The mitochondrial tyrosyl-tRNA synthetases (mtTyrRSs) of Pezizomycotina fungi, a subphylum that includes many pathogenic species, are bifunctional enzymes that both charge mitochondrial tRNA$^{3\text{tyr}}$ and act as splicing cofactors for autocatalytic group I introns. Previous studies showed that one of these proteins, Neurospora crassa CYT-18, binds group I introns by using both its N-terminal catalytic and C-terminal anticodon binding domains and that the catalytic domain uses a newly evolved group I intron binding surface that includes an N-terminal extension and two small insertions (insertions 1 and 2) with distinctive features not found in non-splicing mtTyrRSs. To explore how this RNA binding surface diverged to accommodate different group I introns in other Pezizomycotina fungi, we determined x-ray crystal structures of C-terminally truncated Aspergillus nidulans and Coccidioides posadasii mtTyrRSs. Comparisons with previous N. crassa CYT-18 structures and a structural model of the Aspergillus fumigatus mtTyrRS showed that the overall topology of the group I intron binding surface is conserved but with variations in key intron binding regions, particularly the Pezizomycotina-specific insertions. These insertions, which arose by expansion of flexible termini or internal loops, show greater variation in structure and amino acids potentially involved in group I intron binding than do neighboring protein core regions, which also function in intron binding but may be more constrained to preserve mtTyrRS activity. Our results suggest a structural basis for the intron specificity of different Pezizomycotina mtTyrRSs, highlight flexible terminal and loop regions as major sites for enzyme diversification, and identify targets for therapeutic intervention by disrupting an essential RNA-protein interaction in pathogenic fungi.

Among the best-studied examples of such gain-of-function are the mitochondrial tyrosyl-tRNA synthetases (mtTyrRS) of fungi belonging to the subphylum Pezizomycotina, which evolved to promote the splicing of mitochondrial (mt) group I introns (4). Group I introns are ribozymes that catalyze their own splicing via guanosine-initiated transesterification reactions (5). Although some self-splice in vitro, most have acquired mutations that impair self-splicing and have thus become dependent upon intron-encoded or cellular proteins to promote formation of the catalytically active ribozyme structure (6). The Pezizomycotina mtTyrRS2 evolved to function in splicing via a series structural adaptations, which occurred during or after the divergence of Pezizomycotina from Saccharomycotina and resulted in a new group I intron binding surface distinct from that which binds tRNA$^{3\text{tyr}}$ (4, 7). These structural adaptations included acquisition of a longer N-terminal extension and several small insertions (insertions (Ins) 1–5) resulting from the expansion of flexible terminal or loop regions (4). Thus, these enzymes provide an ideal model system for investigating how proteins in general and aminoacyl-tRNA synthetases in particular acquire new RNA binding functions.

Thus far most studies of how fungal mtTyrRSs function in splicing group I introns have focused on the Neurospora crassa mtTyrRS, denoted CYT-18 protein, which was identified as a group I intron splicing factor in an early genetic screen for splicing-deficient mutants (8). Temperature-sensitive mutants in the cyt-18 gene were found to be defective in splicing the N. crassa mt large subunit rRNA (mtLSU) intron and two other group I introns (8, 9). Further studies showed that CYT-18 is by itself sufficient to promote the splicing of these introns as well as diverse group I introns from other organisms in vivo and in vitro (10–12).

CYT-18 promotes splicing by binding specifically to group I intron RNAs and stabilizing their catalytically active RNA structure (13). Group I introns have minimal sequence conservation but fold into a conserved tertiary structure that is required for ribozyme activity (14–17). This conserved structure consists of two extended helical domains, denoted the P4-P6 and P3-P9 domains, whose juxtaposition creates an active-site cleft that binds the splice sites and the guanosine 2′-O-methylguanosine (G) of intron 1. Among these, the P4-P6 domain is generally conserved but shows some variation in sequence and structure. The P3-P9 domain is more variable and several groups have proposed a structural basis for intron specificity based on divergent C-terminal arm domains (18). However, it is not clear to what extent structure determines specificity.

Aminoacyl-tRNA synthetases are a class of ancient, essential enzymes that catalyze the ligation of amino acids onto cognate tRNAs. In addition to this essential role, many aminoacyl-tRNA synthetases have evolved secondary functions, frequently by the acquisition of new domains or sequence expansions (1–3).
co-factor and uses specifically bound Mg\(^{2+}\) ions for catalysis (15–18). A 4.5-Å co-crystal structure of a C-terminally truncated CYT-18 protein bound to the Twort group I intron RNA ribozyme showed that CYT-18 recognizes the conserved phosphodiester-backbone structure of group I intron RNAs, with few if any base-specific contacts (12). Biochemical and genetic studies suggested a model in which CYT-18 binds first to the P4–P6 domain to promote formation of the correct geometry of the P4–P6-stacked helices and then makes additional contacts with the P3–P9 domain to stabilize the two domains in the correct relative orientation to form the intron RNA active site (13, 19).

mtTyrRSs are class I aminoacyl-tRNA synthetases with a domain architecture similar to that of bacterial TyrRSs. They have N-terminal catalytic and \(\alpha\)-helical domains (together referred to as the NTDs), which are attached via a flexible linker to a C-terminal anticodon binding domain (CTD) (Fig. 1A) (20). Bacterial TyrRSs and mtTyrRSs are homodimers that bind one tRNA\(^{Tyr}\) molecule across the two subunits, with the N-terminal catalytic domain of one subunit binding the acceptor stem and the \(\alpha\)-helical and C-terminal domains of the other subunit binding the variable and anticodon arms (21, 22). The homodimeric CYT-18 protein also binds a single group I intron across both subunits by using both the N-terminal catalytic domain and flexibly attached CTD (23–25). Most contacts to group I intron RNAs are made via the catalytic domain with only some group I introns requiring additional contacts from the CTD for RNA splicing (25).

It was thought initially that CYT-18 might bind group I introns by using its tRNA\(^{Tyr}\) binding site, but structural studies showed that CYT-18 has a distinct group I intron binding surface on the side of the catalytic domain opposite which binds tRNA\(^{Tyr}\) (12, 26). This intron binding surface includes an N-terminal extension (H0) and two small insertions (Ins 1 and 2) that function directly in group I intron binding (12). Three other Pezizomycotina-specific insertions (Ins 3–5) are found in the CTD, whose flexible attachment enables the binding of either a group I intron or tRNA\(^{Tyr}\) on opposite sides of the catalytic domain (12, 26–28). Although full-length CYT-18 could not be crystallized, an NMR structure of the closely related Aspergillus nidulans CTD showed that Ins 3 is an expansion of the flexible linker between the NTDs and CTDs and that Ins 4 and Ins 5 may be involved in both tRNA\(^{Tyr}\) and group I intron binding (28). Small angle x-ray scattering solution structures of full-length CYT-18 by itself and bound to the Twort group I intron showed that the NTDs of the full-length protein bind group I introns in solution in the same manner as the isolated NTDs in the co-crystal structure and that both CTDs of the homodimeric protein move inward to bind oppositely, structurally different ends of the intron RNA (27).

Thus far, biochemical studies have confirmed splicing activity for five Pezizomycotina mtTyrRSs from N. crassa, A. nidulans, Podospora anserina, Coccidioides posadasii, and Histoplasma capsulatum (4, 29, 30). Splicing assays revealed that although all these proteins are closely related and contain the same series of insertions, they preferentially splice their cognate mtLSU group I intron and have variable abilities to splice the mtLSU introns of other fungi (4). This intron specificity suggests that the mtTyrRS in different fungi co-evolved with their cognate introns to splice them efficiently, and in some cases this led to decreased ability to splice non-cognate group I introns.

The Pezizomycotina fungal subphylum encompasses many species of medical and commercial importance, including C. posadasii and A. nidulans whose mtTyrRSs are studied in this work (31). C. posadasii is a causative agent of coccidioidomycosis, a respiratory disease that ranges from mild (asymptomatic) to life threatening, particularly in immunocompromised individuals (32). A. nidulans is closely related to pathogenic Aspergillus species, including Aspergillus fumigatus (Af), the second-most prevalent species in invasive fungal infections and the most common in fatal cases (33). Because Pezizomycotina fungi are generally obligate aerobes, the splicing of mt group I introns by mtTyrRSs is an essential function that is required for the production of respiratory enzymes. Thus, targeting the fungal-specific group I intron binding surface to prevent splicing without inhibiting TyrRS activity of the human enzyme is a potential therapeutic strategy.

Here, we determined x-ray crystal structures of the N-terminal catalytic and \(\alpha\)-helical domains of the mtTyrRSs from A. nidulans (An NTDs) and C. posadasii (Cp NTDs). We then compared these structures to previously determined structures of CYT-18 by itself and bound to a group I intron RNA and used them to construct a structural model of the Af mtTyrRS. The comparisons showed that Ins 1 is structurally similar between the different Pezizomycotina enzymes and likely interacts with intron RNA in a manner similar to Ins 1 of CYT-18. By contrast, the N-terminal H0 extension and Ins 2 vary significantly. Our results suggest a structural basis for the group I intron binding specificities of different Pezizomycotina mtTyrRSs, identify regions of these proteins as potential drug targets, and provide further insight into how fungal mtTyrRSs evolved to function in RNA splicing.

Experimental Procedures

Recombinant Plasmids—Recombinant plasmids used for the expression of fungal mtTyrRSs in Escherichia coli are derivatives of the phage T7 promoter-driven expression vector pET11d (Novagen). All protein constructs lack the N-terminal mt targeting sequence, which was determined experimentally for CYT-18 (24) and predicted for the other proteins by MITOPROT (34). Wild-type CYT-18 (residues 33–669) was expressed in E. coli HMS174(DE3) from plasmid pEX560 (24).

Plasmids pHISTEVAn602 and pHISTEVCP602 express the A. nidulans mtTyrRS NTDs (residues 60–462) and C. posadasii mtTyrRS NTDs (residues 46–429), respectively, with an N-terminal hexahistidine tag that can be removed by cleavage with tobacco etch virus (TEV) protease. These plasmids were constructed by cloning the hexahistidine-tagged A. nidulans and C. posadasii NTD sequences from pHISTEVAnTyrRS and pHISTEVCPtYRS, respectively (4), between the NcoI and BamHI sites of pET11d. Cleavage of the expressed mtTyrRS proteins at the TEV site results in an extra methionine at the N-terminus of the An NTDs and an extra glycine at the N-terminus of the Cp NTDs.

The Af mtTyrRS was expressed from plasmid pAfYRS-DMTP. This plasmid was constructed by separately amplifying

\[11912 \text{ JOURNAL OF BIOLOGICAL CHEMISTRY} \]
each of the five exons of the mtTyrRS coding sequence from A. fumigatus strain 293 genomic DNA (35) (provided by L. Losada, J. Craig Venter Institute), joining them by overlap PCR using a primer that adds a TEV-cleavable N-terminal hexahistidine tag, and cloning the PCR product between the NcoI and BamHI sites of pET11d vector (Novagen). The final construct was confirmed by sequencing.

The Af mtLSU intron was transcribed in vitro from pAfmtLSU-ΔORF. This plasmid was constructed by PCR amplifying the intron and flanking 5′- and 3′-exon sequences (70 and 27 nucleotides, respectively) from A. fumigatus strain 293 DNA (35). The intron was amplified in two segments to remove the ORF encoding mt ribosomal protein S5, which is present in L8, using primers that introduce a phage T7 promoter sequence upstream of the 5′ exon, an overlapping HindIII site for ligation of the two PCR products, and terminal BamHI and EcoRI sites for cloning between the corresponding sites of pUC18. Transcription of the EcoRI-digested plasmid with T7 RNA polymerase yields a 390-nucleotide precursor RNA containing the intron.

Protein Expression and Purification—Protein-expression plasmids were transformed into E. coli Rosetta2(DE3) (EMD Millipore) and stored as frozen glycerol stocks. The An NTDs were expressed by autoinduction (36) in LB medium. A 20-ml starter culture, which was inoculated from a freshly streaked out glycerol stock, was grown at 37 °C in LB overnight, and 5 ml was then used to inoculate cultures containing 500 ml of LB media in 2.5-liter Ultra Yield flasks. Cultures were grown with shaking at 260 rpm for 4 h at 37 °C and then shifted to 25 °C overnight. Selenomethionine (Se-Met)-labeled Cp NTDs protein was expressed by autoinduction in minimal medium for selenium-methionine labeling (36). Briefly, a starter culture was grown in 250-ml PASM-5052 media (36) overnight at 25 °C at 260 rpm, and 18 ml was used to inoculate 500 ml of PASM-5052 media in 2.5-liter Ultra Yield flasks containing 125 µg/ml Se-Met (36). The cells were grown for 16 h at 25 °C with 260 rpm shaking. The selenium methionine-labeled Cp NTDs produced better diffracting crystals than did the native protein.

The wild-type CYT-18 protein was purified as described (12, 23). The An and Cp mtTyrRS NTDs and the full-length Af mtTyrRS were purified by nickel affinity chromatography (His-Trap™ HP column; GE Healthcare) (4). Cells were resuspended in 25 ml of 500 mM KCl, 25 mM Tris-HCl, pH 7.5, and lysed by incubation with lysozyme (1 mg/ml) for 30 min followed by sonication and nucleic acid removal by polyethyleneimine precipitation (final concentration of 1%). The proteins were then precipitated with 40% ammonium sulfate, and the resulting pellet was dissolved in nickel affinity column binding buffer (500 mM KCl, 25 mM Tris-HCl, pH 7.5, 30 mM imidazole). The lysate was then run through a HiTrap™ HP column (GE Healthcare), and hexahistidine-tagged proteins were eluted with elution buffer (500 mM KCl, 25 mM Tris-HCl, pH 7.5, 500 mM imidazole). This was followed by removal of the hexahistidine tag by TEV-protease cleavage (200 µl of 2 mg/ml TEV) during dialysis at 4 °C in 500 mM KCl, 25 mM Tris-HCl, pH 7.5, 5 mM DTT, which also removes imidazole. The proteins were then further purified by another round of nickel affinity chromatography as described above followed by size-exclusion chromatography (HiLoad™ 16/60 Superdex™ 200; GE Healthcare). The size-exclusion column purifications were done in the same buffer used for protein crystallization, after which the proteins were concentrated using Amicon Ultra-15 or Ultra-4 centrifugal filter units (EMD Millipore) in buffers optimized for solubility and stored at 4 °C. The An NTDs protein was stored in 200 mM KCl, 25 mM Tris-HCl, pH 7.5. The Cp NTDs Se-Met protein was stored in 250 mM KCl, 25 mM Tris-HCl, pH 7.5 with 1 mM DTT added before setting up sitting drops.

The purified CYT-18 and An and Cp NTDs used for biochemical assays were dialyzed after purification into 100 mM KCl, 25 mM Tris-HCl, pH 7.5, 50% glycerol (v/v) for storage at −80 °C. The purified Af mtTyrRS was dialyzed against 500 mM NaCl, 25 mM Tris-HCl, pH 7.5, 50% glycerol (v/v), flash-frozen, and stored at −80 °C.

All proteins were >99% pure, as judged by SDSPolyacrylamide gels stained with Coomassie Blue. Protein concentrations were determined by measuring A280 under denaturing conditions (6 M guanidine hydrochloride). Extinctions coefficients used for concentration determination were 1.80 (mg/ml) −1 cm−1 for the CYT-18 NTDs, 1.70 (mg/ml) −1 cm−1 for the An NTDs, and 1.79 (mg/ml) −1 cm−1 for the Cp NTDs. Concentrations of CYT-18, An NTDs, Cp NTDs, and Af mtTyrRS refer to the homodimer.

Tyrosyl Adenylation Assays—Tyrosyl adenylation reactions were done with 100 nM protein in 50 µl of reaction medium containing 5 mM ATP, 100 mM KCl, 10 mM MgCl2, 144 mM Tris-HCl, pH 7.5, 2 mM DTT, 0.1 mg/ml BSA (New England Biolabs), 0.1 unit of yeast inorganic phosphatase (New England Biolabs), and 5 µCi of [l-3,5-3H]tyrosine (53 Ci mmol−1; Amersham Biosciences). The reactions were initiated by adding protein and incubated at 30 °C for 10 min. Reactions were terminated by diluting with 1 ml of reaction buffer and then filtering through a nitrocellulose membrane (Protran BA83, GE Healthcare), which retains protein-bound tyrosyl adenylate. Radioactivity was measured using a Beckman Coulter LS 6500 scintillation counter by dissolving the protein-bound membrane in Ready Protein scintillation mixture (Beckman).

DNA Splicing Assays—[32P]-Labeled intron-containing precursor RNAs were transcribed from linearized pAfmtLSU-ΔORF using a MEGAscript kit (Ambion) with 3.3 mM [α−32P]UTP (3000 Ci mmol−1; PerkinElmer Life Sciences). Splicing assays were carried out in 20 µl of reaction media containing different concentrations of KCl and MgCl2 in 20 mM Tris-HCl, pH 7.4, with 100 nM [32P]-labeled RNA, 200 nM mtTyrRS protein dimer, and 1 mM GTP-Mg2+. The reactions were initiated by adding GTP-Mg2+ and incubated at 30 °C. 4-µl portions were collected at various time points (0, 2, 10, and 60 min), and the reactions were terminated by adding stop buffer (75 mM EDTA, 0.1% SDS). The products were analyzed in denaturing 6% polyacrylamide gels, which were dried and quantified using phosphorimaging.

Protein Crystallization and Data Collection—Initial crystallization screens were done at concentrations of 16.5 mg/ml An NTDs and 6.7 mg/ml Cp NTDs using crystallization screens from Hampton Research and a Phoenix protein crystallization robot (Art Robbins Instruments). Crystals of the An NTDs and Cp NTDs were obtained in crystal screen L1e (Hampton...
Splicing-active Mitochondrial Tyrosyl-tRNA Synthetases

FIGURE 1. Domain architecture of mtTyrRSs and tyrosyl adenylation activity of mtTyrRS constructs. A, domain structure of the mtTyrRSs of N. crassa (CYT-18 protein), An, Cp, S. cerevisiae (Sc), and H. sapiens (Hs) compared with bacterial TyrRSs from G. stearothermophilus (Gs) and E. coli (Ec). Regions conserved between mt and bacterial TyrRSs are black, the mitochondrial targeting sequence (MT) is yellow, the Pezizomycotina-specific insertions are cyan, and the variable length C-terminal extension is purple. The insertions at the position of Ins 1 in the S. cerevisiae and H. sapiens mtTyrRSs (gray) are not homologous to Ins 1 of the Pezizomycotina mtTyrRSs (4). The An and Cp mtTyrRS constructs that were used for crystallization contain the N-terminal catalytic and α-helical domains (NTDs) and lack the CTD and C-tail. B, tyrosyl adenylation assays. The assays show that the CYT-18 and An and Cp NTDs constructs synthesize enzyme-bound tyrosyl adenylate (the first step in aminocacylation) with a stoichiometry close to 1 mol/mol or protein dimer. The bar graph shows the mean for three experiments with the error bars indicating the S.D.

Results and Discussion

Characterization of Crystallization Constructs—To investigate the divergence of the group I intron RNA binding surface in the catalytic domain of different Pezizomycotina mtTyrRSs, we determined the structures of C-terminally truncated An and Cp mtTyrRSs by x-ray crystallography and compared them to previously determined N. crassa CYT-18 structures. The An and Cp mtTyrRS constructs used for crystallography, denoted An and Cp NTDs, contain the N-terminal catalytic and α-helical domains but lack the mt targeting sequence and flexibly attached CTD and C-tail, which impede crystallization of full-length TyrRSs (Fig. 1A) (26, 52, 53). The truncated proteins activated i-tyrosine using ATP to form enzyme-bound tyrosyl adenylate at a stoichiometry of near 1 molecule per protein dimer, as expected for active NTD proteins with half-sites reac-
Splicing-active Mitochondrial Tyrosyl-tRNA Synthetases

Overview of the Pezizomycotina mtTyrRS Structures—As expected, the structures show that the An and Cp NTDs are homodimERIC proteins with overall domain architecture similar to those of bacterial TyrRSs (Fig. 2). The catalytic domain of both proteins displays a characteristic nucleotide binding-fold motif (Rossman fold), consisting of six parallel β-strands (βA-βF) connected by α-helices (H1–H11) with an active-site pocket that binds L-tyrosine, ATP, and the 3′ end of tRNA Tyr. The α-helical domain, which binds the variable and anticodon arms of tRNA Tyr (22), consists of five α-helices (H1′–H5′). As in bacterial TyrRSs, the dimer interface is formed by interactions between the two catalytic domains, leaving two α-helical domains at opposite ends of the dimer (Fig. 2, A and C) (22, 55).

The catalytic domain of class I aminocyl-tRNA synthetases contains two conserved motifs, HIGH and KMSKS, which function in amino acid activation by ATP (20, 56, 57). The HIGH motifs (HVGH and HIGH in the An and Cp mtTyrRSs, respectively) are visible in both structures at the end of H3 (orange; Fig. 2, B and D). The KMSKS motif (KFJKS and KIGKS in the An NTDs and Cp NTDs, respectively) is disordered in the structures, as in other TyrRS structures that lack bound tyrosyl adenylate (22, 41, 53).

Superposition of the An and Cp NTDs and previously determined CYT-18 NTDs structures onto the bacterial G. stearothermophilus TyrRS structure (Fig. 3A) shows that all have structurally similar NTDs, except for the N-terminal extension and two insertions (Ins 1 and 2), which are present only in the Pezizomycotina mtTyrRSs. Excluding these regions, the r.m.s.d. for the fits of the backbone of the CYT-18, An, and Cp NTDs to the bacterial G. stearothermophilus NTDs structure are 1.08, 1.23, and 1.22 Å, respectively. In both the An and Cp mtTyrRSs, the N-terminal extension contains an α-helix denoted H0; Ins 1 is an expansion of H5 followed by a loop, and Ins 2 is an expansion of the loop between helix 10 and βF (Figs. 2 and 3). Superposition of the three mtTyrRSs with each other shows that the structures of the H0 extension and Ins 2 differ significantly in the three fungal enzymes (Fig. 3B), as described in more detail below.

Binding of Tyrosine by the Pezizomycotina mtTyrRSs—All three of the Pezizomycotina mtTyrRSs were crystallized with L-tyrosine in the active site, where it is bound similarly via hydrogen bonds to five residues. Four of these residues (Tyr-99, Tyr-247, Gln-251, and Asp-254 in CYT-18) are conserved in all TyrRSs, with the remaining residue (Asp-143 in CYT-18) is specific to mt and bacterial TyrRSs (Fig. 3C) (53). Although the Cp NTDs crystallized with one monomer in the asymmetric unit, the An NTDs had two monomers of neighboring dimers in the asymmetric unit. Interestingly, one of these monomers contains a bound L-tyrosine, whereas the other monomer does not. The two monomers are highly similar with an overall (all atom) r.m.s.d. of 0.817 Å. However, two regions are visible only in the tyrosine-bound monomer. One is a 15-residue loop connecting βC to H5, which is likely visible due to the conserved stabilizing interaction between the L-tyrosine ligand and Asp-113 (Fig. 3C). The other corresponds to five residues of H9 and the loop connecting H9 to H10, which is near the previously mentioned stabilized 15-residue loop. Similar ordering of regions near the active site upon binding of L-tyrosine has been observed for other TyrRSs (53, 58).

Structural Variations in the Pezizomycotina-specific Insertions—The CYT-18 + Twort intron RNA co-crystal structure showed that the N-terminal H0 extension, Ins 1, and Ins 2 all function directly in binding group I introns (12). In CYT-18 as well as the An and Cp mtTyrRSs, the N-terminal extension consists of a disordered region that is not visible in the crystal structure followed by the α-helix H0, which is flexibly attached to the proteins by a conserved glycine (Gly-62 in CYT-18) and fixed in position by interactions with the protein core (Fig. 4). Although the orientation of H0 is similar in all three proteins, the lengths of the unstructured region and H0 helix differ. In CYT-18, the H0 helix is 23 residues long and is preceded by a 6-residue unstructured region (Figs. 4A and 5), whereas the A. nidulans mtTyrRS structure shows a smaller H0 helix (17 or 18 residues in the two monomers) preceded by 12 or 13 residues that are not visible in the structure (Figs. 4B and 5). The C. posadasii protein has a substantially longer 27-residue helix with only the two N-terminal residues (Gln-48 and Lys-49) not visible in the structure (Figs. 4C and 5).

In all three proteins (CYT-18 and the An and Cp mtTyrRSs), two conserved interactions help stabilize H0 against the body of the protein. One is a stacking interaction between a highly conserved tryptophan in H0 (Trp-52, Trp-81, and Trp-67, respectively) and a conserved arginine in the loop between H2 and βB
(Arg-95, Arg-123, and Arg-110, respectively), and the other is a hydrophobic interaction between a highly conserved isoleucine at the C-terminal end of H0 (Ile-59, Ile-88, and Ile-74) and residues in H2 (Fig. 4) (4). However, the CYT-18 structure shows additional H0 interactions that are not seen in the other proteins. These include (i) a second stacking interaction between the conserved H0 tryptophan (Trp-52) with another arginine in H0 (Arg-55), (ii) a salt bridge between the H0-H1 loop (Lys-64) and H4 of the protein core (Glu-126) (Fig. 4A), and (iii) interactions between H0 and Ins 1 and 2 that are described in more detail in the next section. These differences suggest that the orientation of H0 may be more flexible in the An and Cp mtTyrRSs than in CYT-18.

In all three Pezizomycotina mtTyrRSs, Ins 1 is a 13-residue expansion of H5 followed by a 6-residue loop, which packs against the protein core via salt bridges with the catalytic and α-helical domains (Figs. 4 and 5). The salt bridges between Ins 1 and the catalytic domain are similar in all three TyrRSs, whereas those between Ins 1 and the α-helical domain differ. In CYT-18, the Ins 1 helix interacts with H3’ of the α-helical domain (Fig. 4A), whereas the An and Cp mtTyrRS have a salt bridge between the C-terminal end of the Ins 1 helix and H4’ of the α-helical domain (Fig. 4, B and C).

Surprisingly, the structure of Ins 2 differs markedly between the three proteins. Ins 2 of CYT-18 is 18 residues long and consists of a loop containing a small 7-residue proline-capped helix (Figs. 4A and 5). It is composed of mostly acidic residues and is distinctively proline-rich. By contrast, the Ins 2s of both the An and Cp mtTyrRS lack regular secondary structure (Fig. 4, B and C). The Ins 2 of the Cp mtTyrRS is the same length as that of CYT-18 (18 residues) and is likewise composed largely of acidic and polar residues with a similarly high number of prolines, whereas Ins 2 of the An mtTyrRS is slightly shorter (16 residues) and composed of acidic and polar residues but is not proline-rich (Fig. 5). These differences in the structures and stabilizing interactions of the three insertions, which are most pronounced for H0 and Ins 2, could potentially affect group I intron binding (see below).

**Interactions between the Pezizomycotina-specific Insertions**—A striking feature of the CYT-18 structure is that all three of the catalytic domain insertions interact with each other, forming an extended scaffold for group I intron binding (59). Surprisingly, this fully connected scaffold is not observed in the other two mtTyrRSs. In CYT-18, the N-terminal extension interacts with Ins 1 via the stacking of Lys-44 in the H0 helix with Trp-205 in the Ins 1 loop, which may help stabilize the H0 helix (Fig. 4A). A similar stacking interaction is present in the Cp mtTyrRS between Arg-59 (H0) and Trp-205 (Ins 1) but is absent in the An mtTyrRS, where the corresponding tryptophan in Ins 1 has been replaced by a methionine (Met-207), whose side chain is disordered due to flexibility (Fig. 4, B and C).
In CYT-18 the N-terminal extension also interacts with Ins 2 via a potential salt bridge between a conserved arginine (Arg-35) in the unstructured region upstream of H0 and a glutamic acid (Glu-296) in Ins 2 (59). This interaction may be stabilized by group I intron binding and appears to contribute to splicing activity (25, 59). In the An mtTyrRS, a similar salt bridge is possible as the arginine (Arg-64) in the N-terminal extension and an acidic residue (Asp-332) in Ins 2 are present (Fig. 5). By contrast, in the Cp mtTyrRS, the conserved N-terminal arginine residue (Arg-50) caps the H0 helix, whereas the corresponding residue in Ins 2 is a tryptophan (Trp-309) rather than an acidic residue. Furthermore, the longer H0 helix of Cp mtTyrRS places Arg-50 too far to form a salt bridge with an acidic residue in Ins 2 without a conformational change upon group I intron binding (Fig. 4C). These missing interactions between the three insertions together with the missing interactions between the insertions and protein core described in the preceding section suggest that the group I intron binding sites in both the An and Cp mtTyrRSs may be more flexible than in CYT-18, possibly to accommodate the splicing of different introns.

**Potential Interactions of Cp and An mtTyrRS with a Group I Intron RNA**—The insertions in the catalytic domain of Pezizomycotina mtTyrRSs together with neighboring regions of the protein core form an extended group I intron binding surface on the side of the catalytic domain opposite the TyrRS active site (7, 12, 13, 26). The CYT-18 NTDs + Twort co-crystal structure shows that the protein interacts with both the P4-P6 and P3-P9 domains of the intron RNA, with most of the interactions occurring between the catalytic domain of one protein subunit (denoted subunit B) and the P4-P6 domain, particularly near its junction with the P3-P9 domain (12).

To examine how variations in the insertions might affect RNA binding, we modeled the An NTDs and Cp NTDs onto the CYT-18 NTDs (magenta; An NTDs (purple), and Cp NTDs (green) with the G. stearothermophilus mtTyrRS NTDs (yellow; PDB ID 2TS1 (41)). The superposition highlights that the insertions found in Pezizomycotina mtTyrRSs are absent in the bacterial enzyme. B, superposition of the An NTDs (purple), Cp NTDs (green), and CYT-18 NTDs (magenta), highlighting differences in the Pezizomycotina-specific insertions. The core regions of the TyrRSs are gray, and the insertions are shown in the color for the corresponding mtTyrRS in panel A. C, tyrosine binding sites of the Pezizomycotina mtTyrRSs. L-Tyrosine forms hydrogen bonds with five residues of which four are conserved in all TyrRSs, and one (Asp-143 in CYT-18 and Asp-113 and Asp-158 in the An and Cp mtTyrRSs, respectively) is specific to mtTyrRSs.

**FIGURE 3. Superposition of An NTDs and Cp NTDs with other TyrRSs and structure of the tyrosine binding sites.** A, superposition of the CYT-18 NTDs (magenta), An NTDs (purple), and Cp NTDs (green) with the G. stearothermophilus mtTyrRS NTDs (yellow; PDB ID 2TS1 (41)). The superposition highlights that the insertions found in Pezizomycotina mtTyrRSs are absent in the bacterial enzyme. B, superposition of the An NTDs (purple), Cp NTDs (green), and CYT-18 NTDs (magenta), highlighting differences in the Pezizomycotina-specific insertions. The core regions of the TyrRSs are gray, and the insertions are shown in the color for the corresponding mtTyrRS in panel A. C, tyrosine binding sites of the Pezizomycotina mtTyrRSs. L-Tyrosine forms hydrogen bonds with five residues of which four are conserved in all TyrRSs, and one (Asp-143 in CYT-18 and Asp-113 and Asp-158 in the An and Cp mtTyrRSs, respectively) is specific to mtTyrRSs.
In the CYT-18/H11001 Twort co-crystal structure, Ins 1 of subunit B (Ins 1-B) interacts with the major groove of the P6-P6a region of the intron, with both the helical and loop regions of Ins 1 involved in the interaction (Fig. 6D). These potential interactions of Ins 1-B with P6/P6a are generally similar for all three enzymes, but some residues in Ins 1 that appear to interact with the intron RNA differ between the three proteins (e.g. Glu-206 of the An mtTyrRS and Glu-193 and Trp-203 of the Cp mtTyrRS are potential contact sites in these proteins but not CYT-18) (Fig. 6, E and F).

Finally, Ins 2 of CYT-18 subunit B (Ins 2-B) binds in the P9 minor groove via its proline-capped helix and potentially stabilizes the important long range L9-P5 tertiary interaction that helps orient the P4-P6 and P3-P9 RNA domains of the intron core (Fig. 6D) (12, 60). Ins 2-B of the An mtTyrRS lacks this small helix but is still positioned in the P9 minor groove and may function similarly to Ins 2-B in CYT-18 (Fig. 6E). By contrast, the Cp mtTyrRS Ins 2-B is much farther away from the intron RNA (Fig. 6F, 8 Å compared with 2.7 Å in the CYT-18 NTDs + Twort co-crystal structure), possibly reflecting the missing stabilizing interaction between Ins 2 and H0 (see above). Thus, it is possible that Ins 2-B of the Cp mtTyrRS does not interact with intron RNA. Alternatively, the Cp mtTyrRS may undergo a conformational change upon RNA binding that enables Ins 2-B to interact with the intron RNA as in CYT-18. Such an induced fit scenario for the binding of the Cp Ins 2-B would differ from CYT-18, where the position and structure of Ins 2-B is the same in the presence or absence of bound intron RNA (12). We conclude that whereas Ins 1 may function similarly in group I intron binding in all three mtTyrRSs, both the H0 extension and Ins 2 may interact differently or more dynamically with group I intron RNAs.
Potential Intron RNA Contact Sites with Insertions Are Less Conserved Than Those with the mtTyrRS Core—The CYT-18 NTDs + Twort co-crystal structure identified 23 protein residues as close enough (within 4 Å) to potentially interact with the intron RNA (residues marked by asterisks in Fig. 5). Of the 23 potentially interacting residues, 13 are within the insertions (4, 5, and 4 residues in H0-B, Ins 1-B, and Ins 2-B, respectively), and 10 are in the protein core (Figs. 5 and 7).

Previous sequence alignments showed that none of the 13 putative group I intron binding residues in the insertions are strongly conserved in a group of 18 Pezizomycotina mtTyrRSs (4), and none of these residues is conserved, and only four are similar between CYT-18 and the An and Cp mtTyrRSs (the four similar residues are Tyr-41, Lys-44, Lys-172, and Gln-292 in CYT-18; Figs. 5 and 7). By contrast, 4 of the 10 putative interacting residues in the protein core (His-167, Lys-171, Arg-194, and Lys-205 in CYT-18) are conserved in both the An and Cp mtTyrRSs (residues names in purple in Fig. 7), and another three (Lys-193, Ile-196, and Gln-261 in CYT-18) are similar between CYT-18 and the An and Cp mtTyrRSs (residues names in pink in Fig. 7). Notably, three of the four conserved core residues (His-167, Lys-171, and Lys-205 in CYT-18) and two of the three similar core residues (Lys-193 and Ile-196 in CYT-18) potentially interact with different regions of the intron from both subunits of the protein.

Of the 10 potentially interacting residues that lie in the protein core, seven (His-167, Lys-171, Lys-193, Arg-194, Gly-195, Ile-196, and Gln-262 of subunit B of CYT-18) are clustered near Ins 1 and interact with P4 and P6 of the intron RNA, likely bolstering the interactions of Ins 1-B with the P4-P6 domain.
Four of these residues (His-167, Lys-171, Lys-193, and Ile-196 in CYT-18) may also interact with a different region of the intron RNA (the P5/L9 loop) from subunit A (Fig. 7, right hand panels). Five of the seven core residues near Ins 1 are either conserved (His-167, Lys-171, and Arg-194) or similar (Lys-193 and Ile-196) in other Pezizomycotina mtTyrRSs with the histidine being conserved in all 18 fungal mtTyrRSs compared previously (4).

The remaining three potentially interacting core residues are in the middle of the intron binding surface (His-201 and Gln-261 in CYT-18) or at the dimerization interface (Lys-205 in CYT-18) and are in position to interact with the P4 region of the intron (Fig. 7A). Although Lys-205 and Gln-261 are conserved or similar between the three proteins and likely interact with the intron RNA in a similar manner as in CYT-18, His-201 is dissimilar and replaced by a valine in the An mtTyrRS (Val-229) and a threonine in the Cp mtTyrRS (Thr-216). Although the threonine could interact with intron RNA, the valine in the An NTDs is not in a position to interact with the intron RNA (Fig. 5).

Collectively, these findings indicate that protein core residues that potentially interact with the intron RNA and particularly those near Ins 1 are more conserved than potentially interacting residues found within the insertions. The stronger conservation of potentially interacting core residues may reflect that they are under greater constraint to preserve TyrRS activity than are those within the insertions.

**Use of the Pezizomycotina Structures to Construct a Homology Model of a mtTyrRS from a Pathogenic Fungus**—Our structures of the intron binding surfaces of splicing-active fungal
mtTyrRSs can be used for homology modeling of mtTyrRSs from other pathogenic Pezizomycotina fungi. Such information would enable the rational design of drugs that target the mtTyrRS-intron interaction without perturbing aminoacylation. As an example, we built a homology model of the mtTyrRS from *A. fumigatus*, which is the most common causative agent of aspergillosis (61). The An and Af mtTyrRSs are closely related with a sequence identity of 80.5% and sequence similarity of 87.5% (Fig. 8A). Splicing assays showed that splicing of the Af mtLSU intron is protein-dependent and could be spliced either by the Af mtTyrRS or by CYT-18, which is capable of splicing a diverse range of group I introns from different species (Fig. 8, B and C; Ref. 4). These data support the possibility of targeting the Af mtTyrRS-intron interaction for therapeutic intervention.

An I-TASSER homology model of the Af mtTyrRS showed high structural similarity to the protein core of An NTDs, CYT-18, and the Cp NTDs (Fig. 8D). Ins 1 and Ins 2 are similar in structure to the *A. nidulans* insertions, whereas the N-terminal H0 extension was modeled by the program with a helix length similar to that of CYT-18 (even if the CYT-18 structure was excluded as a template). Like the An mtTyrRS, the Af mtTyrRS enzyme lacks the tryptophan in Ins 1 that stacks with a lysine in the H0 extension, and, therefore, does not have the same interconnected insertion scaffold found in CYT-18. A salt bridge between Ins 2 and the H0 extension may still be possible, however, as the interacting resi-

FIGURE 7. Conservation and divergence of potential group I intron RNA binding residues in the CYT-18 NTDs + Twort co-crystal structure in other Pezizomycotina mtTyrRSs. A–C, CYT-18 NTDs + Twort RNA structure (12) and An and Cp NTDs + Twort RNA models, respectively. The full structure or models are shown to the left, and close-up views are shown to the right. The proteins are shown in space-filling representation with HO, Ins 1, and Ins 2 in cyan and labeled A or B to denote to which protein subunit they belong. The P4-P6 and P3-P9 domains of the Twort intron RNA are colored black and orange, respectively. The dashed line shows the possible location of a portion of the L5 loop that was not visible in the CYT-18 NTDs + Twort structure. Amino acid residues located within 4 Å of the intron RNA in the structure or models are labeled on the protein surface. Potentially interacting residues in the Pezizomycotina-specific insertions are indicated in cyan, and potentially interacting residues in the protein core are indicated in different colors depending upon their degree of conservation (purple, identical in all three Pezizomycotina mtTyrRSs; pink, similar in all three Pezizomycotina mtTyrRSs; green, not conserved in all three Pezizomycotina mtTyrRSs (4)).
Splicing-active Mitochondrial Tyrosyl-tRNA Synthetases

FIGURE 8. A. fumigatus mtTyrRS sequence comparison and biochemical assays. A, sequence alignment of the A. fumigatus and A. nidulans mtTyrRS. The alignment was created by using MUSCLE (72). Residues conserved between the two Aspergillus spp. mtTyrRSs are shown in black boxes with white letters, whereas similar residues defined according to the Henikoff matrix (73) are shown in gray boxes. Residues that potentially interact with intron RNA (located within 4 Å in the CYT-18 NTDs + Twort RNA co-crystal structure and An or Af NTDs + Twort RNA models) are indicated by black asterisks above the alignments if the potential interaction is observed in all three, whereas residues that potentially interact with the intron RNA in only one or two of this group are indicated by magenta (CYT-18), purple (An NTDs), or green (Af NTDs) asterisks. B, splicing assays. 32P-Labeled Af mtLSU (100 nM) was incubated without protein or with the Af mtTyrRS or CYT-18 (200 nM) and with 1 mM GTP-Mg2+ for 60 min at 30 °C in buffer containing 100 mM KCl, 5 mM MgCl2, 20 mM Tris-HCl, pH 7.4. C, splicing time courses. Reactions were done as in panel B in reaction media containing different concentrations of KCl and MgCl2 as indicated in the figure. The plots show the disappearance of precursor RNA as a function of time. End-point splicing reactions (60 min) of the Af mtLSU intron by CYT-18 in the same reaction medium are shown as single time points. D, homology model of A. fumigatus mtTyrRS dimers and monomers built by using I-TASSER (51). E, model of the Af NTDs + Twort intron RNA built by aligning the Af NTDs model onto the CYT-18 NTDs in the co-crystal structure. Insertions are shown in cyan, and the P4-P6 and P3-P9 domains of the Twort intron RNA are colored black and orange, respectively. The remainder of the intron is colored gray. E1-E2, ligated exons; I, excised intron; H0, intron + 3’ exon; P, precursor RNA.

FIGURE 8. A. fumigatus mtTyrRS sequence comparison and biochemical assays. A, sequence alignment of the A. fumigatus and A. nidulans mtTyrRS. The alignment was created by using MUSCLE (72). Residues conserved between the two Aspergillus spp. mtTyrRSs are shown in black boxes with white letters, whereas similar residues defined according to the Henikoff matrix (73) are shown in gray boxes. Residues that potentially interact with intron RNA (located within 4 Å in the CYT-18 NTDs + Twort RNA co-crystal structure and An or Af NTDs + Twort RNA models) are indicated by black asterisks above the alignments if the potential interaction is observed in all three, whereas residues that potentially interact with the intron RNA in only one or two of this group are indicated by magenta (CYT-18), purple (An NTDs), or green (Af NTDs) asterisks. B, splicing assays. 32P-Labeled Af mtLSU (100 nM) was incubated without protein or with the Af mtTyrRS or CYT-18 (200 nM) and with 1 mM GTP-Mg2+ for 60 min at 30 °C in buffer containing 100 mM KCl, 5 mM MgCl2, 20 mM Tris-HCl, pH 7.4. C, splicing time courses. Reactions were done as in panel B in reaction media containing different concentrations of KCl and MgCl2 as indicated in the figure. The plots show the disappearance of precursor RNA as a function of time. End-point splicing reactions (60 min) of the Af mtLSU intron by CYT-18 in the same reaction medium are shown as single time points. D, homology model of A. fumigatus mtTyrRS dimers and monomers built by using I-TASSER (51). E, model of the Af NTDs + Twort intron RNA built by aligning the Af NTDs model onto the CYT-18 NTDs in the co-crystal structure. Insertions are shown in cyan, and the P4-P6 and P3-P9 domains of the Twort intron RNA are colored black and orange, respectively. The remainder of the intron is colored gray. E1-E2, ligated exons; I, excised intron; H0, intron + 3’ exon; P, precursor RNA.

Splicing-active Mitochondrial Tyrosyl-tRNA Synthetases

FIGURE 8. A. fumigatus mtTyrRS sequence comparison and biochemical assays. A, sequence alignment of the A. fumigatus and A. nidulans mtTyrRS. The alignment was created by using MUSCLE (72). Residues conserved between the two Aspergillus spp. mtTyrRSs are shown in black boxes with white letters, whereas similar residues defined according to the Henikoff matrix (73) are shown in gray boxes. Residues that potentially interact with intron RNA (located within 4 Å in the CYT-18 NTDs + Twort RNA co-crystal structure and An or Af NTDs + Twort RNA models) are indicated by black asterisks above the alignments if the potential interaction is observed in all three, whereas residues that potentially interact with the intron RNA in only one or two of this group are indicated by magenta (CYT-18), purple (An NTDs), or green (Af NTDs) asterisks. B, splicing assays. 32P-Labeled Af mtLSU (100 nM) was incubated without protein or with the Af mtTyrRS or CYT-18 (200 nM) and with 1 mM GTP-Mg2+ for 60 min at 30 °C in buffer containing 100 mM KCl, 5 mM MgCl2, 20 mM Tris-HCl, pH 7.4. C, splicing time courses. Reactions were done as in panel B in reaction media containing different concentrations of KCl and MgCl2 as indicated in the figure. The plots show the disappearance of precursor RNA as a function of time. End-point splicing reactions (60 min) of the Af mtLSU intron by CYT-18 in the same reaction medium are shown as single time points. D, homology model of A. fumigatus mtTyrRS dimers and monomers built by using I-TASSER (51). E, model of the Af NTDs + Twort intron RNA built by aligning the Af NTDs model onto the CYT-18 NTDs in the co-crystal structure. Insertions are shown in cyan, and the P4-P6 and P3-P9 domains of the Twort intron RNA are colored black and orange, respectively. The remainder of the intron is colored gray. E1-E2, ligated exons; I, excised intron; H0, intron + 3’ exon; P, precursor RNA.
3 of the 12 interacting residues in the insertions are conserved. These findings again highlight the rapid evolutionary divergence of intron RNA interactions with the less constrained insertion regions (Fig. 8A).

**Comparison with Non-splicing TyrRSs**—Next, we compared the group I intron binding surface of the catalytic domain of the Pezizomycotina mtTyrRSs with the same surface in x-ray crystal structures or models of homologous non-splicing TyrRS, including the bacterial *G. stearothermophilus* and *E. coli* TyrRSs, the human mtTyrRS, and an I-TASSER model of *S. cerevisiae* mtTyrRS. The monomer structures are shown in schematic representation with the Pezizomycotina-insertions in cyan. The yeast and human mtTyrRSs have sequence expansions (gray) in the region of Ins 1 that are structurally similar to but lack residues conserved in Ins 1 of the Pezizomycotina mtTyrRSs. Also, sequence alignment of the H0 extension and Ins 1 region of the four Pezizomycotina mtTyrRSs (*N. crassa*, *A. nidulans*, *A. fumigatus*, and *C. posadasii*) with the same regions of the non-splicing mtTyrRSs. The CYT-18 N-terminal extension, which contains the H0 helix, and Ins 1 are demarcated by cyan bars under the alignment. Residues that were identified as potentially interacting with the Twort intron RNA (located within 4 Å in the CYT-18 NTDs + Twort co-crystal structure; Ref. 12) are indicated with an asterisk. Residues that are identical in three or more of the mtTyrRSs are shown in black boxes with white letters, and residues that are similar in three or more of the mtTyrRSs are shown in gray boxes.

As in the Pezizomycotina mtTyrRSs, Ins 1 in the non-splicing yeast and human mtTyrRSs consists of an expansion of H5 followed by a loop (Fig. 9A). Although the yeast mtTyrRS Ins 1 is similar in size and number of charged residues to Ins 1 of the Pezizomycotina enzymes, the human mtTyrRS Ins 1 is smaller and has fewer charged residues, suggesting less potential for nucleic acid binding (Fig. 9B). Both the yeast and human mtTyrRS Ins 1s lack putative intron-interacting residues found in the Pezizomycotina Ins 1s (Fig. 9B).

Collectively, the most parsimonious interpretation of these observations is that Ins 1 and a short α-helical N-terminal extension predated the divergence of Pezizomycotina and Saccharomycotina fungi and the evolution of splicing activity.
whereas Ins 2 was acquired later, enabling additional contacts with the intron RNA. The N-terminal α-helical extension was presumably lost in the lineage leading to the human mtTyrRS.

**Electrostatic Potential of the Group I Intron Binding Surface in Splicing and Non-Splicing mtTyrRSs**—The evolution of splicing activity can also be traced by comparing the electrostatic potential of the intron binding side of the mtTyrRSs. All of the splicing-active Pezizomycotina mtTyrRSs have a basic, positively charged (blue) surface, consistent with the ability of this surface to bind nucleic acids (Fig. 10A). By contrast, a similar basic surface was not observed in the bacterial TyrRS or the closely related but non-splicing yeast mtTyrRS and was weaker and less extensive in the non-splicing *Homo sapiens* mtTyrRS. Notably, the basic intron binding surface of the Pezizomycotina enzymes is composed not only of the insertions (outlined in Fig. 10A) but also neighboring regions of the protein core, which appear to have evolved toward more positively charged residues as an adaptation for RNA splicing. None of the 10 intron-interacting residues found in the CYT-18 protein core are conserved in the human mtTyrRS or in the Saccharomycotina subphylum (4). Notably, the basic surface in the Cp mtTyrRS extended farther toward the distal ends of NTDs than in CYT-18 or the An or Af mtTyrRSs (Fig. 10B). This longer basic tract may reflect co-evolution of the Cp NTDs with the Cp mtLSU intron, which is spliced by this protein (4) and whose P5 element is larger and extends farther toward the distal regions of the NTDs than in the An and *N. crassa* mtLSU introns (29 nucleotides compared with 16 and 21 nucleotides, respectively) (Fig. 10B).

**Potential Drug Target Sites in Fungal mtTyrRSs**—The novel splicing activity of Pezizomycotina mtTyrRSs, which is required for mitochondrial function, is a potential drug target for therapeutic intervention in fungal infections. The finding
that the group I intron binding site is distinct from the TyrRS active site makes it potentially targetable without inhibiting the essential mtTyrRS activity of the protein. Previous studies showed that the two activities are separable by mutation (24, 26). The most distinctive features of the Pezizomycotina mtTyrRS required for splicing activity are the Pezizomycotina-specific insertions, with the structures of H0 and Ins 1 being more conserved between the different fungal enzymes than that of Ins 2.

As our results suggest variability in the manner in which H0 binds group I introns (Fig. 6), the most suitable drug target in this region may be the highly conserved packing interactions between the H0 helix and the protein core, which helps position the N-terminal extension for RNA splicing. The cyt-18-I mutation G127E is in a highly conserved glycine located away from the protein surface in a loop wedged between H0 and Ins 1 and likely inhibits splicing by destabilizing the RNA binding scaffold created by the Pezizomycotina-specific insertions (26). Interactions mediated by α-helical structural elements are viewed as particularly tractable for targeting by small molecule inhibitors; peptide-backbone mimics that mimic the α-helix; or non-peptide molecules that mimic the side chains of an α-helical element but are not helical (62, 63).

Our findings also identify Ins 1 and neighboring regions of the protein core as particularly promising for drug targeting. The interactions of the protein with the intron RNA in these regions are the most extensive and conserved in different Pezizomycotina mtTyrRSs. Although the human mtTyrRS also has an Ins 1, it is shorter and has fewer basic residues than Ins 1 of the Pezizomycotina mtTyrRSs (Figs. 9 and 10). These differences may make it easier to target this region of the Pezizomycotina enzymes without affecting human mtTyrRS activity. The C-terminal domain insertions, which were not studied here, also provide additional potential drug targets.

**Evolution of Splicing Activity of Fungal mtTyrRSs**—Our comparison of Pezizomycotina mtTyrRSs with each other and with non-splicing mtTyrRSs provides further insight into the evolution of splicing activity. We previously suggested a scenario of pre-adaptive evolution (also referred to as constructive neutral evolution) in which the promiscuous, nonspecific RNA binding properties of the mtTyrRS CTD played a key role in the evolution of splicing activity. According to this scenario, an initial nonspecific interaction between the CTD of the mtTyrRS and a self-splicing group I intron became fixed by a mutation in the intron RNA that made it dependent upon the protein for splicing (27). Once fixed, further adaptive changes in the protein and intron RNA led to specific binding and more efficient protein-dependent splicing. The CTDs of bacterial and mitochondrial TyrRSs have a fold similar to ribosomal protein S4, a promiscuous RNA binding domain that has high nonspecific RNA binding activity and has been incorporated into a variety of proteins that bind structurally different RNAs (64–67).

The finding that interactions between Ins 1 and neighboring core regions with the P4-P6 domain of the intron RNA are similar between Pezizomycotina enzymes indicates that these specific interactions evolved early in the evolution of splicing activity before the divergence of Pezizomycotina mtTyrRSs. Ins 1, an expansion of H5 into the H5-β4 loop, was likely already present in the ancestral mtTyrRs. It could have contributed to the initial pre-adaptive interaction with the intron RNA and would have been more amenable than core regions to adaptive changes necessary for specific binding of group I introns without affecting mtTyrRS activity. The large number of intron interactions in and around Ins 1 is consistent with a ratchet-like mechanism in which new interactions with the intron RNA enable further RNA structural mutations in turn leading to increased binding and greater dependence upon the protein for RNA splicing.

A similar scenario can be imagined for H0, which may have evolved initially from the distal part of the N-terminal mt targeting sequence, as it is also an amphipathic helix (26, 68, 69). However, Ins 2 is an expansion of a smaller loop connecting H10 to βF that is found only in the Pezizomycotina mtTyrRSs and was likely acquired after the evolution of splicing activity to further stabilize the P5-L9 interaction between the P4-P6 and P3-P9 domains. This multistep evolution of the extended group I intron binding scaffold, which precisely fits the conserved phosphodiester backbone structure of the group I intron catalytic core, enabled the protein to interact with and promote the splicing of diverse group I introns with different structural mutations that impair self-splicing.

After the evolution of an extended group I intron binding scaffold, the mtTyrRSs in different Pezizomycotina species continued to co-evolve with cognate group I introns, leading to variations in their ability to splice different group I introns (4). Our results indicate that this further differentiation of the protein occurred more readily in unstructured terminal or loop regions, which may be less constrained to preserve TyrRS activity than are RNA binding regions in the protein core. More generally, our results highlight the role of unstructured terminal and loop regions in the evolution of new protein functions. Such disordered regions have higher rates of mutation to non-conserved amino acid residues than ordered regions and tend to evolve at faster rates, allowing for enzyme modifications that lead to new functions (70, 71).

**Author Contributions**—L. T. L., P. J. P., and A. M. L. designed the study, and L. T. L. and A. M. L. wrote the paper with help and editing from P. J. P. L. T. L. performed and analyzed the tyrosyl adenylation assays, built homology models of the Af NTDs and S. cerevisiae NTDs, crystallized the An and Cp NTDs, and solved their structures. M. S. designed and performed the Af mtTyrRS splicing experiments. All authors reviewed the results and approved the final version of the manuscript.

**Acknowledgments**—Instrumentation for x-ray crystallography and technical assistance for this work were provided by the Macromolecular Crystallography Facility of the Institute for Cellular and Molecular Biology at the University of Texas at Austin. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the United States Department of Energy under Contract DE-AC02-05CH11231.

**References**

1. Wolf, Y. I., Aravind, L., Grishin, N. V., and Koonin, E. V. (1999) Evolution of aminoacyl-tRNA synthetases: analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene trans-
Splicing-active Mitochondrial Tyrosyl-tRNA Synthetases

1. Guo, M., Yang, X.-L., and Schimmel, P. (2010) New functions of aminoacyl-tRNA synthetases beyond translation. Nat. Rev. Mol. Cell Biol. 11, 668–674.

2. Cherniack, A. D., Garriga, G., Kittle, J. D., Jr., Akins, R. A., and Lambowitz, A. M. (1987) A protein required for splicing of group I intron RNA by quantitative binding assays and CYT-18 protein binding site at the junction of stacked helices in a group I catalytic intron RNA. Mol. Biol. 23, 223–42.

3. Saldanha, R. J., Patel, S. S., Surendran, R., Lee, J. C., and Lambowitz, A. M. (1995) Involvement of Neurospora mitochondrial tyrosyl-tRNA synthetase in RNA splicing: a new method for purifying the protein and characterization of physical and enzymic properties pertinent to splicing. Biochemistry 34, 1275–1287.

4. Paukstelis, P. J., Coon, R., Madabushi, L., Nowakowski, J., Monzingo, A., Robertus, J., and Lambowitz, A. M. (2005) A tyrosyl-tRNA synthetase adapted to function in group I intron splicing by acquiring a new RNA binding surface. Mol. Cell 17, 417–428.

5. Lamech, L. T., Mallam, A. L., and Lambowitz, A. M. (2014) Evolution of RNA-protein interactions: non-specific binding led to RNA splicing activity of fungal mitochondrial tyrosyl-tRNA synthetases. PLoS Biol. 12, e1002028.

6. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol. 267, 307–326.

7. Matthews, B. W. (1968) Solvent content of protein crystals. J. Mol. Biol. 33, 491–497.

8. Stein, N. (2008) CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. J. Appl. Crystallogr. 41, 641–643.

9. Wittke, T., and Blow, D. M. (1989) Structure of tyrosyl-tRNA synthetase refined at 2.3 Å resolution. Interaction of the enzyme with the tyrosyl adenylate intermediate. J. Mol. Biol. 208, 83–98.

10. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–667.
Splicing-active Mitochondrial Tyrosyl-tRNA Synthetases

Crystallogr. 40, 658–674
43. Bricogne, G., Vonhein, C., Flensburg, C., Schiltz, M., and Paciorek, W. (2003) Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. Acta Crystallogr. D. Biol. Crystallogr. 59, 2023–2030
44. Vonhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007) Automated structure solution with autoSHARP. Macromolecular Crystallography Protocols: Volume 2: Structure Determination, pp. 215–230, Humana Press, Totowa, NJ
45. Cowtan, K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. D. Biol. Crystallogr. 62, 1002–1011
46. Cowtan, K. (2010) Recent developments in classical density modification. Acta Crystallogr. D. Biol. Crystallogr. 66, 470–478
47. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D. Biol. Crystallogr. 66, 213–221
48. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix. refine. Acta Crystallogr. D. Biol. Crystallogr. 68, 352–367
49. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D. Biol. Crystallogr. 53, 240–255
50. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D. Biol. Crystallogr. 66, 12–21
51. Roy, A., Kucukural, A., and Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 5, 725–738
52. Qiuh, X., Janson, C. A., Smith, W. W., Green, S. M., McDevitt, P., Johanson, K., Carter, P.,hibbs, M., Lewis, C., Chalker, A., Fosberry, A., Lalonde, J., Berge, J., Brown, P., Houge-Frydrych, C. S., and Jarvest, R. L. (2001) Crystal structure of Staphylococcus aureus tyrosyl-tRNA synthetase in complex with a class of potent and specific inhibitors. Protein Sci. 10, 2008–2016
53. Kobayashi, T., Takimura, T., Sekine, R., Kelly, V. P., Vincent, K., Kamata, T., Takimura, T., Sekine, R., Kelly, V. P., Vincent, K., Sakamoto, K., Nishimura, S., and Yokoyama, S. (2005) Structural snapshots of the KMSKS loop rearrangement for amino acid activation by bacterial tyrosyl-tRNA synthetase. J. Mol. Biol. 346, 105–117
54. Jakes, R., and Fersht, A. R. (1975) Tyrosyl-tRNA synthetase from Escherichia coli: stoichiometry of ligand binding and half-of-the-sites reactivity in aminoclonation. Biochemistry 14, 3344–3350
55. Brick, P., and Blow, D. M. (1987) Crystal structure of a deletion mutant of a tyrosyl-tRNA synthetase complexed with tyrosine. J. Mol. Biol. 194, 287–297
56. First, E. A., and Fersht, A. R. (1995) Analysis of the role of the KMSKS loop in the catalytic mechanism of the tyrosyl-tRNA synthetase using mutant cycles. Biochemistry 34, 5030–5043
57. Xin, Y., Li, W., Dwyer, D. S., and First, E. A. (2000) Correlating amino acid conservation with function in tyrosyl-tRNA synthetase. J. Mol. Biol. 303, 287–298
58. Bonnefon, L., Frugier, M., Giege, R., and Rudinger-Thirion, J. (2005) Human mitochondrial TyrRS disobeys the tyrosine identity rules. RNA 11, 558–562
59. Geng, C., and Paukstelis, P. J. (2014) An in vitro peptide complementation assay for CYT-18-dependent group I intron splicing reveals a new role for the N terminus. Biochemistry 53, 1311–1319
60. Chen, X., Mohr, G., and Lambowitz, A. M. (2004) The Neurospora crassa CYT-18 protein C-terminal RNA-binding domain helps stabilize interdomain tertiary interactions in group I introns. RNA 10, 634–644
61. Verweij, P. E., and Brandt, M. E. (2006) Aspergillus, Fusarium, and other opportunistic moniliaceous fungi. Manual of Clinical Microbiology, 9th Ed., pp. 1802–1838, American Society for Microbiology, Washington, D. C.
62. Edwards, T. A., and Wilson, A. J. (2011) Helix-mediated protein-protein interactions as targets for intervention using foldamers. Amino Acids 41, 743–754
63. Dewal, M. B., and Firestone, S. M. (2011) Non-peptidic α-helical mimetics as protein-protein interaction inhibitors. Curr. Med. Chem. 18, 2420–2428
64. Aravind, L., and Koonin, E. V. (1999) Novel predicted RNA-binding domains associated with the translation machinery. J. Mol. Evol. 48, 291–302
65. Korber, P., Zander, T., Herschlag, D., and Bardwell, J. C. (1999) A new heat shock protein that binds nucleic acids. J. Biol. Chem. 274, 249–256
66. Staker, B. L., Korber, P., Bardwell, J. C., and Saper, M. A. (2000) Structure of Hsp15 reveals a novel RNA-binding motif. EMBO J. 19, 749–757
67. Gerstner, R. B., Pak, Y., and Draper, D. E. (2001) Recognition of 16S rRNA by ribosomal protein S4 from Bacillus stearothermophilus. Biochemistry 40, 7165–7173
68. von Heijne, G. (1986) Mitochondrial targeting sequences may form amphiphilic helices. EMBO J. 5, 1335–1342
69. von Heijne, G., Steppuhn, J., and Herrmann, R. G. (1989) Domain structure of mitochondrial and chloroplast targeting peptides. Eur. J. Biochem. 180, 535–545
70. Brown, C. J., Johnson, A. K., Dunker, A. K., and Daughdrill, G. W. (2011) Evolution and disorder. Curr. Opin. Struct. Biol. 21, 441–446
71. Wilke, C. O., and Drummond, D. A. (2010) Signatures of protein biophysics in coding sequence evolution. Curr. Opin. Struct. Biol. 20, 385–389
72. Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797
73. Henikoff, S., and Henikoff, J. G. (1992) Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. U.S.A. 89, 10915–10919
74. Bonnefon, L., Frugier, M., Touzé, E., Lorber, B., Florentz, C., Giege, R., Sauter, C., and Rudinger-Thirion, J. (2007) Crystal structure of human mitochondrial tyrosyl-tRNA synthetase reveals common and idiosyncratic features. Structure 15, 1505–1516
75. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. U.S.A. 98, 10037–10041
76. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004) PDBePQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatic calculations. Nucleic Acids Res. 32, W665–W667