Influence of specific mutations on the thermal stability of the td group I intron in vitro and on its splicing efficiency in vivo: A comparative study

PHILIPPE BRION,1,2 RENÉE SCHROEDER,2 FRANÇOIS MICHEL,3 and ERIC WESTHOF1
1 Institut de Biologie Moléculaire et Cellulaire du CNRS, UPR 9002, F-67084 Strasbourg, France
2 Institute of Microbiology and Genetics, Vienna Biocenter, Dr. Bohrgasse 9, A-1030 Vienna, Austria
3 Centre de Génétique Moléculaire du CNRS, F-91190 Gif-sur-Yvette, France

ABSTRACT
Group I introns constitute excellent systems for analyzing the relationship between RNA tertiary folding and catalysis. Within a hierarchical framework interpretation of RNA folding, secondary structure motifs subdivide RNA threedimensional (3D) architecture. Thus, mutations in two-dimensional motifs are expected to have effects different from those disrupting 3D contacts. Using UV spectroscopy, we have studied the influence of nucleotide substitutions, in both secondary and tertiary structure elements, on the thermal stability of the tertiary folding of the bacteriophage T4 td group I intron. Further, we present a quantitative analysis of the relationship between the splicing efficiency in vivo and the stability of the intron structure as monitored by UV melting curves. We conclude that the stability of the tertiary structure of a group I intron as measured by UV melting is generally a good indication of its ability to splice in vivo.

Keywords: cooperative folding; group I intron; secondary structure; tertiary structure; UV melting

INTRODUCTION
Group I self-splicing introns constitute particularly attractive model systems for studying the interrelationship between sequence, folding, structure, stability, and activity of an RNA molecule. Group I introns are natural RNA catalysts that carry out self-splicing via two consecutive transesterification reactions, leading to their own excision from a pre-messenger RNA and the ligation of the flanking exon sequences, generally in the absence of proteins but in the presence of guanosine as a cofactor (reviewed by Cech, 1990). These catalytic activities are associated with a relatively small conserved core (about 100 nt), which is organized into two domains, P4-P5-P6 and P3-P7-P8, that form the active site and bind the substrates (Michel & Westhof, 1990); Pn designates double-stranded helices in the well-established group I secondary structure. This core is surrounded in turn and stabilized by additional structural components, the number and nature of which differ in the various group I intron subgroups. The structure of a molecule composed of domain P4-P5-P6 and the optional subdomain that stabilizes it in the Tetrahymena pre-ribosomal intron has been determined at atomic resolution (Cate et al., 1996). Many additional tertiary interactions have been identified by comparative sequence analysis and biochemical experiments within the core of group I introns (Lehnert et al., 1996) and additional detail has recently become available with the publication of a medium-resolution crystallographic structure of the entire core of the Tetrahymena intron (Golden et al., 1998). A complete understanding of group I intron structure would require not only the determination of the thermodynamically stable and native structure adopted, but also a knowledge of the kinetic folding processes. Folding pathways and intermediates are beginning to be identified for both the wild-type version and variants of the Tetrahymena intron, but we are not yet able to predict the consequences of any particular base substitution on the final fraction of correctly folded molecules.
and their stability. We are still further away from understanding the requirements for activity in vivo (e.g., Coetzee et al., 1994; Myers et al., 1996).

One question of general interest with regard to RNA folding which could be tackled by using group I introns is whether the disruption of tertiary interactions has distinct consequences from the destabilization of secondary structure helices. As originally observed on tRNA (Cole et al., 1972; Crothers et al., 1974), and now confirmed by a number of studies (Banerjee et al., 1993; Jaeger et al., 1993; Laing & Draper, 1994; Costa et al., 1998), there is a clear thermodynamic basis for distinguishing RNA secondary and tertiary structures. Three-dimensional structure tends to melt early and cooperatively, whereas a majority of secondary structure helices constitute separate melting units, the stability of which is generally higher than that of the overall tertiary structure. However, some double-stranded helices are also stabilized, at least partly, by the three-dimensional fold of the molecule (Banerjee et al., 1993; Jaeger et al., 1993) with their exact number a function of experimental conditions (Draper, 1996). Kinetically as well, calculations and experiments indicate that secondary structures should, and do, tend to form earlier, and also that a majority of them are likely to survive into the final structure (or at least so in molecules that self-assemble in vitro). Hence, the current consensus views RNA folding primarily as a hierarchical process. But does the fact that secondary structure normally predates tertiary structure mean that base substitutions that destabilize double-stranded helices will tend to have less severe effects on the fraction of correctly folded molecules than those that impair tertiary interactions? And which type of mutations should have the greatest effects on the thermal stability of those molecules that managed to fold correctly?

Group I introns capable of forming homogeneous populations of correctly folded molecules in salt solutions can clearly be utilized to try to answer these questions. A further advantage of some group I introns is that they can easily be introduced back into their original in vivo settings, so that the effects of specific mutations on in vitro folding and stability can directly be compared with the ability of the same mutations to impair splicing in vivo. In the present study, we chose to introduce a number of base substitutions in secondary structures and known tertiary interactions of the td intron of bacteriophage T4. We then determined (1) the in vitro thermal stability, as estimated by UV melting curves, of substituted molecules that retain a three-dimensional structure and (2) the extent to which splicing of the same intron variants is affected in vivo (for comparison purposes, melting curves were also determined for some of the splicing-deficient mutants selected by Belfort et al. (1987)). Our main conclusion is that the stability of the tertiary fold of a group I intron, as measured by UV melting experiments, provides a generally good indication of its ability to splice in vivo, although the reverse proposition is not necessarily true.

RESULTS

Phenotype of td intron mutants

We used a plating assay (Belfort et al., 1987) to determine the phenotype at 37 °C of the mutants we had chosen to investigate. Splicing efficiency was assayed by the ability of a plasmid-borne td gene to provide thymidylate synthase (TS) to Escherichia coli cells lacking this enzyme because of the disruption of the endogenous thy A gene (Galloway-Salvo et al., 1990). E. coli C600 thyA::Km' cells were transformed with td plasmids containing the mutant intron. Transformants were replated on minimal medium (MM), MM with thymine (MMT) and MM with trimethoprim and thymine (TTM). Cells containing introns that cannot splice fail to produce TS and thus cannot grow on MM, but are able to grow on TTM. Inefficient splicing results in the synthesis of small amounts of TS and is reflected by growth on both MM and TTM, whereas cells that are splicing efficiently can grow on MM, but are unable to grow on TTM (Amyes & Smith, 1974).

The location and nature of substitutions we introduced in the td intron, in the td ∆P6-2 context, are shown in Figure 1, as well as the mutants recovered by Belfort et al. (1987) in the td ∆P6-1 context. Figure 2 shows a representative plating assay for the different td constructs and summarizes the results of plating assays for all the td intron mutants. As expected, the Thy− control, harboring a pTZ18U plasmid lacking the td gene, grew on trimethoprim-containing TTM medium, in contrast to the Thy+ control containing the unmutagenized td+ plasmid that was sensitive to trimethoprim. On the other hand, the Thy+ control grew on medium lacking thymine, whereas the Thy− control did not. Two other td mutants with mutations in the P7 helix of the catalytic core, namely C873U (Belfort et al., 1987) and C870U (Schroeder et al., 1991), were used as additional splicing-deficient controls. These mutants also exhibit the expected phenotype. Although C873U did not grow on thymine-less media, C870U grows slightly in the absence of thymine; both constructs grew on TTM medium. The G944A mutant, which also harbors a mutation in P7, exhibits a Thy− phenotype, whereas mutants U912G (in helix P3) and G948A (in an isolated tertiary Watson–Crick pair called P9.0a), which grew on both MM and TTM media, have to be considered as leaky Thy− mutants, producing enough TS to grow in the absence of thymine, but too little to become sensitive to trimethoprim. As illustrated by the fact that they can grow on MM but are unable to grow on TTM, all the other mutants have a Thy+ phenotype.
Ribozyme stability and splicing efficiency

FIGURE 1. Secondary structure model of the td group I intron of bacteriophage T4 (Shub et al., 1988) (redrawn after Cech et al., 1994). Heavy arrows point to splice sites, exon sequences are in lower case. Tertiary interactions are indicated by broken lines. The numbering of the intron sequence is according to Belfort et al. (1987). The specific mutations analyzed in this study are in cyan boldface letters, the mutations recovered by Belfort et al. (1987) are in red boldface letters, and the C873U mutation, analyzed in both studies, is in violet. The bases homologous to part of the nucleotides studied by Michel et al. (1992) and Jaeger et al. (1993) in the sun Y intron are circled in green. The truncated ribozyme form of the intron used for melting curves begins with an additional G (shaded) after the // symbol (Costa & Michel, 1995). Bases marked with * denote the position of the open reading frame deleted in the construct used (Galloway-Salvo et al., 1990). The junction between P7.1 and P7.2 follows the proposal of Leonitis & Westhof (1998).

**Primer extension assay of in vivo splicing**

To estimate the efficiency of splicing, we carried out primer extension assays using total cellular RNA isolated from bacteria grown at 37 °C and containing wild-type or mutant td introns. A 5’ end-labeled primer complementary to a sequence immediately downstream of the 3’ splice site was used to synthesize short cDNAs in the presence of three dNTPs and high levels of ddTTP to terminate the cDNA opposite to the first A-residue in the template (Mohr et al., 1992; Semrad & Schroeder, 1998). Figure 3A shows a representative primer extension assay. Extension of the pre-mRNA results in a product that is 5 nt longer than the primer, whereas the mature mRNA results in a product 16 nt longer. In addition to the mature splicing product, a cryptic (crp) splice site located 29 nt upstream of the 5’ splice site is also utilized, resulting in an extension product 8 nt longer than the primer (Chandry & Belfort, 1987). The bands 12 and 13 nt longer than the primer are unidentified products that might result from a premature stop of reverse transcription on the mature template (Semrad & Schroeder, 1998). As the total amount of counts detected for the various mutants in Figure 3A is different, although a constant amount of total RNA was used, we conclude that the amount of td RNA in some mutants is lower than in the wild-type. It is therefore possible that those mutants have either lower transcription rates or increased RNA degradation rates, which could have an influence on the td phenotype in vivo.

Figure 3B shows the results of the primer extension assay performed. Splicing activities were calculated using the ratio mRNA/mRNA + crp-mRNA + pre-mRNA. Splicing activities of the Thy’ mutants C873U and G944A are severely reduced, yielding less than 1.5% of the wild-type splicing activity. On the other hand, the td constructs with Thy’ phenotypes accumulated between 38% and 84% of the amount of mature mRNA
Inefficient splicing of the mutant td introns might be due to folding defects. If this is the case, do the folding defects revealed by inefficient splicing detectably affect the thermal stability of the td intron molecule? To answer this question, the UV absorbance of wild-type and mutant versions of the td intron was followed as a function of temperature. To prevent splicing during the experiment, we chose to work with a truncated ribozyme form of td that is lacking the P1–P2 substrate (Costa & Michel, 1995; Fig. 1).

When wild-type transcripts are heated in a buffer containing 5 mM MgCl₂, an early unfolding transition that gives rise to a major peak at 56 °C in the first-derivative absorbance versus temperature profile is observed (Fig. 4A). Similar early occurring and highly cooperative transitions were previously observed for other group I introns (Banerjee et al., 1993; Jaeger et al., 1993) and reflect the disruption of the three-dimensional structure of the molecule (for a review, see Brion & Westhof, 1997).

All but one of the mutants having a Thy⁺ phenotype show a melting profile comparable to the melting profile of the wild-type molecule, with an early and rather distinct cooperative transition. The exception is the G949A mutant in helix P6a, which shows instead a twin-peak transition (see Figs. 4A,C). In contrast, the two mutants with a Thy⁻ phenotype, C873U and G944A (both with a mutation in P7), show no clear peak that could be attributed to the melting of the three-dimensional structure (see Fig. 4A). These mutant molecules are therefore assumed either to have no stable three-dimensional fold at all or to be unable to form a conformationally homogeneous population after renaturation. Among the mutants having an intermediate phenotype, G948A (in the P9.0a base pair) shows a discrete three-dimensional melting peak downshifted to 49 °C (Fig. 4B), whereas the U912C mutant in P3 exhibits an additional early transition at 42 °C (Fig. 4A).

The absorbance versus temperature profiles of all mutants with a wild-type-like melting profile differ nevertheless from that of the wild-type in that the leading edge of the first peak is invariably shifted to a lower temperature (see Figs. 4A,B,C; Table 1). Interestingly, a connection exists between this observation and the decreased splicing efficiency observed in Figure 3. When the fraction of in vivo spliced molecules is compared with the value of ΔTm, as measured on melting profiles, it becomes apparent that the decrease in activity correlates in general with the extent of thermal destabilization; notably, the mutant with the highest ΔTm (G948A) shows only 36% splicing, whereas the two mutants with a ΔTm of 1.0 °C or less have splicing efficiencies above 80%. However, the molecule carrying the C870U mutation in the internal bulge of the P7 helix stands apart from other mutant molecules in this respect: this mutation appears to have very little effect on the thermal stability of the intron (ΔTm = −0.4 °C), whereas its in vivo splicing efficiency is severely affected.

A number of splicing-deficient mutations of the td ΔP6-1 form of intron td were selected in vivo by Belfort et al. (1987) and some of these mutations were further characterized for their ability to be rescued in vitro by the product of the CYT-18 gene of Neurospora crassa, a protein that contributes to the stabilization of the native fold of a number of group I introns (Mohr et al., 1992). The in vivo phenotypes and mRNA levels of some of these mutants are listed in Table 2, together with the extent of rescue by CYT-18 (Mohr et al., 1992). The Tm values corresponding to the initial melting tran-

---

**FIGURE 3.** Primer extension assay of td RNA synthesized in vivo. A: Reverse transcription of RNA isolated from transformed cells after 30 min at 42 °C. Extension products were separated by electrophoresis. The primer is designed such that extension on the mRNA template produces a band 16 nt longer than the primer, extension of the cryptic (crp)-mRNA produces a band 8 nt longer, and extension of the unspliced pre-mRNA a band 5 nt longer. B: Extension products were quantified with a PhosphorImager (Molecular Dynamics). The percentage of splicing is calculated as the percentage of mRNA relative to the total td RNA (mRNA + crp-mRNA + pre-mRNA). Splicing efficiencies are given relative to wild-type activity. The results presented are mean values from at least three independent experiments.

accumulated by the wild-type. Finally, those with intermediate phenotypes gave ratios ranging between 5% of the wild-type splicing activity for C870U to 30% and 36% for U912C and G948A, respectively. Altogether, these results show that there is good agreement between the plating assay and splicing activity using the primer extension assay.

**Probing individual interactions by UV-melting curves**

Inefficient splicing of the mutant td introns might be due to folding defects. If this is the case, do the folding defects revealed by inefficient splicing detectably affect the thermal stability of the td intron molecule? To answer this question, the UV absorbance of wild-type and mutant versions of the td intron was followed as a function of temperature. To prevent splicing during the experiment, we chose to work with a truncated ribozyme form of td that is lacking the P1–P2 substrate (Costa & Michel, 1995; Fig. 1).

When wild-type transcripts are heated in a buffer containing 5 mM MgCl₂, an early unfolding transition that
Ribozyme stability and splicing efficiency

Mutations affecting known tertiary interactions are again seen to result in a strong in vivo phenotype and low mRNA level, whereas mutations not known to affect tertiary interactions have milder phenotypes and higher mRNA levels. However, the melting profiles of these mutants (Figs. 4D,E,F,G) are more complex than those of the engineered intron as 122 nt

**FIGURE 4.** First derivative profile of absorbance versus temperature curves for various *td* group I intron transcripts. Derivative absorbance melting profiles were determined in 50 mM Na cacodylate (pH 7.5 at 25°C), 50 mM NH₄Cl, and 5 mM MgCl₂. For A, B, and C, the transcripts used are a truncated form of the *td* intron, lacking the ORF in L6a (Galloway-Salvo et al., 1990), and beginning with an additional G (see Fig. 1 and Materials and Methods; Costa & Michel, 1995). For D, E, F, and G, the transcripts used are a truncated form of the *td* Δ 1-3 intron (Belfort et al., 1987), with part of the natural ORF being present in L6a, and beginning with an additional G (see Fig. 1 and Materials and Methods; Costa & Michel, 1995). A: Melting profiles of the wild-type and various mutants affecting the P7 and P3 pairing. B: Melting profiles of *td* transcripts carrying mutations known to affect tertiary interactions. C: Melting profiles of *td* transcripts carrying mutations not known to affect tertiary interactions. D: Melting profiles of the wild-type and various mutants affecting the P7 and P3 pairing. E,F: Melting profiles of *td* transcripts carrying mutations known to affect tertiary interactions. G: Melting profiles of *td* transcripts carrying mutations not known to affect tertiary interactions.
of the open reading frame (ORF) naturally present in the td intron gene are present in loop L6a at the tip of P6a. First, the peak corresponding to secondary structure, which is present in all profiles, appears downshifted to 64°C. Second, the peak corresponding to the three-dimensional structure is broad and appears to include two transitions, one around 56°C, assigned to the melting of the three-dimensional structure of the intron, and one around 58°C. Indeed the deconvolution of the melting profiles of the various mutants displays two additional transitions, centered around 58°C and 65°C, compared to what is observed in the td ΔP6-2 context. These transitions, because they do not exist in the td ΔP6-2 context, must be due to structures present in the 122 nt of the ORF or to long-range interactions between those nucleotides and the intron core (Figs. 5A,B). A computer-predicted two-dimensional model of this region shows two very stable secondary structures (Fig. 5C), one of them already noticed by Chu et al. (1986).

Three mutations—C873U, G898A, and G891A—essentially abolish the initial peak. In those three mutants, the fraction of mRNA is low (<11%) and the rescue by CYT-18 ≤ 47%. In three other mutants, the peak centered around 52–54°C is conspicuous: G932A, C866U, and C1000U. The first two of those mutations are known to affect tertiary interactions and present low levels of mRNA, but a good rescue by CYT-18 (~72%). The third mutation, not known to be involved in tertiary contacts, presents excellent growth with high % weight and % CYT-18 rescue. A third set of mutants presents a broad shoulder of low intensity around 52–53°C: G67A, G51A, and G968A, with C79U having no clear change. The G67A and C79U mutants have severe phenotypes and are rescued by CYT-18 only up to 49%. Both of the mutated nucleotides are involved in important tertiary contacts, whereas the third mutant, G968A, has an intermediate phenotype and could only indirectly be involved in tertiary contacts. The low CYT-18 rescue of the C79U mutant could be due to the fact that the mutated residue belongs to the recognition site of CYT-18 on P4–P6 (Caprara et al., 1996). The mutation G51A also presents an intermediate phenotype, but with an excellent rescue by CYT-18. Although

### TABLE 1. In vivo and in vitro phenotypes of the various mutants tested, arranged in each subgroup by increasing biological activity.

<table>
<thead>
<tr>
<th>T4 intron td</th>
<th>Phenotype</th>
<th>% wt&lt;sup&gt;a&lt;/sup&gt;</th>
<th>T&lt;sub&gt;m&lt;/sub&gt; (± 0.2) (°C)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>ΔT&lt;sub&gt;m&lt;/sub&gt; (°C)&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type</td>
<td>+++</td>
<td>100</td>
<td>55.6</td>
<td>—</td>
</tr>
<tr>
<td>Mutation in the substrate</td>
<td>A19U (substrate)</td>
<td>+ + +</td>
<td>70</td>
<td>nd&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Mutations in P7 or P3</td>
<td>C873U P7 (3D)</td>
<td>—</td>
<td>&lt;1.5</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>G944A P7 (3D)</td>
<td>—</td>
<td>&lt;1.5</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>C870U P7 (3D)</td>
<td>+ +</td>
<td>5</td>
<td>55.2</td>
</tr>
<tr>
<td></td>
<td>U912C P3 (2D)</td>
<td>+ +</td>
<td>30</td>
<td>41.6/55.6&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>Mutations known to affect tertiary interactions</td>
<td>G948A P9.0a (3D)</td>
<td>+ +</td>
<td>36</td>
<td>49.1</td>
</tr>
<tr>
<td></td>
<td>U899C L7.2–L9.2 (3D)</td>
<td>+ + +</td>
<td>40</td>
<td>51.2</td>
</tr>
<tr>
<td></td>
<td>U920C L2–P8 (2D/3D)</td>
<td>+ + +</td>
<td>80</td>
<td>54.8</td>
</tr>
<tr>
<td></td>
<td>G960A L9–P5 (3D)</td>
<td>+ + +</td>
<td>84</td>
<td>54.6</td>
</tr>
<tr>
<td>Mutations not known to affect tertiary interactions</td>
<td>G885A P7.1 (2D)</td>
<td>+ + +</td>
<td>38</td>
<td>51.2</td>
</tr>
<tr>
<td></td>
<td>G972A P8.1 (2D)</td>
<td>+ + +</td>
<td>44</td>
<td>52.0</td>
</tr>
<tr>
<td></td>
<td>G953A P9 (2D)</td>
<td>+ + +</td>
<td>45</td>
<td>53.2</td>
</tr>
<tr>
<td></td>
<td>G89A P6a (2D)</td>
<td>+ + +</td>
<td>69</td>
<td>51.4&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup>Growth phenotypes at 37°C are indicated as follows. +++: growth on MM equivalent to MMT; no growth in the presence of trimethoprim (TTM). + + : growth on MM equivalent to MMT; incomplete growth inhibition on TTM. −: no growth on MM; no growth inhibition on TTM. Growth phenotypes are defined as in Mohr et al. (1992).

<sup>b</sup>Percentage of mRNA relative to the total td RNA (mRNA + crp-mRNA + pre-mRNA). The percentage of splicing is given relative to wild-type activity.

<sup>c</sup>Melting temperature of the three-dimensional structure measured at 5 mM MgCl<sub>2</sub>.

<sup>d</sup>Substrate not present in the truncated ribozyme form of the intron used for the melting experiments.

<sup>e</sup>No discernable absorbance change corresponding to the melting of the three-dimensional structure.

<sup>f</sup>See Figure 4A and text.

<sup>g</sup>This RNA partially dimerizes (Brion et al., 1999) and the melting temperature of the dimer is 51.2°C (see Fig. 4C).
covariations involving bp 4 of P4 were noticed (Michel & Westhof, 1990), recent reports indicate that it is bp 5 of P4 that is involved in tertiary contacts with J8/7 (Tanner et al., 1997).

DISCUSSION

Despite the fact that the group I intron family as a whole is well known for possessing extremely few conserved nucleotides (Cech, 1988; Michel & Westhof, 1990; Damberger & Gutell, 1994; Lisacek et al., 1994), the sequence of any particular self-splicing group I intron is nevertheless highly constrained by natural selection. Related group I introns tend to have very similar secondary structures, even in peripheral domains, and, except at sites where mismatched base pairs or bulges are under clear positive selection, imperfect helices are rare. In any particular subgroup of closely related introns, variations tend to be confined, in fact, to a few specific positions in single-stranded segments, or else to consist of occasional insertions and deletions in the most peripheral parts of the molecule. This picture is reinforced by selection experiments that show that even in vitro, core segments are essentially intolerant of base substitutions, unless strong selection pressures are applied for many cycles (Beaudry & Joyce, 1992; Lehman & Joyce, 1993).

Base substitutions can interfere with group I intron function in a variety of ways. Mutations in or in the vicinity of the active site could wreck catalysis. Those that affect directly or indirectly substrate recognition sites could prevent substrate binding. Some single-base substitutions in tertiary interactions or at ion-binding sites, and possibly even in secondary structure helices, could affect the stability of the final correct fold of the molecule to such an extent that it could no longer exist under in vivo conditions, or else become less viable than alternate incorrect structures. Finally, a number of substitutions, by interfering with folding pathways, could generate molecules that remain trapped in meta-

### TABLE 2

Phenotypic data gathered from the literature on the td intron (Mohr et al., 1992) and the sunY intron (Michel et al., 1992; Jaeger et al., 1993).

<table>
<thead>
<tr>
<th>T4 intron td</th>
<th>Phenytype*</th>
<th>% wt</th>
<th>Trn (°C)</th>
<th>CYT-18 rescue (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type</td>
<td>++</td>
<td>100</td>
<td>−56</td>
<td>100</td>
</tr>
<tr>
<td>Mutation in the substrate</td>
<td>G-3A P1 (substitute)</td>
<td>−&lt;1.7</td>
<td>nd*</td>
<td>&lt;1.8</td>
</tr>
<tr>
<td>Mutations in P7</td>
<td>C873U P7 (3D)</td>
<td>−&lt;1.7</td>
<td>−</td>
<td>47.2</td>
</tr>
<tr>
<td>Mutations known to affect tertiary interactions</td>
<td>G67A P5–L9 (3D)</td>
<td>−&lt;1.7</td>
<td>−52</td>
<td>45.1</td>
</tr>
<tr>
<td></td>
<td>G989A L7.2–L9.2 (3D)</td>
<td>+&lt;1.7</td>
<td>−46</td>
<td>32.8</td>
</tr>
<tr>
<td></td>
<td>G932A P8–L2 (3D)</td>
<td>+&lt;1.7</td>
<td>53.8</td>
<td>71.9</td>
</tr>
<tr>
<td></td>
<td>C79U P6 (2D/3D)</td>
<td>+5.4</td>
<td>−</td>
<td>48.8</td>
</tr>
<tr>
<td></td>
<td>C866U J6/7 (3D)</td>
<td>++</td>
<td>7.1</td>
<td>53.6</td>
</tr>
<tr>
<td></td>
<td>G891A P7.2 (2D/3D)</td>
<td>++</td>
<td>10.8</td>
<td>49</td>
</tr>
<tr>
<td>Mutations not known to affect tertiary interactions</td>
<td>G968A P9 (2D)</td>
<td>++</td>
<td>12.2</td>
<td>55.4</td>
</tr>
<tr>
<td></td>
<td>G51A P4 (2D)</td>
<td>++</td>
<td>16.5</td>
<td>80.4</td>
</tr>
<tr>
<td></td>
<td>C1000U P9.2 (2D)</td>
<td>++</td>
<td>41.5</td>
<td>81.2</td>
</tr>
<tr>
<td>Data on the sunY intron</td>
<td>Wild-type</td>
<td>100</td>
<td>56.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G131C L7.2–L9.2 (3D)</td>
<td>8</td>
<td>44.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C1013G L7.2–L9.2 (3D)</td>
<td>6</td>
<td>44.6</td>
<td></td>
</tr>
</tbody>
</table>

*Growth phenotypes are indicated as follows: +++: growth on MM equivalent to MMT; no growth in the presence of trimethoprim (TTM). ++: growth on MM equivalent to MMT; incomplete growth inhibition on TTM. +: slow growth on MM; no growth inhibition on TTM. +−: no growth on MM; no growth inhibition on TTM. Definition of growth phenotypes and data from Mohr et al. (1992).

*Percentage of mRNA relative to the total td RNA (mRNA + crp-mRNA + pre-mRNA). Data from Mohr et al. (1992).

*Melting temperature of the three-dimensional structure. The symbol − indicates a broad shoulder. Otherwise, there is a clearly identified peak. (This work).

*Data from Mohr et al. (1992).

*Substrate not present in the truncated ribozyme form of the intron used for the melting experiments.

*No discernable absorbance change corresponding to the melting of the three-dimensional structure.

*Data from Michel et al. (1992).

*Data from Jaeger et al. (1993).
stable conformations or fall prey to nucleases before they can ever reach their final correct structure.

In this study, we have used two sets of point mutants of the td intron to try and answer two possibly related questions: do mutations that impair known tertiary interactions have more severe effects than those that weaken secondary structure pairings (traditionally defined as composed of contiguous Watson–Crick or wobble base pairs between antiparallel strands), and is there a clear correlation between the extent of structural destabilization in vitro and the degree of splicing impairment in vivo?

**Mutations with a severe phenotype tend to be in tertiary interactions**

In this work, in vivo splicing efficiencies were directly estimated by primer extension and, indirectly so, by a plating assay that makes it possible to evaluate levels of the protein product of the td gene (Belfort et al., 1987). The same tests were previously used by Mohr et al. (1992) to characterize the splicing-deficient td mutants isolated by Belfort et al. (1987). A comparison between our data (Table 1) and data of Mohr et al. (1992) (Table 2) shows that, in both studies, mutations with clear Thy' phenotypes turned out to have splicing efficiencies lower than 1.7%.

Our own set of mutants was generated by deliberately introducing a series of single base substitutions in proven tertiary interactions, whereas another series affected sites located in secondary structure helices and not known to participate in tertiary structure. The four mutations of the latter type were found to have a wild-type growth phenotype and in vivo splicing efficiencies between 38% and 69% (80% if mutant U920C, which is located in helix P8 but could affect binding of the GUGA loop of the P1–P2 substrate by its receptor, is included). Three of the mutations in tertiary interactions also have wild-type-like phenotypes and splicing efficiencies in the same range. However, mutant G948A, which disrupts an isolated Watson–Crick base pair (P9.0a; Jaeger et al., 1993), is somewhat more severely affected and base substitutions in helices P3 and especially P7 have even more drastic effects. Helices P3 and P7 have in common that their individual strands are located relatively far from one another on the sequence of the intron; that, at least in the Tetrahymena group I intron, they form at a late stage in the folding process (Zarrinkar & Williamson, 1994); and that they cannot be simultaneously included in a planar tree-like representation of the group I intron secondary structure—that is, they form a pseudoknot. Formally, one of the two helices constituting a pseudoknot belongs to the secondary structure whereas the other belongs to the tertiary structure. In Table 1, P3 and P7 are regarded as belonging to secondary and tertiary structure, respectively, following evidence from melting

---

**FIGURE 5.** Example of deconvolution of a melting profile. The distributions are characterized by three parameters: \( Tm \), the melting temperature; \( \sigma \), the standard deviation; and \( S \), a correction factor for the surface. **A.** Deconvolution of the first derivative melting profile of the wild-type transcript in the td AP6-2 context. A two Gaussian distribution model is used. The distributions, of parameters \( (Tm, \sigma, S) = (55.6, 2.2, 0.2) \) and \( (71.5, 13.3, 1.1) \), are assigned to tertiary and secondary structures unfolding, respectively. **B.** Deconvolution of the first derivative melting profile of the wild-type transcript in the td AP6-1 context. A four Gaussian distribution model is used. The trends, of parameters \( (Tm, \sigma, S) = (55.3, 2.8, 0.2) \) and \( (64.6, 10.9, 0.9) \), are assigned to the unfolding of the tertiary and secondary structures of the intron, respectively. The other two, of parameters \( (Tm, \sigma, S) = (57.7, 2.1, 0.05) \) and \( (64.6, 2.3, 0.1) \), are assigned to additional structures present in the ORF. **C.** Secondary structure model of natural ORF fragment remaining in L6a of td AP6-1 transcripts. This model was established using the program SAPSSARN (Gaspin & Westhof, 1995). The base marked with * denotes the position of the sequence deleted in the construct used and the numbering of the sequence is according to Belfort et al. (1987).
studies of some of their mutants (Brion et al., 1999). Furthermore, helix P7 has a semiconserved base sequence that may contribute to its forming a non-A-form helix (Golden et al., 1998), thereby making it more susceptible to mutations. The U912C substitution in P3 has an intermediate phenotype, the C870U mutation at the P7 bulge position is more severe, with a splicing efficiency of 5%, whereas the other two P7 mutations, which both affect the 873–944 base pair, are even more deleterious.

In summary, only a fraction of the mutations we introduced in tertiary interactions turned out to have serious effects on splicing activity, but there is a marked excess of substitutions affecting tertiary structure among the most severely affected mutants. Preferential localization of mutations that prevent in vivo splicing at or close to sites involved in tertiary interactions was already observed for group I introns (Michel et al., 1992) and recently described for the Sc. cox 1/1 (a11) group II intron as well (Robineau et al., 1997). As evidenced by superimposing the mutagenesis data of Belfort et al. (1987) on the three-dimensional model of Michel & Westhof (1990), a majority of the sites at which splicing-deficient mutations of the td intron of bacteriophage T4 were recovered participate, in fact, in tertiary interactions (see also Fig. 1 and Table 2). The two receptor sequences, in stem P5 and P8, for the two GUGA tetraloops L9 and L2 of the td intron (Jaeger et al., 1994; Costa & Michel, 1995) constitute particularly striking examples: homologous G-to-A substitutions were recovered at these two sites and both mutations have splicing efficiencies below 1.7% when compared to the wild-type (Mohr et al., 1992; Table 2). However, why then in this context were mutations like C1000U, G891A, and G968A, which are located on double-stranded helices, also selected by Belfort et al. (1987)? In fact, the C1000U mutation has, according to Mohr et al. (1992), a clear Thy+ phenotype, and the other two mutants, although not fully sensitive to trimethoprim, grow well on minimal medium and have intermediate levels of mRNA (Table 2).

Correlation between structural destabilization and impairment of in vivo splicing

In the case of the sun Y group I intron, a close relative of the td intron, it was shown that a variety of base substitutions in tertiary interactions lowers the temperature at which the molecule becomes inactive (Jaeger et al., 1993, 1994), unless the mutation is compensated by raising the concentration of magnesium ions (Michel et al., 1992). One of these substitutions was also investigated by UV-temperature profiles and directly shown to decrease the temperature of the first unfolding transition (Jaeger et al., 1993), and similar effects were recently reported for mutations in a group II intron (Costa et al., 1998) and the thiostrepton-binding domain of ribosomal RNA (Conn et al., 1998). We therefore resorted to UV melting curves to characterize both our mutations and some of the ones studied by Mohr et al. (1992), some of which were reported to be suppressed by higher magnesium ions concentrations or spermidine.

All but two of our own base-substituted molecules retain the ability to form a tertiary structure, as evidenced by the fact that a well-individualized peak is present to the left of a broad domain corresponding to the melting of secondary structure elements (Fig. 4A,B,C). However, in all mutants, the temperature at which the transition from a fully folded to a partly unfolded structure takes place is decreased, with \( \Delta T_m \) values ranging from \(-0.4^\circ C \) to \(-10^\circ C \) compared to the wild-type (the U912C mutation does not result in a significant \( \Delta T_m \), but the presence of an additional peak at about \( 42^\circ C \) may indicate that only a fraction of molecules carrying it can fold correctly). Three of the mutations we examined are at sites where base substitutions had also been introduced in the sun Y intron. Replacement of the P9.0a G:C pair by an A:C mismatch (mutant G948A) results in a \( \Delta T_m \) of \(-6.5^\circ C \) in both molecules (Table 1; Jaeger et al., 1993). Disruption of the L7.2/3-L9.2/6 tertiary base pair has similar effects in the two introns: an A:C mismatch in the td intron shows a \( \Delta T_m \) of \(-10^\circ C \), whereas a C:C mismatch in the sun Y intron was reported to have a \( \Delta T_m \) of \(-12^\circ C \) (Table 2; Jaeger et al., 1993). In the L9:P5 contact as well, replacement of the G960:A66 pair by an A:A combination, which results in a \( \Delta T_m \) of \(-1.0^\circ C \) in the td intron (Table 1), had been shown to decrease the temperature at which half of a population of sun Y molecules are inactive by precisely the same value (Jaeger et al., 1994). On the other hand, not all sites within a particular tertiary contact are equally sensitive to substitutions. Thus, the U899C mutation in the L7.2–L9.2 pairing (Table 1) is much less damaging than the neighboring G898A mutation, even though both result in an A:C mismatch; mutations at position 898 might change the preferred conformation of the L7.2 loop to one unfit for interaction with the L9.2 loop.

Mutations in the P5 and P8 receptors of the L9 and L2 loops are also worth discussing. Both the U920C (Table 1) and G932A (Table 2) substitutions create A:C mismatches in helix P8, and have relatively small effects on melting profiles. However, molecules that carry the latter mutation are essentially unable to splice in vivo, presumably because they cannot bind their P1–P2 substrate (which is absent from the construct used for UV melting profiles). The homologous mutation in P5, G67A (Table 2) has more devastating effects, which are likely due to the fact that it does not merely affect directly the L9:P5 contact, but also disrupts the short P5 helix and, as a consequence, may well alter the structure of the J4/5 loop, which is also involved in substrate binding (Michel & Westhof, 1990; Strobel et al., 1998).
Because the initial unfolding transition corresponds to the disruption of tertiary structure, it might have been thought of as rather insensitive to the introduction of a single mismatch in one of the double-stranded helices of the secondary structure. However, since the enthalpy associated with an individual helix is usually much smaller than the enthalpy corresponding to the dissociation of the entire three-dimensional structure, even a single-base substitution may bring the melting temperature of any but the longest helices below the temperature range over which tertiary structure is lost, at which point the existence of this helix and the formation of tertiary structure may become interdependent. Thus, the fact that disruption of helix P7 (G885A), P9.1 (G972A), or P8 (U920C and G932A) lowers the temperature of the initial unfolding transition may simply mean that formation of the overall tertiary structure of the molecule requires correct closure of the internal loops at the base of each of these secondary structure components, even though they are not directly involved in tertiary interactions (or, for P8, at least so in the absence of a P1–P2 substrate). Alternatively, some of the mutations thought to affect only secondary structure could actually disrupt some yet unknown tertiary interactions. For instance, residue G891 may be part of a complex structure at the junction between P7.1 and P7.2 (Leontis & Westhof, 1998), which could explain the particularly strong effects of the G891A mutation (Table 2 and Fig. 1).

Whatever the cause for structural destabilization, how well do in vitro measured $\Delta Tm$ values correlate with the extent of splicing impairment in vivo? It is readily apparent from Tables 1 and 2 that structural stability is necessary for efficient in vivo splicing. All mutants with a Thy$^+$ phenotype and a relative mRNA level above 40% have in vitro melting profiles that reveal a markedly cooperative initial transition with a $\Delta Tm$ of at most $-5^\circ C$ compared to the wild-type. Moreover, mutations with mRNA levels above 80% all have a $\Delta Tm$ of $-1^\circ C$ or smaller. Conversely, mutants C873U and G944A, which show no evidence of tertiary structure in vitro, are unable to splice in vivo. However, structural stability is not sufficient by itself to ensure a good in vivo performance and this is best exemplified by the case of mutant C870U at the P7 bulge site. Consistent with chemical modification data, which had failed to reveal major structural defects in molecules carrying the mutation (Streicher et al., 1996), the C870U melting profile is almost identical to the wild-type one. In fact, position 870 is next to the catalytic center of group I introns, and can only be A or C in natural group I introns, so that the C870U mutation is a particularly good candidate for being a mutation that affects catalysis, even though an additional structural function of the bulge residue cannot be totally excluded because P1 and P2 are absent in the constructs used in the UV melting experiments (Schroeder et al., 1991).

In conclusion, a mutation can be severe because it destroys a tertiary long-range interaction, globally destabilizing the molecule—U899C and, probably, G948A—or interfering with the correct folding (U912C). But a mutation can also be severe because it affects the catalytic center (C870U) without significantly affecting stability or folding. Substituting one base in a helix may be essentially without effect on stability and folding or, at the other extreme, can abolish the global structure when that helix constitutes, in fact, a crucial tertiary interaction (e.g., the central P7 helix of the pseudoknot). Between these two extremes, the disruption of local secondary structure elements can be of unexpected importance by affecting directly or indirectly known or unknown tertiary interactions, ion binding sites, or a nucleotide crucial on the folding pathway. UV melting profiles, by making it possible to quickly assess the quality of folding and overall extent of structural destabilization of any particular variant molecule, can be of invaluable help in choosing between these various hypotheses.

MATERIALS AND METHODS

E. coli strains and growth media

E. coli strains used in vivo studies were thyA::Km' derivatives of C600 (Bell-Pedersen et al., 1991). Growth medium was TYBET, that is, tryptone broth (1% tryptone and 0.5% NaCl) supplemented with 0.5% yeast extract and 50 mg/L thymine. Bacto agar (1.5%) was added for solid media. Minimum medium supplemented with 0.1% Norit A-treated casamino acids and 0.2% glucose (Belfort et al., 1983) but lacking thymine was used to select for the Thy$^+$ (TS$^+$) phenotype. TTM, a minimum medium supplemented with thymine (50 mg/L) and trimethoprim (20 mg/L) was used to screen for the Thy$^-$ (TS$^-$) and leaky Thy$^+$ phenotypes, as only cells lacking TS activity or producing low levels of TS are able to grow on media containing the folate analog trimethoprim. MMT, a minimal medium supplemented with thymine (50 mg/L) was used as nonselective medium for the plating assays. The cells were grown at 37°C.

DNA constructs

All td constructs were inserted in the vector pTZ18U (U.S. Biochemical). The parental td construct pTZtd $\Delta P$6-2, containing a 751-nt deletion in the L6a region, was described in Galloway-Salvo et al. (1990). The C870U td construct derived therefrom was described in Schroeder et al. (1991). All the other substitutions were introduced by site-directed mutagenesis (Kunkel et al., 1987). All constructs were verified by sequencing the entire length of the intron.

Colony assay for splicing of td introns in E. coli

E. coli C600 thyA::Km' cells were transformed with a pTZtd plasmid containing a td gene and transformants were grown...
overnight at 37°C on plates containing TBYET supplemented with ampicillin (100 mg/L) and kanamycin (40 mg/L). Single colonies were inoculated into liquid TBYET and grown to saturation overnight. Two microliters of cells were spotted on MM, MMT, and TTM plates, and these plates were incubated overnight at 37°C prior to scoring phenotypes.

RNA extraction and primer extension assays

Cell growth and RNA extraction were done as described previously (Belfort et al., 1990) with minor modifications: cells were grown in the presence of 1 mM IPTG (isopropyl-β-D-galactoside) and RNA was precipitated after the addition of 100 vol of Na₂EDTA 0.5 M (pH 8.0) instead of Mg(OAc)₂ 1 M.

In vivo splicing was allowed to cool slowly to room temperature (25°C) within its melting range for 10 min and then subjected to two rounds of precipitation with 70% ethanol before drying and resuspension in water.

Samples were adjusted to an approximate absorbance of 1.2 at 260 nm with a heating rate of 0.2°C/min on a U-2000 spectrophotometer (Hitachi, Ltd.) equipped with the electronic thermostated cell holder (Hitachi, Ltd.) or on a Uvikon 941 plus spectrophotometer (Kontron Instruments) equipped with two cells thermostated by a Neslab RTE 100 circulating bath coupled to a thermoprogrammer (Neslab Instruments). Data were collected from 20°C/30°C to 80°C at 0.05°C intervals. Derivatives of absorbance versus temperature were generated using the locally weighted least squared error method provided by the Kaleidagraph program (Abelbeck Software) on a Macintosh computer.

ACKNOWLEDGMENTS

We are indebted to M. Belfort for the gift of the pTZ lfd Δ1-3 plasmids (Belfort et al., 1987) and to Uwe von Ahlsen and Katharina Semrad for fruitful discussions on primer extension assays. We are grateful to T.R. Cech and an anonymous reviewer for critical reading of the manuscript and precious advice. P.B. thanks Philippe Benas for his input in the deconvolution of melting profiles. P.B. was supported by a Ministère de l’Enseignement Supérieur et de la Recherche fellowship. Part of this work was done by P.B. during a stay at the Vienna Biocenter (Institut für Mikrobiologie und Genetik, Universität Wien) in the laboratory of R.S. with support from the Bundesministerium für Wissenschaft Forschung und Kunst (Vertrag G2200.016/1-IV/6/96).

Received March 3, 1999; returned for revision March 22, 1999; revised manuscript received April 23, 1999

REFERENCES


Lehner V, Jaeger L, Michel F, Westhof E. 1996. New loop-loop terti-
Michel F, Westhof E. 1990. Modelling of the three-dimensional archi-
tecture of group I catalytic introns based on comparative se-
Stern S, Moazed D, Noller HF. 1988. Structural analysis of RNA using chemical and enzymatic probing monitored by primer ex-