EDTA (1 min, 90 °C), incubation on ice (1 min), addition of reaction buffer (10 min, 37 °C) and slow cooling (30 min, 25 °C) before addition of maturase protein. The ΔCys mutant and the single-cysteine maturase variants (G284C, M322C, G325C, M425C and F459C) were created by oligonucleotide-directed mutagenesis (Stratagene QuikChange). Proteins were purified in parallel via a $(His)_6$ affinity tag, as described in ref. 7, except that 3 mM tris-(2-carboxyethyl)-phosphine was substituted for 2 mM DTT in the final dialysis steps. Equilibrium dissociation constants for the RNA– (**Fig. 5a**) or DNA (see ref. 7)–protein complexes were measured using a dual filter system[20] in reaction buffer supplemented with 100 μg ml^{-1} BSA; reactions contained 10 pM ^{32}P-radiolabeled RNA or DNA and 30 pM–1 μM maturase.

Solvent-based hydroxyl radical footprinting. Refolded end-labeled P5-P4-P6 RNA and maturase were incubated (10 μl, 10 min, 25 °C) at final concentrations of 64 and 320 nM, respectively. Hydroxyl radical cleavage (15 min, 25 °C) was initiated by the addition of one-tenth volume each of (i) 30 mM $(NH_4)_2Fe(SO_4)_2$ and 45 mM EDTA, (ii) 15 mM DTT and (iii) 50 mM sodium ascorbate. Reactions were quenched by the addition of 1 μl 2 M thiourea, 1 μl 0.5 M EDTA and 1.3 μl proteinase K (1 mg ml^{-1}; 30 min at 25 °C) and then diluted with formamide loading buffer before resolution by gel electrophoresis. Individual band intensities were integrated using SAFA[31] and protection was calculated as $I - I_{maturase}$, where I is the band intensity minus background (from a reaction in which the quench solution was added prior to cleavage reagents) in the free RNA and $I_{maturase}$ is the net intensity upon maturase binding.

Site-directed hydroxyl radical cleavage. Site-directed mutagenesis was used to convert the two cysteine residues at positions 219 and 302 in the native sequence to valine and serine, respectively; this parent construct was used to introduce unique cysteine residues at solvent-accessible positions. Proteins (25 μl) were treated with Fe(II)-BABE (5 μl, 5 mM; bromoacetamidobenzyl-EDTA-Fe(II), Pierce) at 37 °C for 45 min. Excess Fe(II)-BABE was removed by dialysis (10 h, 4 °C, twice) against 5,000 volumes of 40 mM HEPES (pH 7.7) and 600 mM NaCl. Final protein concentrations were determined by fluorimetry (NanoOrange dye, Molecular Probes). The RNA-binding activity for each mutant and for the Fe(II)-BABE adducts was determined using equilibrium filter binding experiments. All derivatized variants have equilibrium binding affinities within two-fold that of the parent ΔCys protein (K_d = 2.5 nM), except the derivatized N1 variant, which binds eight times more weakly than the native protein. All proteins have binding profiles consistent with formation of equimolar RNA–protein complexes (**Fig. 5a** and data not shown) and yield ≥90% bound RNA under the conditions of our cleavage experiments. For site-directed cleavage experiments, refolded 5′ ^{32}P end-labeled P5-P4-P6 RNA (24 nM) was mixed with 2 μl maturase or protein dilution buffer to give (in 10 μl) 20 nM RNA, and 125 or 250 nM maturase in reaction buffer supplemented with 8 mM HEPES (pH 7.6) and 120 mM NaCl (from the protein dilution buffer). After an initial incubation (15 min, 25 °C), hydrogen peroxide and ascorbic acid were added (2 μl) to final concentrations of 0.05% (v/v) and 5 mM, respectively. The cleavage reaction (3 min, 25 °C) was quenched by the addition of 12 μl 85% (v/v) formamide, 1 M thiourea, 60 mM EDTA, 2% (w/v) SDS and dyes. Samples were resolved on a series of 8–25% (w/v) denaturing polyacrylamide gels to achieve nucleotide resolution at almost every RNA position. Absolute cleavage intensities for each maturase mutant were calculated by band integration[31] and then subtraction of the background observed for the mock-labeled ΔCys protein. Absolute cleavage intensities three-fold and greater than five-fold above background were judged to be medium-to-weak and strong, respectively.

Modeling the bI3 maturase–RNA complex. The crystallographic structure of the bI3 maturase (molecule A) was used as a rigid unit. RNA building and docking steps were performed with Sybyl (Tripos) and PyMOL (http://www.pymol.org).

Accession codes. Protein Data Bank: Coordinates have been deposited with the accession code 2AB5. BIND identifier (http://bind.ca): 315515.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

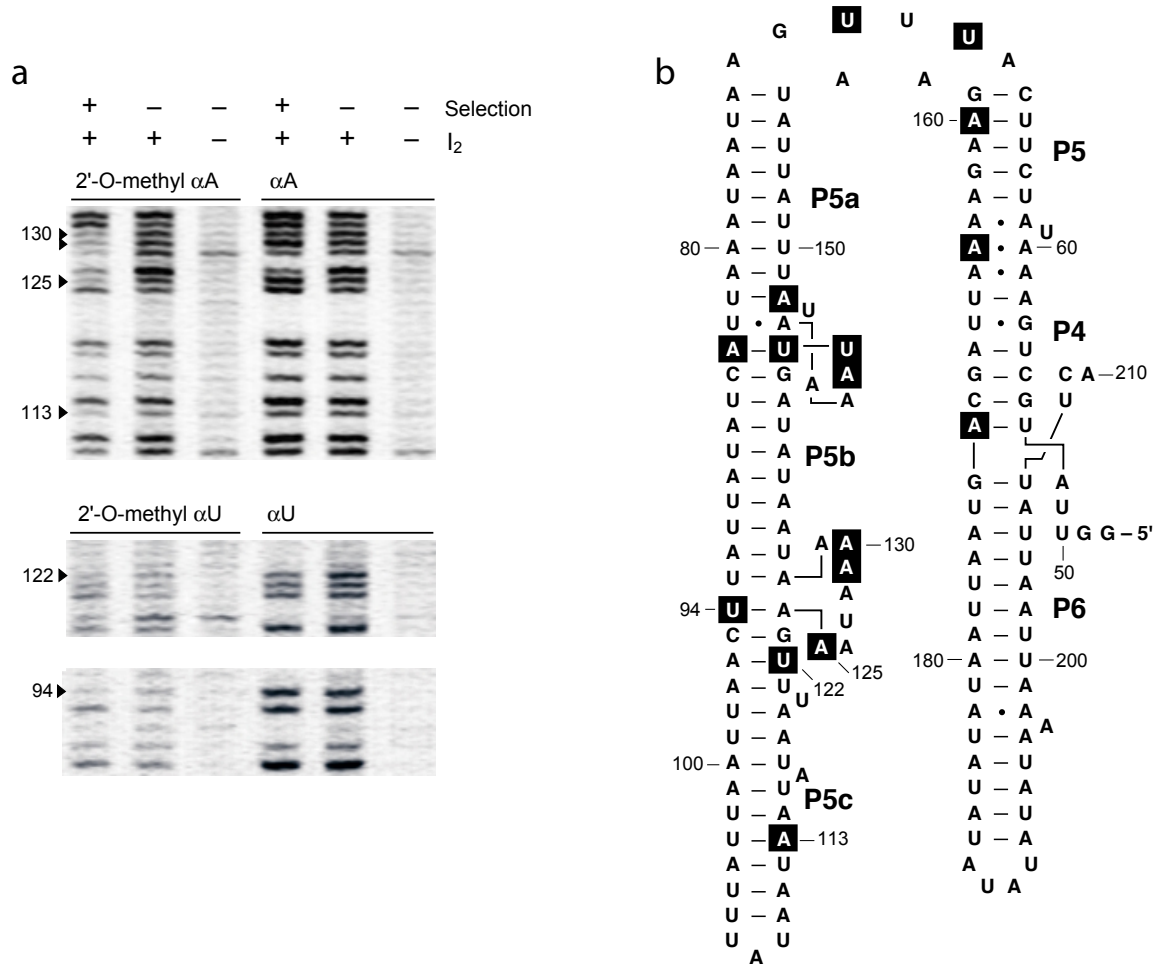
We are indebted to L. Pedersen and Z. Dauter for assistance with synchrotron data collection at beamline X9B at Brookhaven National Laboratory, D. de Oliveira for performing DNA binding assays and our colleagues for critical comments on the manuscript. This work was supported by the US National Institutes of Health (NIH) grant GM56222 to K.M.W. and by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (T.M.T.H.).

COMPETING INTERESTS STATEMENT

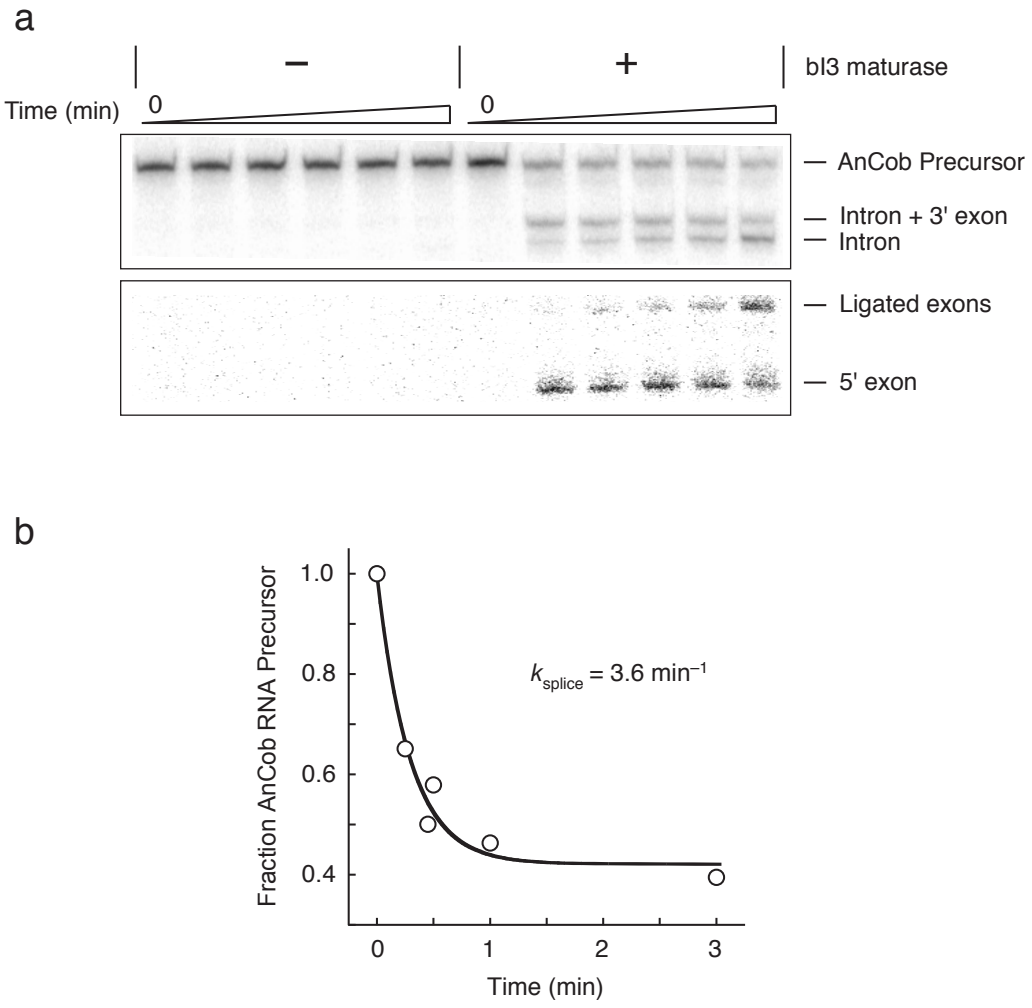
The authors declare that they have no competing financial interests.

Received 17 June; accepted 19 July 2005
Published online at http://www.nature.com/nsmb/

1. Cech, T.R. Self splicing of group I introns. *Annu. Rev. Biochem.* **59**, 543–568 (1990).
2. Michel, F. & Westhof, E. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* **216**, 585–610 (1990).
3. Adams, P.L., Stahley, M.R., Kosek, A.B., Wang, J. & Strobel, S.A. Crystal structure of a self-splicing group I intron with both exons. *Nature* **430**, 45–50 (2004).
4. Golden, B.L., Kim, H. & Chase, E. Crystal structure of a phage Twort group I ribozyme-product complex. *Nat. Struct. Mol. Biol.* **12**, 82–89 (2004).
5. Weeks, K.M. Protein-facilitated RNA folding. *Curr. Opin. Struct. Biol.* **7**, 336–342 (1997).
6. Hur, M., Geese, W.J. & Waring, R.B. Self-splicing activity of the mitochondrial group-I introns from *Aspergillus nidulans* and related introns from other species. *Curr. Genet.* **32**, 399–407 (1997).
7. Bassi, G.S., de Oliveira, D.M., White, M.F. & Weeks, K.M. Recruitment of intron-encoded and co-opted proteins in splicing of the bI3 group I intron RNA. *Proc. Natl. Acad. Sci. USA* **99**, 128–133 (2002).
8. Lambowitz, A.M., Caprara, M.G., Zimmerly, S. & Perlman, P.S. Group I and Group II ribozymes as RNPs: clues to the past and guides to the future. in *The RNA World* (eds. Gesteland, R.F., Cech, T.R. & Atkins, J.F.) 451–485 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1999).
9. Rho, S.B. & Martinis, S.A. The bI4 group I intron binds directly to both its protein splicing partners, a tRNA synthetase and maturase, to facilitate RNA splicing activity. *RNA* **6**, 1882–1894 (2000).
10. Lazowska, J., Jacq, C. & Slonimski, P.P. Sequence of introns and flanking exons in wild-type and *box3* mutants of cytochrome *b* reveals an interlaced splicing protein coded by an intron. *Cell* **22**, 333–348 (1980).
11. Szczepanek, T. & Lazowska, J. Replacement of two non-adjacent amino acids in the *S. cerevisiae* bi2 intron-encoded RNA maturase is sufficient to gain a homing-endonuclease activity. *EMBO J.* **15**, 3758–3767 (1996).
12. Ho, Y., Kim, S.J. & Waring, R.B. A protein encoded by a group I intron in *Aspergillus nidulans* directly assists RNA splicing and is a DNA endonuclease. *Proc. Natl. Acad. Sci. USA* **94**, 8994–8999 (1997).
13. Gampel, A. & Tzagoloff, A. *In vitro* splicing of the terminal intervening sequence of *Saccharomyces cerevisiae* cytochrome b pre-mRNA. *Mol. Cell. Biol.* **7**, 2545–2551 (1987).
14. Weeks, K.M. & Cech, T.R. Efficient protein-facilitated splicing of the yeast mitochondrial bI5 intron. *Biochemistry* **34**, 7728–7738 (1995).
15. Bolduc, J.M. *et al.* Structural and biochemical analyses of DNA and RNA binding by a bifunctional homing endonuclease and group I intron splicing factor. *Genes Dev.* **17**, 2875–2888 (2003).
16. Paukstelis, P.J. *et al.* A tyrosyl-tRNA synthetase adapted to function in group I intron splicing by acquiring a new RNA binding surface. *Mol. Cell* **17**, 417–428 (2005).
17. Dalgaard, J. *et al.* Statistical modeling and analysis of the LAGLIDADG family of site-specific endonucleases and identification of an intein that encodes a site-specific endonuclease of the HNH family. *Nucleic Acids Res.* **25**, 4626–4638 (1997).
18. Chevalier, B.S. & Stoddard, B.L. Homing endonucleases: structural and functional insight into the catalysis of intron/intein mobility. *Nucleic Acids Res.* **29**, 3757–3774 (2001).
19. Belfort, M. Two for the price of one: a bifunctional intron-encoded DNA endonuclease-RNA maturase. *Genes Dev.* **17**, 2860–2863 (2003).
20. Bassi, G.S. & Weeks, K.M. Kinetic and thermodynamic framework for assembly of the six-component bI3 group I intron ribonucleoprotein catalyst. *Biochemistry* **42**, 9980–9988 (2003).
21. Lazowska, J. *et al.* Protein encoded by the third intron of cytochrome b gene in *Saccharomyces cerevisiae* is an mRNA maturase. Analysis of mitochondrial mutants, RNA transcripts proteins and evolutionary relationships. *J. Mol. Biol.* **205**, 275–289 (1989).
22. Galburt, E.A. & Stoddard, B.L. Catalytic mechanisms of restriction and homing endonucleases. *Biochemistry* **41**, 13851–13860 (2002).
23. Silva, G.H., Dalgaard, J.Z., Belfort, M. & Van Roey, P. Crystal structure of the thermostable archaeal intron-encoded endonuclease I-Dmol. *J. Mol. Biol.* **286**, 1123–1136 (1999).
24. Heath, P.J., Stephens, K.M., Monnat, R.J.J. & Stoddard, B.L. The structure of I–CreI, a group I intron-encoded homing endonuclease. *Nat. Struct. Biol.* **4**, 468–476 (1997).
25. Chevalier, B.S., Monnat, R.J.J. & Stoddard, B.L. The homing endonuclease I–CreI used three metals, one of which is shared between the two active sites. *Nat. Struct. Biol.* **8**, 312–316 (2001).
26. Chatterjee, P., Brady, K.L., Solem, A., Ho, Y. & Caprara, M.G. Functionally distinct nucleic acid binding sites for a group I intron encoded RNA maturase/DNA homing endonuclease. *J. Mol. Biol.* **329**, 239–251 (2003).
27. Goddard, M.R. & Burt, A. Recurrent invasion and extinction of a selfish gene. *Proc. Natl. Acad. Sci. USA* **96**, 13880–13885 (1999).
28. Chevalier, B. *et al.* Metal-dependent DNA cleavage mechanism of the I–CreI LAGLIDADG homing endonuclease. *Biochemistry* **43**, 14015–14026 (2004).
29. Seligman, L.M. *et al.* Mutations altering the cleavage specificity of a homing endonuclease. *Nucleic Acids Res.* **30**, 3870–3879 (2002).
30. Latham, J.A. & Cech, T.R. Defining the inside and outside of a catalytic RNA molecule. *Science* **245**, 276–282 (1989).
31. Das, R., Laederach, A., Pearlman, S.M., Herschlag, D. & Altman, R.B. SAFA: Semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. *RNA* **11**, 344–354 (2005).
32. Murphy, F.L. & Cech, T.R. GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J. Mol. Biol.* **236**, 49–63 (1994).
33. Cate, J.H. *et al.* Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* **273**, 1678–1685 (1996).
34. Culver, G.M. & Noller, H.F. Directed hydroxyl radical probing of RNA from iron(II) tethered to proteins in ribonucleoprotein complexes. *Methods Enzymol.* **318**, 461–475 (2000).
35. Ryder, S.P., Ortoleva-Donnelly, L., Kosek, A.B. & Strobel, S.A. Chemical probing of RNA by nucleotide analog interference mapping. *Methods Enzymol.* **317**, 92–109 (2000).
36. Weeks, K.M. & Crothers, D.M. Major groove accessibility of RNA. *Science* **261**, 1574–1577 (1993).
37. Draper, D.E. Themes in RNA-protein recognition. *J. Mol. Biol.* **293**, 255–270 (1999).
38. Huang, D.B. *et al.* Crystal structure of NF-κB $(p50)_2$ complexed to a high-affinity RNA aptamer. *Proc. Natl. Acad. Sci. USA* **100**, 9268–9273 (2003).
39. Nolte, R.T., Conlin, R.M., Harrison, S.C. & Brown, R.S. Differing roles for zinc fingers in DNA recognition: structure of a six-finger transcription factor IIIA complex. *Proc. Natl. Acad. Sci. USA* **95**, 2938–2943 (1998).
40. Lu, D., Searles, M.A. & Klug, A. Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. *Nature* **426**, 96–100 (2003).
41. Downing, M.E., Brady, K.L. & Caprara, M.G. A C-terminal fragment of an intron-encoded maturase is sufficient for promoting group I intron splicing. *RNA* **11**, 437–446 (2005).
42. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
43. Brunger, A.T. *et al.* Crystallography and NMR system (CNS): A new software system for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998).
44. Collaborative Computational Project. Number 4 The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
45. Ramakrishnan, V. & Biou, V. Treatment of multiwavelength anomalous diffraction data as a special case of multiple isomorphous replacement. *Methods Enzymol.* **276**, 538–557 (1997).
46. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and location of errors in these models. *Acta Crystallogr. D Biol. Crystallogr.* **47**, 110–119 (1991).
47. McRee, D.E. XtalView/Xfit - A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165 (1999).
48. Lovell, S.C. *et al.* Structure validation by C-alpha geometry: phi, psi, and C-beta deviation. *Proteins* **50**, 437–450 (2003).
49. Kleywegt, G.J. & Jones, T.A. Detecting folding motifs and similarities in protein structures. *Methods Enzymol.* **277**, 525–545 (1997).



Suppl. Figure 1. 2'-O-methyl interference analysis of the bI3 maturase-RNA interaction. (**a**) Representative interference experiments at adenosine and uridine residues. Phosphorothioate (αA and αU) nucleotides were identified by iodine (I_2) cleavage. Phosphorothioate substituted RNAs were synthesized using a Y639F mutant of T7 RNA polymerase [Padilla et al., Nucl. Acids Res. 27, 1561-1653 (1999)]. RNA (50 nM) –maturase (10 nM) complexes were selected by filter partitioning [Garcia & Weeks, J. Mol. Biol. 331, 57-73 (2003)]. Interference values [Ryder et al., Methods Enz. 317, 92-109 (2000)] greater than two were taken to indicate that the 2'-O-methyl substitution impaired maturase binding. Interfering nucleotides are boxed in (**b**).



Suppl. Figure 2. bI3 maturase-facilitated splicing of the AnCOB group I intron RNA in the cytochrome *b* gene of Aspergillus nidulans. (**a**) Splicing reactions employed an AnCOB-□L8 precursor [Hur et al. Curr. Genet. 32, 399-407 (1997); generously provided by Richard Waring] and the bI3 maturase as the sole protein cofactor. Reactions were performed as described by Bassi et al. [Proc. Natl. Acad. Sci USA 99, 128-133 (2002)] at 20 mM $MgCl_2$ and 3 mM guanosine monophosphate. Reactions were monitored using 1 nM internally [^{32}P]-labelled RNA precursor and 10 nM bI3 maturase, if present. For clarity, figure includes two excerpted sections from a longer gel; all bands apparent on gel are shown. (**b**) Single turnover kinetics for splicing of the AnCOB intron facilitated by the bI3 maturase. Fraction RNA precursor was fit to $f\exp(-k_{splice}t + (1-f)$, where k_{splice} is the rate constant for splicing and f is the fraction reactive RNA.