Small molecules that target group II introns are potent antifungal agents

Olga Fedorova1,2,4, G. Erik Jagdmann Jr.3,4, Rebecca L. Adams2, Lin Yuan3, Michael C. Van Zandt3 and Anna Marie Pyle1,2*

Specific RNA structures control numerous metabolic processes that impact human health, and yet efforts to target RNA structures de novo have been limited. In eukaryotes, the self-splicing group II intron is a mitochondrial RNA tertiary structure that is absent in vertebrates but essential for respiration in plants, fungi and yeast. Here we show that this RNA can be targeted through a process of high-throughput in vitro screening, SAR and lead optimization, resulting in high-affinity compounds that specifically inhibit group II intron splicing in vitro and in vivo and lack toxicity in human cells. The compounds are potent growth inhibitors of the pathogen Candida parapsilosis, displaying antifungal activity comparable to that of amphotericin B. These studies demonstrate that RNA tertiary structures can be successfully targeted de novo, resulting in pharmacologically valuable compounds.

It is becoming increasingly clear that large, highly structured RNA molecules are essential for most metabolic functions. In theory, RNA molecules present novel targets for drug discovery, but, with the exception of bacterial riboswitch inhibitors1, small-molecule inhibitors of RNA tertiary structures have not been successfully identified using the high-throughput screening and classical medicinal chemistry campaigns that have yielded most protein inhibitors. There are, to our knowledge, no new classes of antimicrobial compounds that target RNA and none that are designed to target eukaryotic pathogens. Though most inhibitors operate at the protein level, small molecules that target specific RNA structures can potentially modulate gene expression through a diversity of mechanisms, representing a powerful orthogonal strategy for the treatment of disease.

Previous studies on bacterial riboswitches have demonstrated that folded RNA molecules often bind small molecules such as metabolites with high affinity, and these ligands have been optimized in efforts to develop new antibiotics1-7. Likewise, group I introns and HIV-1 TAR RNA are known to bind small molecules8-11. Indeed, antibacterial compounds have long been known to target ribosomal RNA12,13, and given the availability of high-resolution structural data, these have been the subject of continuing optimization14,15. There has also been a growing interest in understanding structural and physicochemical properties of small molecules that selectively bind RNA motifs16,17. Though these studies provide important precedents, they were conducted on RNA molecules with known small-molecule ligands. De novo RNA targeting efforts by HTS and other methods have been limited, and they have focused primarily on small RNA secondary structural elements, such as junctions, miRNAs, RNA hairpins from triplet repeat diseases and stem-loops in viral RNA genomes18,19,20. It would be ideal to identify small molecules that selectively bind RNA tertiary structures, as these more complex RNA motifs provide a route to specific molecular recognition, and they are often found in biomolecules of high medical relevance.

To accomplish this goal, we have focused on RNA targets within pathogenic yeasts, as these organisms have become increasingly problematic, particularly for patients with compromised immune systems such as recipients of implanted devices, neonatal patients, and cancer patients21. For example, there has been a marked increase in pathologies associated with non-albicans strains, particularly C. parapsilosis22. The availability of potent antifungals that lack toxicity in mammals is a major unmet medical need and is of value for industrial and agricultural applications. The development of new antifungals is difficult as eukaryotes, fungi and yeast cells have enzymes and biochemical pathways that are similar to those of humans. However, fungal RNA metabolism differs substantially, thereby providing a potential route toward new therapeutics23.

To meet these challenges, we set out to identify small-molecule inhibitors of group II introns, which are large self-splicing ribozymes that are found in the mitochondrial genomes of plants, fungi, and yeast, but are not present in mammals. These autocatalytic RNA molecules adopt an elaborate tertiary structure that has been crystallographically characterized and contains an active site for RNA cleavage and ligation, as well as solvent-accessible pockets for potential inhibitor binding24,25. In yeasts such as Saccharomyces cerevisiae and the pathogen C. parapsilosis, group II introns are found within genes that are essential for respiration, such as cytochrome oxidase subunit genes of the mitochondria26,27. Importantly, respiration is essential for pathogenic yeast to differentiate into biofilms, which colonize medical implant surfaces, are relatively resistant to antifungals, and contribute to pathogenic virulence28,29,30. Thus, based on the complexity of their structures, and their essential role in fungal metabolism, group II introns represent outstanding targets for the development of highly specific antifungal agents.

**Results**

**High-throughput screening.** To identify group II intron splicing inhibitors, we developed a sensitive, high-throughput fluorescence assay for monitoring ribozyme activity of the well-characterized ai5y group II intron from S. cerevisiae (Supplementary Figs. 1a and 2; Supplementary Table 1; Fig. 1a). The self-splicing ai5y group II intron was transformed into a multiple-turnover ribozyme by

---

1Howard Hughes Medical Institute, New Haven, CT, USA. 2Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT, USA. 3New England Discovery Partners, Branford, CT, USA. 4These authors contributed equally: Olga Fedorova, G. Erik Jagdmann Jr.

*e-mail: anna.pyle@yale.edu

ARTICLES

https://doi.org/10.1038/s41589-018-0142-0
removing the flanking exons and consolidating catalytic domains to create the D135 ribozyme, which efficiently and specifically catalyzes cleavage of RNA 'substrate' oligonucleotides that contain sequences of the original 5′ splice site. The RNA substrate for the screening assay was an oligonucleotide containing the last 17 nucleotides of the 5′ exon (which encompasses the 'intron binding sequences' that base pair with intron domain 1) and the first two nucleotides of the intron (substrate 17/2 DL, Supplementary Fig. 2a). A fluorophore and a fluorescence quencher were conjugated at positions on opposite sides of the scissile linkage so that fluorescence is observed when the oligonucleotide substrate is cleaved by the ribozyme (Supplementary Fig. 2b,c).

Using this assay for ribozyme activity, we screened a curated library of 10,000 compounds (see Methods) and identified 16 reproducible hits, some of which shared common structural elements that suggested a shared mechanism of action (Supplementary Fig. 2d). We then analyzed a series of commercially available derivatives of the major hits to identify more suitable scaffolds for further optimization. The most potent scaffold identified during this phase of the study was compound 1, which exhibited an IC_{50} of 2 μM (Supplementary Table 2). It was used as a starting point for the design of additional compounds that were used to define structure–activity relationships (SAR) and to optimize potency (Supplementary Table 2).

**In vitro SAR and optimization of potency.** To carry out the SAR studies in vitro, we complemented the primary fluorimetric assay with a robust secondary radioanalytic self-splicing assay of the precursor RNA containing the full-length αβγ intron and short exons, which enabled the determination of K<sub>i</sub> values for all compounds of interest (Fig. 1b; Supplementary Fig. 3a–c; Supplementary Table 2).

As part of an initial strategy to optimize the early leads, a series of compounds was selected to determine the critical structural components required for activity, i.e., the pharmacophore. Three regions of compound 1 (A, B and C) were defined, and substituents in these sections of the molecule were evaluated for their effect on inhibition (Fig. 2). In region A, replacing any of the hydroxyl groups with hydrogen or methoxy or boronic acid substituents (compounds 2–5, Supplementary Table 2) resulted in nearly complete loss of activity. Notably, when the trihydroxyl was replaced with a dihydroxyl catechol moiety, the molecule was inactivated both in vitro and in vivo (2, Supplementary Table 2), indicating that the presence of the catechol motif by itself cannot explain reactivity of the group II intron inhibitors, unlike certain classes of promiscuous molecules that are collectively classified as 'PAINS' compounds.

In contrast, each of the hydroxy groups in region C could be removed or replaced with halogen atoms without significant loss of function (compounds 6–8, Fig. 2; Supplementary Table 2). To minimize any risk of reactivity from the α,β-unsaturated ketone, we replaced the 2-benzylidenbenzofuran-3(2H)-one moiety with a more chemically and metabolically stable benzofuran-2-y(phenyl) methanone (region B). This more drug-like molecular template resulted in a two-fold increase in in vitro potency compared to the parental molecule 7 (compound 9, Fig. 2; Supplementary Table 2). Other attempts to introduce changes in this region (for example, amide or thiazole derivatives) resulted in a substantial loss of activity (compounds 10 and 11, Supplementary Table 2). The fact that potency was sensitive to certain modifications in region B, and in distal parts of region C (Supplementary Table 2), indicates that inhibitory activity is not solely attributable to functional groups in region A.

With the new lead structure (compound 9) in hand, additional substituents were added to the benzofuran moiety to further develop the structure–activity relationship for the series. Introductions of a wide variety of substituents, including aryl, heteroaryl, amino or halogens at the 5′ or 6′ positions were all generally tolerated (compounds 8, 12, 13 and 14–19; Fig. 2; Supplementary Fig. 3c; Supplementary Table 2). Notably, our data suggest that adding large substituents or positively charged residues to region C can increase inhibitory activity (lower the K<sub>i</sub>) of the respective compounds (Fig. 2; Supplementary Table 2). Examples of such effects are observed for compounds 13, 14 and 19 (Figs. 1b and 2; Supplementary Fig. 3c; Supplementary Table 2).

Reversibility of inhibitor binding to the intron was established using pulse-chase dilution experiments (Supplementary Fig. 3d, see Methods). The clear reactivity patterns evident from this SAR...
In vitro splicing inhibition constants \( (K_i) \) for the ai5γ intron, MIC values for C. parapsilosis and IC\(_{50} \) values for cytotoxicity in HEK-293 cells are shown next to each molecule. Color coding is used to highlight functional regions of the lead compound. Changes in region A of compound 1 are shown in blue, changes in region B are shown in green, and changes in region C are shown in yellow. Changes that were introduced for optimization of inhibitory activity are highlighted in pink.

To directly monitor the effect of our most potent compounds on group II intron splicing in vivo, we developed a qRT-PCR assay for monitoring the splicing of the ai5γ intron in S. cerevisiae in the presence of small molecules. We found that our most potent compounds cause a severe splicing defect in vivo, which is evident from substantial accumulation of precursor RNA molecules containing the 5′-exon–intron junction (Fig. 3b; Supplementary Fig. 4b). Importantly, unspliced COX1 transcripts are targeted for rapid degradation in cells\(^{1,2} \), and thus the direct observation of substantial precursor accumulation suggests a considerable effect on splicing. Taken together, our results in cells demonstrate that the highest affinity compounds specifically target the ai5γ intron in vivo, selectively disrupting splicing of the COX1 gene and thereby reducing yeast growth.

**Selectivity of the group II intron inhibitors.** Although the SAR, reversibility analysis and gene specificity of the compounds are consistent with selectivity, it was important to determine whether the compounds can also bind other highly structured RNA molecules and other RNA splicing systems. In addition, we wondered whether activity requires the fully folded group II intron RNA tertiary structure or if individual intron domains can bind the inhibitors with high affinity. To this end, we monitored inhibitory activity of one of the most promising compounds (19) in the presence of a large excess of various RNAs, including separate ai5γ intron domains D1, D3, D5, the U2–U6 snRNA stem-loop (analogous to group II intron D5) and yeast tRNA\(^{\text{5hu}} \) (Supplementary Fig. 1, see Methods). The latter was an important control, because tRNA molecules possess many archetypal elements of RNA tertiary structure such as kissing loops involving canonical and noncanonical base pairs, coaxially stacked helices, base triples and U-turn motifs, which make them
commonly used specificity controls for RNA targeting. However, we observe that none of these RNAs, presented in a 1,000-fold excess relative to intron RNA (2 nM), affect the inhibitory activity of compound 19 (Fig. 4a). The only RNA that competed with radiolabeled SE group II intron RNA for binding of 19 was the same unlabeled group IIb intron RNA added in excess (Fig. 4a).

To evaluate inhibition of the other two known RNA splicing systems (group I and spliceosome), we monitored splicing of the Aszoraccs pre-rRNA (Ile) group I intron in the presence of compound 19. We observed that splicing of the Aszoraccs intron is unaffected even at 100 μM compound 19 (Fig. 4b). In addition, we used qRT-PCR to monitor inhibition of group I intron and spliceosomal splicing in S. cerevisiae in vivo. Consistent with our previous results, we found that the inhibitors only affect splicing of the yeast aI5 group II intron (Fig. 3b). Spliceosomal processing, group I intron splicing and even splicing of group II introns from subclasses that differ from subclass IIB are unaffected by the small-molecule inhibitors. These results are consistent with the in vitro results and suggest that the inhibitors bind selectively to group IIB introns. Our data also indicate that the inhibitors bind tertiary structural elements formed by the entire intron and not individual intronic domains.

Small-molecule growth inhibition of C. parapsilosis. The yeast pathogen C. parapsilosis contains a single group IIB intron in its COXI gene. The active site of this intron (DS5) is almost identical to that of the aI5 intron (Supplementary Fig. 1b), suggesting that compounds that inhibit the S. cerevisiae aI5 intron may also inhibit splicing by the group II intron in C. parapsilosis. To evaluate efficacy of the compounds against this pathogen, we measured the minimum inhibitory concentrations (MIC values) required for growth inhibition of C. parapsilosis. We observed that the high-affinity aI5 intron inhibitors significantly reduce the growth of C. parapsilosis (Fig. 2; Supplementary Table 2). Correlations between IC_{50} K_{i} and MIC values suggest that these compounds employ the same mechanism of action both in vitro and in vivo (Fig. 2; Supplementary Table 2). Although many of the highest affinity compounds displayed strong MIC values, compounds 18 (K_{i} = 2.1 ± 0.2 μM) and 19 (K_{i} = 0.36 ± 0.02 μM) were particularly notable because their MIC values (2–4 μg/ml) are comparable to that of amphotericin B, which is still commonly used for acute C. parapsilosis infection (MIC is 0.5–1 μg/ml; Fig. 2; Supplementary Table 2). Given their potency and antifungal effects, we have named compound 18 Intronistat A and compound 19 Intronistat B.

To directly monitor the behavior of the compounds in vivo, we used qRT-PCR to quantify levels of splicing for the C. parapsilosis COXI precursor mRNA (which contains the group IIB intron) in the presence of Intronistat B and inactive compound 4. We observed a moderate splicing defect caused by Intronistat B, as indicated by increased levels of unspliced C. parapsilosis COXI relative to total; this effect was not observed in the presence of inactive compound 4 (Supplementary Fig. 4c). To determine whether inhibition is specific to yeast, we evaluated toxicity of the most potent compounds in human cells, determining the IC_{50} for inhibition of HEK-293T cells (Fig. 2; Supplementary Fig. 5; Supplementary Table 2). Although some of the compounds are broadly toxic to all eukaryotic cells tested, our most potent compounds, including Intronistat A and Intronistat B, did not show toxicity in human cells after 24 h incubation (Fig. 2; Supplementary Fig. 5a; Supplementary Table 2). Even after 72 h of incubation, Intronistat B had only mild effects on cell viability (Supplementary Fig. 5b), suggesting that this compound specifically targets yeast strains that contain group II introns in an essential gene and that it lacks cross-reactivity with the nuclear spliceosome, or other targets, in yeast and humans.

Discussion

In this study, we have demonstrated that the tertiary structure of a large, complex RNA molecule can be targeted in de novo and that high-affinity compounds can be identified using classical methods for developing pharmacologically active compounds. The success of this study implies that high-complexity RNA tertiary structures, such as those found in viral genomes, in the untranslated regions of human mRNAs, pre-miRNA clusters, long intergenic noncoding RNAs and other RNAs essential for the regulation of gene expression, can potentially be targeted with specific, high-affinity small molecules to regulate their function.

An equally significant aspect of this project was the experimental approach used for RNA targeting. Most recent examples of RNA targeting by small molecules involve the use of rational drug design or the serendipitous discovery of natural products. However, in this example, we used a classical method that has been refined through decades of investigations aimed at identifying inhibitors of
The most potent inhibitors that we identified contain a gallate moiety, which is notable because related polyphenolic compounds, such as catechols, are sometimes found in promiscuous PAINS compounds. Although the presence of a gallate or catechol moiety necessitates caution, it should not disqualify a compound scaffold, because catechol reactivity is highly dependent on the architecture of the surrounding molecule. Selectivity and appropriate behavior of such molecules must be established through rigorous SAR, reversibility, binding selectivity, cytotoxicity and other applicable methods, as in this study. Although caution is appropriate, recent studies demonstrate that more than 90% of small molecules that are flagged in automated PAINS screens (such as FAF-drugs3) are well-behaved compounds that have a low hit rate in conventional α-screen assays. Indeed, there is considerable recent concern that valuable bioactive compounds are being removed from the pool of informative chemical space by PAINS screens. This is a particularly serious problem in the development of RNA ligands, as the chemical space for RNA recognition is not well understood, and what is known from studies of RNA binding to macromolecules suggests that base stacking and hydrogen bonding, particularly to hydroxyls, drives tight binding. For these reasons, the gallate moiety in our group II intron inhibitors may be representative of a useful RNA binding motif, but its inclusion was necessarily accompanied by extensive selectivity analysis.

Finally, this study is significant because of its practical applications. Having shown that group II introns can be targeted with small molecules, we used the resulting compounds to inhibit the growth of yeast pathogens that uniquely depend on group II introns for metabolic function. Indeed, the best inhibitors we identified in this study have MIC values (~2 μg/ml) comparable to that of amphotericin B (0.5–1 μg/ml), which is a drug that is still used to treat severe fungal infections despite serious side effects. Consistent with the lack of group II introns in vertebrate animals, our most potent inhibitor, Intronistat B, is not toxic in mammalian cells and does not influence spliceosomal RNA processing, suggesting that this molecule has potential for further development as an antifungal therapeutic. MIC values for C. parapsilosis are in good agreement with in vitro K, values for the S. cerevisiae α5Y intron, indicating that the inhibitors function similarly in different yeast species that contain group II introns, suggesting a route for broad-spectrum fungal inhibitors. Given the unique RNA metabolism of plants, fungi and yeast, our results demonstrate that RNA targeting may provide a much needed approach for developing therapeutics against eukaryotic pathogens.

Given the vast number of physiologically important RNA tertiary structures that control gene expression in all domains of life, the ability to regulate RNA function with small molecules represents a new frontier in molecular medicine. Here we present a discovery pipeline for targeting RNA tertiary structures de novo and we demonstrate the pharmacological utility of small-molecule modulators by using them to specifically inhibit self-splicing group II introns, thereby disrupting the growth of pathogenic yeast. This demonstrates that RNA tertiary structures are attractive, accessible targets for the development of new probes and therapeutics.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41589-018-0142-0.

Received: 16 January 2018; Accepted: 29 August 2018; Published online: 15 October 2018

References
1. Blount, K. F. & Breaker, R. R. Riboswitches as antibacterial drug targets. Nat. Biotechnol. 24, 1558–1564 (2006).
2. Howe, J. A. et al. Selective small-molecule inhibition of an RNA structural element. Nature 526, 672–677 (2015).
ARTICLES

3. Howe, J. A. et al. Atomic resolution mechanistic studies of ribocil: a highly selective unnatural ladinic mimic of the E. coli FMN riboswitch. *RNA Biol.* **13**, 946–954 (2016).

4. Wang, H. et al. Dual-targeting small-molecule inhibitors of the *Staphylococcus aureus* FMN riboswitch disrupt riboflavin homeostasis in an infectious setting. *Cell Chem. Biol.* **24**, 576–588 (2017).

5. Kim, J. N. et al. Design and antimicrobial action of purine analogues that bind guanine riboswitches. *ACS Chem. Biol.* **4**, 915–927 (2009).

6. Blount, K. F. et al. Novel riboswitch-binding flavin analogue that protects mice against *Clostridium difficile* infection without inhibiting cecal flora. *Antimicrob. Agents Chemother.* **59**, 5736–5746 (2015).

7. von Ahlsen, U., Davies, J. & Schroeder, R. Antibiotic inhibition of group I ribonucleases. *Nature* **353**, 368–370 (1991).

8. Disney, M. D., Childs, J. L. & Turner, D. H. Hoechst 33258 selectively inhibits group I intron self-splicing by affecting RNA folding. *ChemBioChem* **5**, 1647–1652 (2004).

9. Mei, H. Y., Cui, M., Lemrow, S. M. & Czarnik, A. W. Discovery of selective, small-molecule inhibitors of RNA complexes. II. Self-splicing group I intron riboswitch. *Bioorg. Med. Chem.* **25**, 1185–1195 (1997).

10. Patwardhan, N. N. et al. Amiloride as a new RNA-binding scaffold with activity against HIV-1 TAR. *MedChemComm* **8**, 1022–1036 (2017).

11. Schroeder, R., Waldsich, C. & Wank, H. Modulation of RNA function by aminoglycoside antibiotics. *EMBO J.* **19**, 1–9 (2000).

12. Colameco, S. & Elliot, M. A. Non-coding RNAs as antibiotic targets. *Biochem. Pharmacol.* **133**, 29–42 (2017).

13. Wilson, D. N. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat. Rev. Microbiol.* **12**, 35–48 (2014).

14. Blaha, G. M., Polonov, Y. S. & Steitz, J. A. Elements of ribosomal drug resistance and specificity. *Curr. Opin. Struct. Biol.* **22**, 750–758 (2012).

15. Eubanks, C. S., Forte, J. E., Kapral, G. J. & Hargrove, A. E. Small molecule-based pattern recognition to classify RNA structure. *J. Am. Chem. Soc.* **139**, 409–416 (2017).

16. Morgan, B. S., Forte, J. E., Culver, R. N., Zhang, Y. & Hargrove, A. E. Discovery of key physicochemical, structural, and spatial properties of RNA-targeted bioactive ligands. *Angew. Chem. Int. Ed.* **56**, 13498–13502 (2017).

17. Bernat, V. & Disney, M. D. RNA structures as mediators of neurological diseases and as drug targets. *Neuron* **87**, 28–46 (2015).

18. Costales, M. G. et al. Small molecule inhibition of microRNA-210 molecules; M.C.V. designed small-molecule inhibitors; A.M.P., O.F., G.E.J. and R.L.A. followed: O.F. performed in vitro biochemical studies of splicing inhibition, carried out cytotoxicity and MIC experiments; G.E.J. designed small-molecule inhibitors and carried out the role in mitochondrial RNA metabolism. *J. Biol. Chem.* **282**, 1603–1611 (2003).

19. Schroeder, R., Waldsich, C. & Pyle, A. M. An oblige intermediate along the slow folding pathway of a group II intron ribozyme. *Nucleic Acids Res.* **33**, 6674–6687 (2005).

20. Perez-Martinez, X., Broadway, S. A. & Fox, T. D. Mas51p promotes mitochondrial Cox1p synthesis and interacts with newly synthesized Cox1p. *EMBO J.* **22**, 5951–5961 (2003).

21. Dziembowski, A. et al. Yeast mitochondrial degradosome. Its composition, interplay between RNA helicase and RNase activities and the role in mitochondrial RNA metabolism. *J. Biol. Chem.* **278**, 1603–1611 (2003).

22. Luedtke, N. W., Liu, Q. & Tor, Y. RNA-ligand interactions: affinity and specificity of aminoglycoside dimers and acridine conjugates to the HIV-1 Rev response element. *Biochemistry* **42**, 11391–11403 (2003).

23. Tanner, M. & Cech, T. Activity and thermostability of the small self-splicing group I intron in the pre-rRNA(ile) of the purple bacterium *Azorarcus* RNA 2, 74–83 (1996).

24. Li, C. F., Costa, M., Bassi, G., Lai, Y. K. & Michel, F. Recurrent insertion of 5′-terminal nucleotides and loss of the branchpoint motif in lineages of group II introns inserted in mitochondrial preribosomal RNAs. RNA **17**, 1321–1335 (2011).

25. Moen, M. D., Lyseng-Williamson, K. A. & Scott, L. J. Liposomal amphotericin B: a review of its use as empirical therapy in febrile neutropenia and in the treatment of invasive fungal infections. *Drugs* **69**, 361–392 (2009).

26. Velagapudi, S. P. et al. Design of a small molecule against an oncogenic microRNA. *Proc. Natl. Acad. Sci. USA* **113**, 5898–5903 (2016).

27. Mullbach, J. et al. Novel riboswitch ligand analogs as selective inhibitors of guanine-related metabolic pathways. *PLoS Pathog.* **6**, e1000685 (2010).

28. Baell, J. & Walters, M. A. Chemistry: chemical con artists foil drug discovery. *Nature* **513**, 481–483 (2014).

29. Baell, J. B. & Nissim, J. W. M. Seven year itch: pan-assay interference compounds (PAINS) in 2017—utility and limitations. *ACS Chem. Biol.* **13**, 36–44 (2018).

30. Jaisal, S., Hu, Y. & Bajorath, J. How frequently are pan-assay interference compounds active? Large-scale analysis of screening data reveals diverse activity profiles, low global hit frequency, and many consistently inactive compounds. *J. Med. Chem.* **59**, 3879–3886 (2016).

31. Capuzzi, S. J., Muratov, E. N. & Tropsha, A. Phantom PAINS: problems with the utility of alerts for pan-assay INterference Compounds. *J. Chem. Inf. Model.* **57**, 417–427 (2017).

32. Re, A., Joshi, T., Kulbryktye, E., Morris, Q. & Workman, C. T. RNA-protein interactions: an overview. *Methods Mol. Biol.* **1097**, 491–521 (2014).

Acknowledgements

We thank S. Umlauf, P. Gareiss and J. Merkel at Yale Center for Molecular Discovery for their help with high-throughput screening. We are grateful to K. Blount and R. Breaker for help with setting up MiC experiments. We thank J. Sinclair and A. Schepartz for sharing their expertise on cytotoxicity experiments. We gratefully acknowledge T. Fox (Cornell University) for sharing wild-type and mtDNA intronless *Candida* strains. We thank S. Woodsen (Johns Hopkins University) for sharing the Azo-pre-tRNA plasmid. We thank S. Herzon and R. Breaker for comments on the manuscript. We are grateful to C. Zhao (Cornell University) for sharing wild-type and mtDNA intronless *Candida* strains. We thank S. Woodsen (Johns Hopkins University) for sharing the Azo-pre-tRNA plasmid. We thank S. Herzon and R. Breaker for comments on the manuscript. We are grateful to C. Zhao for help in making Supplementary Fig. 1. A.M.P. is an Investigator, and O.F. is a Research Specialist in the Howard Hughes Medical Institute. This work was supported by NIH grants RO1GM50313 to A.M.P. and R43 AI115951 to M.V.Z.

Author contributions

A.M.P., O.F. and G.E.J. designed the study. All authors contributed to this work as follows: O.F. performed in vitro biochemical studies of splicing inhibition, carried out cytotoxicity and MiC experiments; G.E.J. designed small-molecule inhibitors and carried out the role in mitochondrial RNA metabolism. R.L.A. wrote the paper.

Competing interests

Yale University has filed a provisional patent application on the work developed in this manuscript.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41558-018-0142-0.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to A.M.P.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2018
Methods

Yeast strains. Strains of Candida parapsilosis (ATCC 22019) and Sacharomyces cerevisiae (ATCC 18824) were purchased from American Type Culture Collection (ATCC) and cultured according to the manufactures (ATCC; S. cerevisiae-type (NP40-36a) and mtDNA intromless (XPM46)). Strains were kindly provided by T. Fox.

RNA preparation. Synthesis of the RNA oligo substrates. RNA oligonucleotides containing 3'-terminal Black Hole Quencher 2 label and amino-modifier C6dT nucleotide (Glen Research) as well as U2-6 RNA oligonucleotide (5'-AGC AGU UCC CUC GCA AAU AUG AUG AAC CGC U) were synthesized on a MerMade 12 DNA-RNA synthesizer (BioAutomation) using TBDSMs phosphoramidites (Glen Research). Base deprotection was carried out in a 3:1 mixture of 30% ammoniumhydroxide (JT Baker) and ethanol at room temperature (25-25 oC) for 24h. Subsequent 2'-OH deprotection and purification was performed on a 20% denaturing polyacrylamide gel as previously described (33). RNA oligonucleotides U2-6 were deprotected and purified as described (33).

In vitro transcription. Large-scale transcription of the SE, D1 (domain 1 in isolation), D3 (domain 3 in isolation), D56 (RNA molecule containing both domains 5 and 6) and D135 RNAs (34) was carried out using T7 RNA polymerase and purified on a 20% denaturing polyacrylamide gel. Products were purified on a 20% denaturing polyacrylamide gel.

Fluorescent labeling of RNA oligonucleotides. Purified RNA oligonucleotides containing 3'-terminal Black Hole Quencher 2 label and amino-modifier C6dT nucleotide (Glen Research) were fluoresce fluoro labeled with the NHS ester of AlexaFluor 555 dye (Life Technologies Corp.) at the primary amino group located on the amino-modifier C6dT nucleotide. RNA oligonucleotides were dissolved in 200 μl of 0.25 M sodium bicarbonate buffer (pH 9.2) and then combined with a solution containing 0.5 mg of AlexaFluor 555 NHS ester in 200 μl formamide. The reaction was incubated at room temperature (25-25 oC) for 2 h, and the labeled products were purified on a 20% denaturing polyacrylamide gel.

High-throughput screening. High-throughput screening was carried out using a library of 10,000 compounds selected from the following collections: NCI Oncology (85 compounds), NCI Diversity (1,356 compounds), ENZO kinase inhibitors (80 compounds), 640 FDA-approved drugs, ENZO phosphatase inhibitors (33 compounds), BML-ENZO ion channel ligands (72 compounds), BML-metabotropic glutamate ligands (56 compounds), BML nuclear receptor ligands (76 compounds), protease inhibitors (53 compounds) and additional compounds from ChemBridge MW and ChemDiv to bring the total number of compounds to 10,000.

The D135 ribozyme (31), fluoresce fluoro labeled oligo substrate 17/2 D1 (Supplementary Fig. 2) (both at 20 nM final concentrations in 50 mM MOPS, pH 7.0, 100 mM MgCl2, 0.5 mM KCl), and the test compounds (final concentration of 10 μM in 50 mM MOPS, pH 7.0, 100 mM MgCl2, 0.5 mM KCl) were combined in black nonbinding 384-well plates (Corning 3757) using multicolor combi (Thermo Fisher Scientific) and pinto (V & P Scientific) pipettors. The reaction mixtures were incubated at 37 oC for 45 min, and then quenched with 100 mM EDTA. Fluorescence intensity was analyzed on a Tecan Infinite multimode plate reader (λex = 520nm, λem = 560 nm, 5-mm bandwidth). Data were normalized relative to the untreated wells and to the wells lacking the D135 ribozyme, and percent of inhibition was calculated using ActivityBase (IDBS). The Z′-factor was calculated as described (32). Z′ = 1 - 3(σo + σp) / (σo - σp), where σo and σp are mean value and s.d. for positive control (compound) and negative control (water), respectively. Percent inhibition is defined as (100 × (control - compound) / control).

Inhibition of group I intron splicing. Splicing of the Azo-Pre-tRNA intron was carried out essentially as previously described (31). Internally labeled Azo-Pre-tRNA (20 nM) was incubated in 25 mM HEPES, pH 7.5, 10 mM MgCl2, at 30 °C for 10 min, and then at 32 °C for 2 min. Then (19) (Intronistat B) was added to the final concentrations of 1, 0.5, 0.25, and 0.125 μM and reaction was initiated by addition of 100 μM GTP (final concentration). The final concentration of DMSO in all samples was 10%. Aliquots at different time points were quenched and analyzed on a 5% denaturing polyacrylamide gel as previously described (33). Data were fit with a single-exponential equation to determine the first order rate constants (kobs). Data were performed in triplicate to ensure reproducibility.

Effect of exposure of various RNAs on splicing inhibition. Internally labeled SE RNA (2 nM final concentration) and unlabeled SE, D1, D3, D56 RNAs, yeast tRNA Phe and U2-6 RNA oligo (2 μM final concentrations) were preincubated separately in 50 mM MOPS, pH 7.5, 100 mM KCl and 5 mM MgCl2 (10 mM MgCl2, for yeast tRNA), at 30 °C for 20 min. The labeled SE RNA and unlabeled SE RNA solutions were mixed together with simultaneous addition of 19 (Intronistat B) (300 nM final concentration) to initiate the reaction. Reaction was carried out under near-physiological conditions (in 50 mM MOPS, pH 7.5, 8 mM MgCl2, 100 mM KCl) at 30 °C. Aliquots at different time points were quenched and analyzed on a 5% denaturing polyacrylamide gel as previously described (33). Data were fit with a single-exponential equation to determine the first order rate constants (kobs). Data were performed in triplicate to ensure reproducibility.

Determination of Ki values for small-molecule inhibition of the self-splicing reaction. Internally labeled SE RNA (2 nM) was incubated with various concentrations of inhibitor compound under near-physiological conditions (in 50 mM MOPS, pH 7.5, 8 mM MgCl2, 100 mM KCl) at 30 °C. Aliquots at different time points were quenched and analyzed on a 5% denaturing polyacrylamide gel as previously described (33). Data were fit to a single-exponential equation to determine the first order rate constants (kobs). The latter were plotted against the concentration of inhibitor and fit to the equation for noncompetitive inhibition to determine K_i values: kobs = k∞ / ([I] + [I]/K_i), where k∞ and K_i are the first order rate constants measured in the presence and in the absence of the inhibitor, respectively. [I] is the concentration of the inhibitor and K_i is the inhibition constant. Experiments were performed four times for compound 3, three times for 14, 18 (Intronistat A) and 19 (Intronistat B) and twice for the remaining compounds to ensure reproducibility. Data represent average ± s.e.m.

Effect of concentration of protein on self-splicing rate.

YEAST RESPIRATION ASSAY IN S. CEREVISIAE. Yeast respiration assays were conducted either in liquid YPD media (BD Bacto Yeast Extract, BD Bacto Peptone, 2% glucose) or in YPG media (Bio-Express Yeast Extract, BD Bacto Peptone, 3% glycerol, 3% ethanol), essentially as described for the antifungal MIC assays (see below) using the guidelines from the Clinical and Laboratory Standards Institute (31). To initiate these experiments, 100 μl of a fresh stock solution was prepared for each compound (3.2 mg/ml for amphotericin B and 12.8 mg/ml for each test compound) in DMSO stock solution was placed in the well of a 96-well plate on a 96-well plate. The subsequent nine wells in this row contained 50 μl of DMSO for serial dilution. From the first well, 50 μl of the compound stock was withdrawn, transferred to the second well and mixed. This process was repeated from the second and subsequent wells, resulting in a 1:2 serial dilution of the stock into DMSO. Wells 11 and 12 contained DMSO only, to be used as the no-compound control and sterility control, with no inoculum present (sterility substrate only), and where μo and σo are the mean value and s.d. for negative control (no compound). The average Z′-factor from the screen was 0.83 ± 0.03.

Determination of K_i values for small-molecule inhibition of the self-splicing reaction. Internally labeled SE RNA (2 nM) was incubated with various concentrations of inhibitor compound under near-physiological conditions (in 50 mM MOPS, pH 7.5, 8 mM MgCl2, 100 mM KCl) at 30 °C. Aliquots at different time points were quenched and analyzed on a 5% denaturing polyacrylamide gel as previously described (33). Data were fit to a single-exponential equation to determine the first order rate constants (kobs). The latter were plotted against the concentration of inhibitor and fit to the equation for noncompetitive inhibition to determine K_i values: kobs = k∞ / ([I] + [I]/K_i), where k∞ and K_i are the first order rate constants measured in the presence and in the absence of the inhibitor, respectively. [I] is the concentration of the inhibitor and K_i is the inhibition constant. Experiments were performed four times for compound 3, three times for 14, 18 (Intronistat A) and 19 (Intronistat B) and twice for the remaining compounds to ensure reproducibility. Data represent average ± s.e.m.

Effect of concentration of protein on self-splicing rate.

Identification of candidate compounds.

C. elegans measurement for the small-molecule inhibition of the D135 ribozyme cleavage reaction. Black 96-well plates (Corning 3902) were filled with 50 μl of solution containing 20 nM D135 ribozyme, 20 nM double-labeled substrate 17/2 D1 and small-molecule inhibitor in 50 mM MOPS, pH 7.0, 100 mM MgCl2, and 500 mM KCl. Small-molecule inhibitors were tested at 19 different concentrations ranging from 5 nM to 1 μM. Plates were incubated at 37 °C for 10 min, and reaction mixtures were quenched with 100 mM EDTA and then analyzed on a Synergy H1 plate reader (BioTek). Each experiment was performed in triplicate. Data were fit to a 4-parameter logistic function c = a + (d - c) / (1 + (x/a)^n), where a is the IC50, c is the slope parameter, ε is the minimum response and d is the maximum response. Data are reported as average ± s.e.m.
or 18 (YPGE 18 (Intronist A) plate) at 30 °C. Experiment was replicated twice to ensure reproducibility of growth phenotype.

**Analysis of splicing in S. cerevisiae by qRT-PCR.** A culture of *S. cerevisiae* was grown overnight with shaking in YPGE (see above) at 30 °C. The saturated culture was then diluted and grown to mid-log phase in YPGE. Equal volumes of DMSO or compounds dissolved in DMSO were added to individual cultures (final concentrations of 19 (Intronist B): 64 μg/ml; 18 (Intronist A): 52 μg/ml; 12: 32 μg/ml; 6: 64 μg/ml; amphotericin B: 8 μg/ml) and incubated for 4 h. Cells were harvested by centrifugation, washed with 500 μl ice-cold Milli-Q water, centrifuged again, and snap-frozen in liquid nitrogen. Total RNA was isolated using Z.N.A. Yeast Purification Kit (Omega Bio-tek) according to the manufacturer’s procedures.

The RNA was eluted in 50 μl of RNase/DNase-free water (Thermo Fisher), followed by DNase treatment with RQ1 DNase (Promega), for 1 h at 37 °C. Then, it was mixed with 6 μl of 5 M NaOAc and precipitated with 75% EtOH at −20 °C overnight. The RNA was then reverse transcribed with SuperScript III (Thermo Fisher) following the manufacturer’s recommendations with 200 ng Random Hexamer Primers (Thermo Fisher) and 40 U RNasin (Thermo Fisher) in a 20 μl reaction. After reverse transcription, RNA was degraded by addition of 2 N NaOH (2 μl) and incubation at 95 °C for 5 min. After cooling on ice for 5 min, 2 μl of 1 M HCl and 2 μl of 3 M NaOAc were added to the reaction mixture, and cDNA was precipitated with 75% EtOH at −20 °C for 30 min and resuspended in RNase/DNase-free water. The cDNA levels were quantified with real-time PCR using LightCycler 480 SYBR Green I (Roche) and a CFX384 Real-Time PCR Detection System (Bio-Rad) in triplicate with independent samples. Relative levels of indicated RNA species from different conditions were normalized according to the ΔΔCt method following the equation:

\[ \Delta \Delta C_t = \left[ C_{\text{total}} \text{ or unspliced} - C_{\text{ACTI or PGK1}} \right] \]

**Cytoxicity in HEK-293T cells.** HEK-293T cells were cultured in DMEM (Gibco) supplemented with 10% Fetal Bovine Serum (FBS, Gibco) and 100 U/ml of Penicillin-Streptomycin (Gibco) at 37 °C and 5% CO₂. For the cytoxicity experiments, cells were aliquoted into black 96-well plates with a clear bottom (Corning 3603) at a concentration of 10,000 cells per well. The cells were grown at 37 °C, 5% CO₂ for 5–6 h, then the medium was replaced with the same medium without FBS and cells were grown at 37 °C, 5% CO₂ for 24 h. Freshly prepared stock solutions of test compounds (12.8 mg/ml in DMSO) were serially diluted 1:2 with DMSO into the first 11 successive wells of a 12-well plate on a 96-well plate as described above. Well 12 was used as a growth control (no compound). After dilution, 1 μl from each well of the compound plate was added to 94 μl of Penicillin-Streptomycin (Gibco) containing 2 μl of the assay plate, and cells were incubated at 37 °C, 5% CO₂ for another 24 or 72 h. After incubation, cell viability was determined using the luminescence Cell Titer Glo cell viability assay (Promega), in which 100 μl of the assay reagent was added to each well of the assay plate. After gentle shaking for 7–10 min, the plates were analyzed on a Synergy H1 plate reader (BioTek). Luminescence was plotted against the compound concentration and IC₅₀ values were determined by fitting the data to a 4-parameter logistic function c + (d−c)/(1 + (x/a)), where a is the IC₅₀, b is the slope parameter, c is the minimum response and d is the maximum response. Experiments were replicated four times for compounds 2, 3, 5, 7, 9, 11, 12, 15, 18 and 19, three times for compounds 1, 6, 8, 10, 13, 14, 16 and 17, and twice for compound 4 to ensure reproducibility.

**Statistics and reproducibility.** The following reported data represent average of n = 3 independent experiments: all IC₅₀ values for ribozyme cleavage, K values for compounds 14, 18 and 19, kₐ values for inhibition of splicing in the presence of the excess of various RNAs, time courses for inhibition of the group I intron splicing. ΔΔCₕ values obtained from qRT-PCR analysis of group II intron splicing in *C. parapsilosis*, MIC values from yeast respiration assay, MIC values for small molecules against *C. parapsilosis*, IC₅₀ values for cytotoxicity in HEK-293T cells for compounds 1, 6, 8, 10, 13, 14, 16 and 17.

The following reported data represent average of n = 4 independent experiments: K for compound 3, ΔΔCₕ values obtained from qRT-PCR analysis of group II intron splicing in *S. cerevisiae*, IC₅₀ values for cytotoxicity in HEK-293T cells for compounds 1, 2, 3, 5, 7, 9, 11, 12, 15, 18 and 19.

The following experiments were repeated twice to ensure reproducibility: test of reversibility of compound binding, *S. cerevisiae* growth assays, cytotoxicity of compound 4 in HEK-293T cells, K determination for compounds 1, 2, 4–13 and 15–17.

All values are reported as mean ± s.e.m.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability.** Authors can confirm that all relevant data are included in the paper and/or its supplementary information files.

**References.**

47. Wincott, F. et al. Synthesis, deprotection, analysis and purification of ribozymes and ribonucleases. *Nucleic Acids Res.* 23, 2677–2684 (1995).

48. Dickey, T. H. & Pyle, A. M. The SMAD3 transcription factor binds complex RNA structures with high affinity. *Nucleic Acids Res.* 45, 11980–11988 (2017).

49. Chin, K. & Pyle, A. M. Branch-point attack in group II introns is a highly cooperative process that enables ribozyme activity: quantitation of interdomain binding and reaction rate. *Biochemistry* 33, 2716–2725 (1994).

50. Ziegler, N., Solem, A. & Pyle, A. M. Dual roles for the Mss116 cofactor during the assembly of reverse transcriptase and the qRT-PCR assay. 8: 6602–6609 (2010).
53. Daniels, D. L., Michels, W. J. Jr & Pyle, A. M. Two competing pathways for self-splicing by group II introns: a quantitative analysis of in vitro reaction rates and products. *J. Mol. Biol.* **256**, 31–49 (1996).

54. Zhang, J. H., Chung, T. D. & Oldenburg, K. R. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen.* **4**, 67–73 (1999).

55. Clinical and Laboratory Standards Institute. *Reference Method For Broth Dilution Antifungal Susceptibility Testing of Yeasts; Approved Standard—Third Edition* (Clinical and Laboratory Standards Institute, Wayne, PA, 2008).

56. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2^−ΔΔCT method. *Methods* **25**, 402–408 (2001).
Experimental design

1. Sample size
   Describe how sample size was determined.
   We aimed to replicate each experiment sufficient number of times to ensure reproducibility of the results. Most data presented in the paper represent an average of n=3 independent experiments.

2. Data exclusions
   Describe any data exclusions.
   No data were excluded from the analyses

3. Replication
   Describe the measures taken to verify the reproducibility of the experimental findings.
   All experimental findings presented in the paper have been successfully replicated 2-4 times to ensure reproducibility.

4. Randomization
   Describe how samples/organisms/participants were allocated into experimental groups.
   This is not relevant to our study, because it does not involve animal or human testing

5. Blinding
   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.
   This is not relevant to our study, because it does not involve animal or human testing

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

6. Statistical parameters
   For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

   n/a        Confirmed
   
   □ ☑ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
   A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
   □ ☑ A statement indicating how many times each experiment was replicated
   □ ☑ The statistical test(s) used and whether they are one- or two-sided
     Only common tests should be described solely by name; describe more complex techniques in the Methods section.
   □ ☑ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
   □ ☑ Test values indicating whether an effect is present
     Provide confidence intervals or give results of significance tests (e.g. P values) as exact values whenever appropriate and with effect sizes noted.
   □ ☑ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
   □ ☑ Clearly defined error bars in all relevant figure captions (with explicit mention of central tendency and variation)

   See the web collection on statistics for biologists for further resources and guidance.
Software

7. Software

Describe the software used to analyze the data in this study.

Commercial software ImageQuant, Microsoft Excel, KaleidaGraph and GraphPad Prism was used for data analysis.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). Nature Methods guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.

The only unique materials used in this paper are chemical compounds. The structure and detailed description of synthesis of each compound is provided in Methods.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

b. Describe the method of cell line authentication used.

c. Report whether the cell lines were tested for mycoplasma contamination.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

Human cell line HEK-293T, C. parapsilosis strain ATCC 22019 and S. cerevisiae strain ATCC 18824 were purchased from American Type Culture Collection (ATCC). S. cerevisiae Wild-type (NP40-36a) and mtDNA intronless (XPM46) strains were kindly provided by Dr. Thomas Fox.

The cell line HEK-293T purchased from ATCC has been authenticated using STR matching analysis at the DNA analysis facility on Science Hill at Yale University. All yeast strains were validated by PCR.

All cell lines tested negative for mycoplasma contamination.

No misidentified cell lines were used.

Animals and human research participants

11. Description of research animals

Provide all relevant details on animals and/or animal-derived materials used in the study.

No animals were used.

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

The study did not involve human research participants.