Downloaded from www.rnajournal.org on February 14, 2006

Coupled nucleotide covariations reveal dynamic RNA interaction patterns

A. P. Gultyaev, T. Franch and K. Gerdes

RNA 2000 6: 1483-1491

References

Article cited in:
http://www.rnajournal.org/cgi/content/abstract/6/11/1483#otherarticles

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

Notes

To subscribe to *RNA* go to:
http://www.rnajournal.org/subscriptions/

© 2000 RNA Society

Downloaded from www.rnajournal.org on February 14, 2006

RNA (2000), *6*:1483–1491. Cambridge University Press. Printed in the USA.
Copyright © 2000 RNA Society.

PERSPECTIVE

Coupled nucleotide covariations reveal dynamic RNA interaction patterns

ALEXANDER P. GULTYAEV,[1,2] THOMAS FRANCH,[3] and KENN GERDES[3]

[1]Section Theoretical Biology and Phylogenetics, Institute of Evolutionary and Ecological Sciences of Leiden University, 2311 GP Leiden, The Netherlands
[2]Leiden Institute of Chemistry, Leiden, The Netherlands
[3]Department of Molecular Biology, Odense University, DK-5230 Odense M, Denmark

ABSTRACT

Evolutionarily conserved structures in related RNA molecules contain coordinated variations (covariations) of paired nucleotides. Analysis of covariations is a very powerful approach to deduce phylogenetically conserved (i.e., functional) conformations, including tertiary interactions. Here we discuss conserved RNA folding pathways that are revealed by covariation patterns. In such pathways, structural requirements for alternative pairings cause some nucleotides to covary with two different partners. Such "coupled" covariations between three or more nucleotides were found in various types of RNAs. The analysis of coupled covariations can unravel important features of RNA folding dynamics and improve phylogeny reconstruction in some cases. Importantly, it is necessary to distinguish between multiple covariations determined by mutually exclusive structures and those determined by tertiary contacts.

Keywords: antisense RNA; hok; RNA evolution; RNA structure

INTRODUCTION

Comparative analysis of RNA sequences is a well-established approach for elucidating nucleotides that pair in RNA molecules (James et al., 1988; Michel & Westhof, 1990; Woese & Pace, 1993; Gutell & Damberger, 1996; Williams & Bartel, 1996). The phylogenetic approach is successful in those cases where homologous sequences contain nucleotide covariations (i.e., base positions that covary with another position elsewhere in an RNA). It is not only models of conserved secondary structures that can be verified in this way, but also tertiary contacts may be revealed, including noncanonical base pairs.

The examination of secondary structure models to find nucleotide covariations usually precedes the modeling of higher-order structures. Many of the known tertiary interactions were first revealed by analysis of statistically significant nucleotide covariations (Michel & Westhof, 1990; Gutell et al., 1992; Gutell, 1993). A sequential process of three-dimensional modeling often follows the hierarchy of folding, with local secondary structure contacts predicted first, followed by longer-range pairings, and finally tertiary interactions (Michel & Westhof, 1990; Brion & Westhof, 1997). Such hierarchical models generally assume that the interactions suggested by covariations are not disrupted during the folding process and that all of them exist simultaneously. However, the covariation analysis as such does not provide evidence that the postulated interactions coexist in the final structure of a given RNA (Michel & Westhof, 1990).

Several families of phylogenetically related RNAs that contain coupled nucleotide covariations have been described (e.g., ribosomal RNAs and group I introns). Such RNAs contain base positions that covary with two other positions elsewhere in the RNA. In those cases coupled covariations indicate tertiary interactions between three nucleotides (Michel & Westhof, 1990; Michel et al., 1990; Gutell, 1993). However, coupled covariations can also arise from the existence of multiple RNA conformations that are mutually exclusive. Here we describe novel and known examples of coupled nucleotide covariations that reflect mutually exclusive secondary structures.

Reprint requests to: A.P. Gultyaev, Section Theoretical Biology and Phylogenetics, Institute of Evolutionary and Ecological Sciences, Leiden University, Kaiserstraat 63, 2311 GP Leiden, The Netherlands; e-mail: sgultyaev@rulsfb.leidenuniv.nl.

Downloaded from www.rnajournal.org on February 14, 2006

PHYLOGENETICALLY CONSERVED FOLDING PATHWAY IN mRNAs OF THE *HOK* KILLER GENE FAMILY

A unique example of transitions between alternative RNA structures is provided by the family of genes that mediate stabilization of plasmids by killing newborn plasmid-free cells (for reviews, see Gerdes et al., 1990, 1997). The prototype locus, *hok/sok* from plasmid R1, produces a stable toxin-encoding mRNA and an unstable (antidote) antisense RNA. The Sok antisense RNA (suppression of killing) inhibits translation of the toxin *hok* mRNA (host killing). The postsegregational killing stabilizes the plasmids carrying the *hok/sok* system, because the toxin encoded by the mRNA that remains in a cell even after the plasmid loss kills only plasmid-free cells whereas cells retaining the plasmid are rescued by the replenished Sok RNA molecules.

Activation of *hok* translation is regulated by alternative secondary structures in the *hok* mRNA translation initiation region. One of these structures is both translationally active and competent in antisense RNA binding, whereas the other is inert in both respects (Franch et al., 1997). Furthermore, the functionally important structural rearrangements turned out to be triggered by refoldings at the very ends of the mRNA. In these regions, phylogenetic conservation of the specific RNA folding pathway (Gultyaev et al., 1997) results in a unique pattern of coupled covariations that reflects three different functional conformations, present at different stages in the life cycle of the RNA (Fig. 1A).

The folding pathway works as a clock mechanism, keeping the toxin mRNA inactive for a considerable time while providing its activation at a proper moment. During transcription, the metastable hairpin at the very 5′ end (Fig. 1A) induces formation of inactive folding in the translation initiation region, located further downstream in the sequence (not shown in the Fig. 1A). Upon completion of transcription, the inactive mRNA is locked by pairing of the 5′ end to a so-called fold-back-inhibition element (*fbi*), located at the very 3′ end. The *fbi* interaction disrupts the metastable hairpin with the formation of the new stem-loop structure in the 5′-proximal region. In this stable conformation, the full-length mRNA is neither translated nor does it bind the inhibitory antisense RNA. This allows for the accumulation of a pool of inactive mRNAs ready to be activated if the plasmid-born gene is lost. Translation of *hok* mRNA is activated by slow 3′-end processing (Thisted et al., 1994; Nielsen & Gerdes, 1995). The removal of the 3′-terminal nucleotides, including the inhibitory *fbi* element, results in the third conformation at the 5′ end, characterized by the extension of the stable stem (Fig. 1A). This final rearrangement in the truncated RNA triggers the formation of the active structure downstream (not shown), which results either in translation or in Sok antisense RNA binding.

All three conformations, shown in Figure 1A, are vital for the mechanism and conserved in other *hok*-homologous gene systems (Gerdes et al., 1997; Pedersen & Gerdes, 1999). The constituent stems are supported by covariations that are interdependent due to overlaps of mutually exclusive structures. For example (Fig. 1A), the covariation of pair 3–28 in the metastable hairpin (C-G in *hok* and U-A in related gene *srnB*) is coupled with nt 394 (*hok* numbering) in the *fbi* element, which pairs to nt 3 in the full-length RNA and is, correspondingly, G in *hok* and A in *srnB*. Similarly, the covariation of pair 11–19 in the metastable hairpin is coupled with that of pair 19–57 in the stable stem due to the alternative pairings of nt 19.

The folding pathway of *hok*-related mRNAs was predicted by RNA folding simulations using a genetic algorithm (Gultyaev et al., 1995a). The folding pathway was supported by numerous covariations in all regulatory secondary structure elements (Gultyaev et al., 1997). Furthermore, the predicted structures were verified experimentally by mutational analyses, structure probing, and functional assays using translation and antisense RNA binding (Thisted et al., 1995; Nielsen & Gerdes, 1995; Franch & Gerdes, 1996; Franch et al., 1997). Recently, the formation of the metastable hairpin (which, according to the model, folds only transiently), was studied in vitro by structure probing of RNA fragments from *hok* and *pnd* genes, captured in metastable state (Nagel et al., 1999). This approach also allowed the observation of relatively slow kinetics of the hairpin unfolding, consistent with the assumption about its metastable character.

COUPLED COVARIATIONS IN METASTABLE SECONDARY STRUCTURES

The remarkable pattern of multiple covariations in the mRNAs of the *hok* gene family is the result of strong conservation of mutually exclusive secondary structure elements together with surprisingly high sequence variability. We inspected the literature for examples of coupled covariations that reflect alternative secondary structures rather than coexisting interactions in the three-dimensional configuration. The evolution of alternative structures has been analyzed in some detail for the RNA primers involved in the replication of ColE1-related plasmids (Tomizawa, 1993). The primer RNA (RNA II) undergoes a conformational transition during transcription, which is important for the binding of the antisense RNA (RNA I). The transition involves alternative stem-loop structures that are also found in sequences related to ColE1. However, the regions involved in alternative pairing schemes are organized in such a way that the nucleotides covarying in one of them (Tomi-

Downloaded from www.rnajournal.org on February 14, 2006

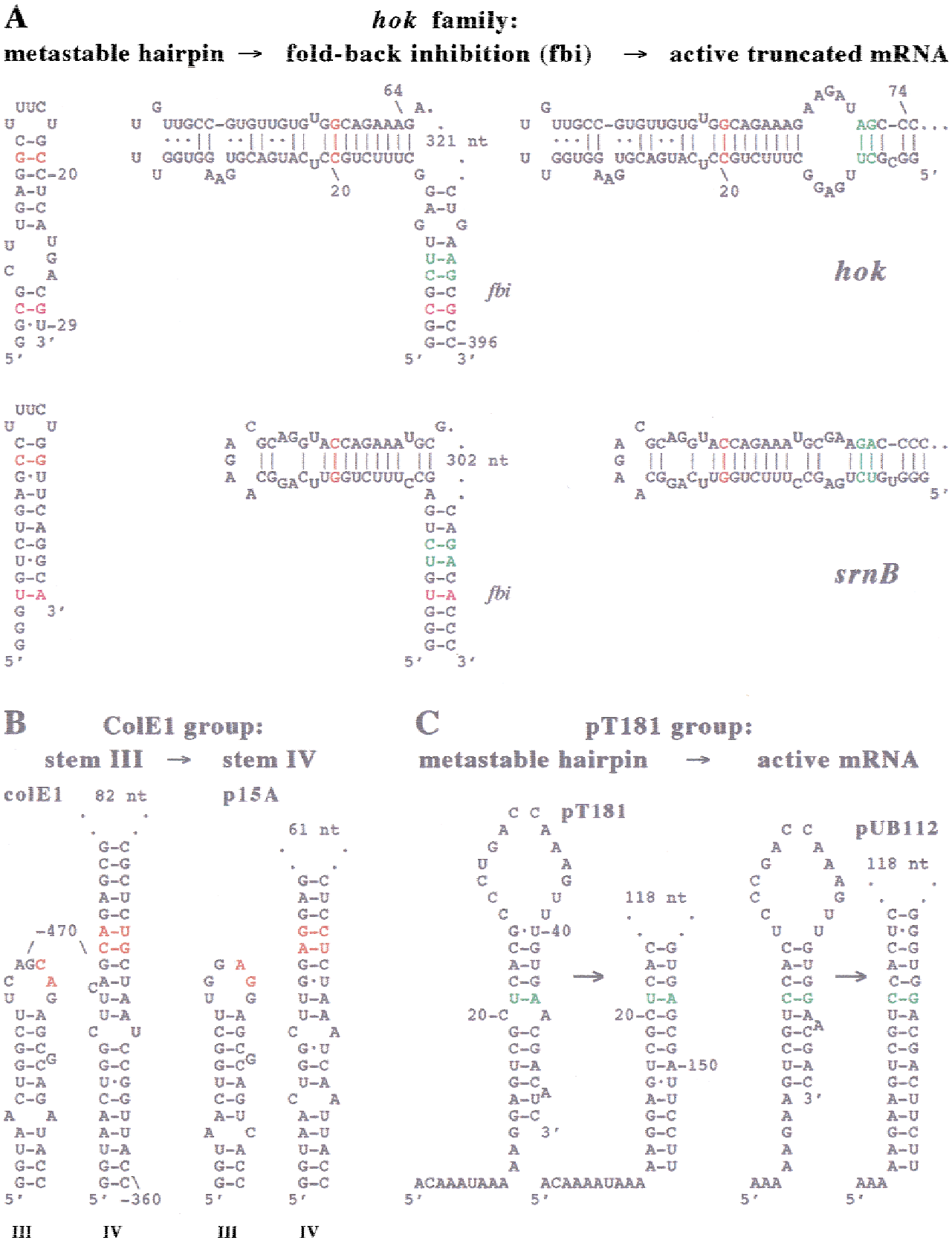

FIGURE 1. Coupled covariations in RNAs undergoing conformational transitions related to antisense RNA regulation. **A**: Refoldings at the 5′ ends of mRNAs from the *hok* gene family (Gultyaev et al., 1997). **B**: Alternative stems in RNA primers regulating copy numbers of plasmids ColE1 and p15A from the ColE1 group (Tomizawa, 1993). **C**: Alternative stems in mRNAs encoding RepC proteins from two plasmids belonging to the pT181 group (Novick et al., 1989). The numbering in the ColE1 primer is in 3′–5′ direction from the replication origin.

Downloaded from www.rnajournal.org on February 14, 2006

zawa, 1993) are located in a loop of the alternative structure and therefore do not result in a coupled covariation (Fig. 1B). On the other hand, the nucleotides that do participate in two mutually exclusive stems are conserved.

More variable sequences were found in the alternative RNA secondary structures involved in the replication control of plasmids from the pT181 family, also regulated by antisense RNAs (Novick et al., 1989). Similar to the ColE1 system, the binding of the antisense RNA is modulated by alternative structures of the target. In turn, RNA–RNA duplex formation influences functional structures downstream of the target region. The target is located in the leader region of the mRNA encoding the replication initiator protein RepC (RNA II); binding of the antisense RNA I to its target in RNA II suppresses translation whereas absence of the binding results in an alternative folding of the mRNA leader that allows translation (Novick et al., 1989).

Analysis of the mutually exclusive structures in the target regions of mRNAs from pT181 and related plasmids did reveal a coupled covariation (Fig. 1C). During transcription, the target hairpin is formed, and, in the absence of antisense RNA, this structure may be refolded into the translatable conformation. Conservation of both structures results in coordinated change of nucleotides: in pT181 the nucleotide U21 pairs either to A44 or to A146, and the substitution of U21 by C in the pUB112 plasmid is accompanied by two A-to-G substitutions (Fig. 1C). Other "simple" covariations in one of the alternative structures do not lead to a change of the third nucleotide because the covarying bases are not involved in the other alternative. Alternative metastable structures are also involved in other antisense RNA-regulated systems (Ma et al., 1994; Pouwels et al., 1994), but we could not find coupled covariations in these cases because of insufficient sequence variation in related sequences.

Coupled covariations are not a unique feature of antisense RNA-regulated systems and may be found in other RNAs that have conserved alternative structures. For example, metastable hairpin conformations were shown to be very important for replication of some viroid RNAs, small plant pathogens (Loss et al., 1991). Analyzing covariations in the base pairs of these hairpins, we have found (not shown) that some of them correlate with base pairing of the stable rodlike folding that is also essential (Wassenegger et al., 1994).

It is noteworthy that in the case of viroids, coupled base–base covariations could be a result of functional structures folded in the RNA strands of opposite polarities rather than derived from a pathway with conformational transitions. Although the rodlike structure seems to be essential for formation of an infectious replicon (Wassenegger et al., 1994), metastable multihairpin conformations are important for the template properties of the viroid (−) strand, where they are formed during transcription (Qu et al., 1993; Repsilber et al., 1999). Recently, another metastable hairpin, presumably functional in the (−) strands of viroids, was revealed in the terminal domain T1 (Gultyaev et al., 1998). This hairpin is not the mirror image of the structure formed by the (+) RNA; therefore nucleotides at some positions should be paired to different partners in the strands of opposite polarities. Considering the high sequence variability in the region and requirements for the (+) strand folding, conservation of this metastable hairpin is remarkable (Gultyaev et al., 1998). Although alignment of nucleotide sequences in this region is not straightforward and it is difficult to indicate unequivocally homologous nucleotide positions, it is clear that variations in base pairs are coordinated so as to maintain different pairings in complementary strands.

CONFORMATIONAL SWITCHES

Alternative secondary structures are not necessarily sequential steps in the folding pathway of an RNA; they may instead form a conformational switch whose operation is dependent on some conditions.

Recently, such a switch in the small ribosomal subunit RNA (SSU rRNA) was shown to be an essential part of the ribosome decoding mechanism (Lodmell & Dahlberg, 1997). The transition occurs in the central part of the molecule, involving alternative pairings of nt 910–912 (*Escherichia coli* numbering) to either positions 888–890 or 885–887. Initially both arrangements were suggested on the basis of comparative analysis and coupled covariations may easily be seen when various sequences are superimposed on the switch model (Lodmell & Dahlberg, 1997). Moreover, the principle of triple complementary mutations may even be used in elegant experiments to show that indeed some balance between the two positions of the switch is important (Lodmell & Dahlberg, 1997). It is interesting that different ribosomal proteins favor the suggested alternatives.

Analysis of nucleotide covariations in a pseudoknot, which is specific for eukaryotic SSU rRNAs (the one formed by so-called helices 21-7 and 21-8) also suggests a dynamic character (Neefs & De Wachter, 1990). Although both stems of the pseudoknot are well supported, some coupled covariations in the junction region indicate the possible extension of one stem at the expense of another. The function of this switch and whether ribosomal proteins can modulate it remains to be elucidated.

A similar pseudoknot-like arrangement of helices was found by sequence comparisons in self-splicing group I introns, where part of a pairing at the 5′ end of the intron (P1 stem) may compete with a long-range interaction with the 3′ exon, stem P10 (Davies et al., 1982). Although the P10 pairing is supported by coupled co-

Downloaded from www.rnajournal.org on February 14, 2006

variations, it is not present in all group I introns (Michel & Westhof, 1990). Nevertheless, it was shown to be involved in a redundant mechanism of organizing the 3′ intron–exon junction for the splicing reaction (Michel et al., 1989).

A conformational switch, regulated by coat-protein binding, has been recently proposed to occur at the 3′ end of alfamovirus and ilarvirus RNAs (Olsthoorn et al., 1999). The structure of the 3′ UTR of the RNA of alfalfa mosaic virus (the type member of alfamoviruses), which is favored by protein binding, was known for a long time and is supported by phylogenetic comparisons. However, an alternative pseudoknot conformation is also possible, and it is supported by covariations that are actually coupled to covariations in the 3′-terminal hairpin (which is strongly conserved in AMV and ilarviruses and recognized by the coat protein). Furthermore, the suggested pseudoknotted conformation, resembling some known 3′-terminal tRNA-like structures, was shown to be favored by Mg^{2+} and involved in the minus-strand synthesis, which is presumably inhibited by coat-protein binding (Olsthoorn et al., 1999). Thus switching between two RNA conformations seems to be related to a switch between different steps of the viral life cycle.

COUPLED COVARIATIONS IN INTERMOLECULAR STRUCTURES

Recently, an analysis of covariations was used to support a functional intermolecular interaction between U3 small nucleolar RNA (snoRNA) and the SSU rRNA (Hughes, 1996). This analysis suggested that U3 snoRNAs interact with highly conserved elements in the 18S-like SSU rRNAs and that this interaction facilitates the correct folding of the central pseudoknot of the rRNA. Presumably the U3 snoRNA binds to the pre-rRNA and brings together distant rRNA regions that comprise the pseudoknot in the mature molecule. Despite very strong conservation of rRNA sequences in these regions, some covariations can be found and they are accompanied by corresponding changes in the U3 snoRNA, thus preserving the possibility for intermolecular pairing. It is remarkable that this pairing scheme is consistent for quite distant groups of eukaryotes and also convincingly supported by coupled covariations in the archaeon *Sulfolobus* (Hughes, 1996). Furthermore, experiments (Mereau et al., 1997) later confirmed this model in a slightly revised version, and binding of U3 to the 5′ end of SSU rRNA was shown to be required for pre-rRNA processing (Sharma & Tollervey, 1999).

Dynamic RNA–RNA interactions occur in spliceosomes, where several conformational transitions involving inter- and intramolecular pairings of small nuclear RNAs (snRNAs) are necessary for nuclear pre-mRNA splicing (for a review, see, e.g., Nilsen, 1998; Staley & Guthrie, 1998). The majority of these transitions are well documented at the level of secondary structure and constituent pairings are strongly conserved. Although many simple covariations may be found in the structures of snRNAs or in the intermolecular stems, the regions involved in alternative pairing schemes are so conserved that they mostly do not contain coupled covariations. Nevertheless, these may be found when very distant species are considered, and serve as strong evidence for similarities in splicing mechanisms (Fast et al., 1998). In Figure 2 we compare the rearrangement from the U2 snRNA 5′-stem to the complex with U6 snRNA, that

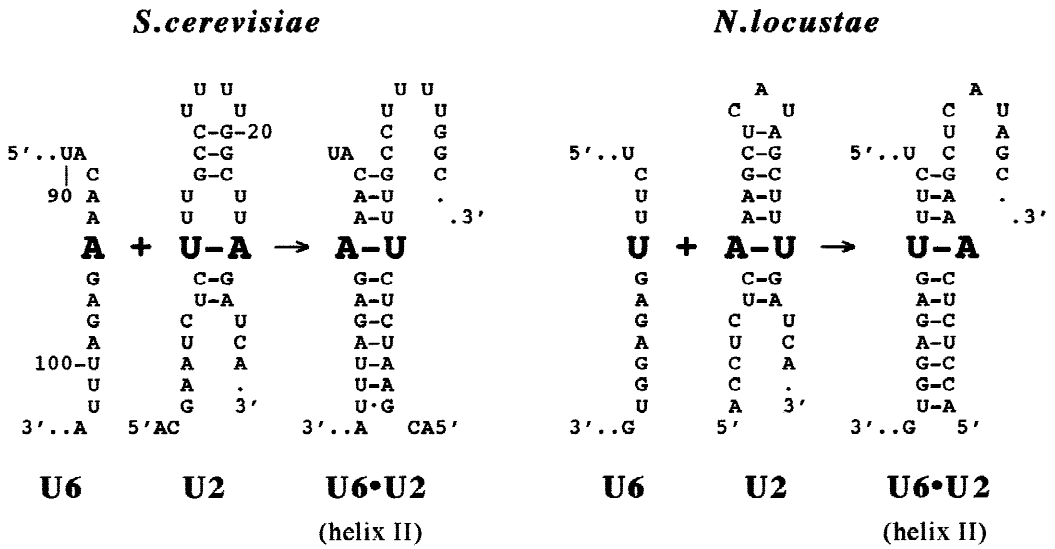

FIGURE 2. Coupled covariations in the spliceosomal RNAs. The formation of the intermolecular pairing between U6 snRNA and U2 snRNA is compared for *S. cerevisiae* and *N. locustae* (Fast et al., 1998). Shown are the sequences involved in the 5′-hairpins of the U2 snRNAs and U2/U6 helices II (see, e.g., Staley & Guthrie, 1998).

Downloaded from www.rnajournal.org on February 14, 2006

is, the so-called U2/U6 helix II (Staley & Guthrie, 1998), in *Saccharomyces cerevisiae* and the suggested similar transition in the microsporidian *Nosema locustae* (Fast et al., 1998). As can be seen, one of the covariations noted for these sequences (Fast et al., 1998) is actually a coupled covariation.

COUPLED COVARIATIONS AND HIERARCHICAL FOLDING

In the examples given above, the intermediate RNA conformations undergo essential rearrangements at the level of secondary structure, breaking Watson–Crick base pairs that are not present in the final folding. Such folding pathways are different from the hierarchical folding that has been described for catalytic RNAs (Zarrinkar & Williamson, 1996; Zarrinkar et al., 1996; Brion & Westhof, 1997).

The ribozyme-folding process resembles that of proteins, with rapid formation of secondary structures, starting from short-range hairpins, followed by a hierarchical folding of higher-order structures. Slow kinetics is observed at the later steps of tertiary structure formation (Zarrinkar & Williamson, 1996; Zarrinkar et al., 1996) and kinetic traps seem to be mostly determined by tertiary contacts (Treiber et al., 1998). For a ribozyme, slow kinetics of folding may be a negative factor, increasing the fraction of nonactive molecules. A misfolded secondary structure can trap a ribozyme in a nonfunctional state for a considerable time due to a very high energy barrier (Thirumalai & Woodson, 1996; Pan et al., 1997). Thus, although parallel folding pathways are probably common for large ribozymes (Pan et al., 1997; Russell & Herschlag, 1999), the paths with kinetic traps should be mostly considered as misfolding and are unlikely to be selected for during evolution.

Nevertheless, conserved alternative secondary structures do exist in ribozymes. The intron/exon interaction P10 in group I introns was mentioned above. Another alternative pairing was shown to affect the splicing of the *Tetrahymena* group I intron from the large ribosomal subunit (LSU) rRNA (Woodson & Cech, 1991). The inhibitory hairpin, designated as P(−1), is located in the 5′ exon immediately preceding the splice site and competes with the P1 hairpin at the splice junction, which is required for splicing. The equilibrium between P(1) and P(−1) turns out to be dependent on another alternative hairpin PX, which suppresses P(−1) (Woodson & Emerick, 1993; Cao & Woodson, 1998). Interestingly, coupled covariations, indicative of a competition between PX and P(−1), were found even in LSU rRNAs without introns, with compensatory changes between eukaryotic, archaebacterial, eubacterial, and chloroplast sequences (Woodson & Emerick, 1993). It was suggested that positions of intron insertions may correlate with regions of conformational rearrangements (Woodson & Emerick, 1993).

Although the role of alternative secondary structures is apparent in regulatory RNA regions, the reason for their presence in catalytic molecules is less clear. A long-range pairing between the 5′ exon and 3′ terminus of the group I intron of the T4 td gene, which has been suggested to inhibit splicing, is disrupted in vivo by the translating ribosomes (Semrad & Schroeder, 1998). Such findings raise interesting questions about regulatory functions of alternative secondary structures in splicing and translation (Woodson, 1998).

The activity of small ribozymes in viroids and viroidlike RNAs seems to be also finely tuned at different steps of the replication cycle by alternative pairings in stable rodlike or branched configurations, or metastable intermediates (Di Serio et al., 1997; Flores et al., 1997; Matysiak et al., 1999). Structural constraints due to such pairings restrict variations in viroid quasi-species, so that the majority of substitutions are located in loops (Ambros et al., 1998) and coupled covariations are unlikely to be observed here. On the other hand, high variability within viroid groups in most cases does not allow straightforward sequence alignments of different species. It may be anticipated that upon accumulation of more sequence data, the analysis of covariations may help to elucidate structural rearrangements during the replication cycle of viroids and viroidlike replicons.

USING BASE COVARIATIONS FOR STRUCTURAL MODELING

It is apparent that quite different types of selective pressures on RNA sequence can lead to very similar patterns of nucleotide covariations. In regulatory sequences, mutually exclusive secondary structures mostly do not involve tertiary contacts, because the regulation is usually dependent on simply engaging some functional regions in structures inhibiting the function. For example, translation can be regulated by base pairing of the Shine–Dalgarno region (de Smit & van Duin, 1990; de Smit, 1998). This may result in coupled covariations, as, for example, in *hok* family RNAs (Franch et al., 1997). However, triple covariations, as described above, may also reflect just the opposite scenario: the need to avoid alternative pairings for efficient formation of more complex tertiary folds as, for example, in ribosomal RNAs or ribozymes. Here coordinated substitutions may arise from base triples (e.g., Michel et al., 1990) or even more complex structures like, for example, tetraloop interactions with various motifs (e.g., Michel & Westhof, 1990; Costa & Michel, 1995; Costa et al., 1997). Although these interactions can be dynamic (Costa et al., 1997), formation of such structures usually does not require refolding at the level of secondary structure, that is, breaking Watson–Crick base pairs, because the accurate formation of some secondary structure is mostly a prerequisite for the tertiary contact.

Downloaded from www.rnajournal.org on February 14, 2006

Based solely on covariation patterns, one cannot distinguish decisively between tertiary contacts and mutually exclusive secondary structures without additional data from molecular modeling, structure probing, or mutational analysis. For example, when on the basis of covariation analysis a base-triple interaction in RNA tertiary structure is envisaged, an analysis of the molecular geometry of suggested base triples is very useful. Invariance of this geometry with respect to underlying covariations ("isomorphism") may be considered as independent support for this element of tertiary structure (Michel et al., 1990; Gautheret & Gutell, 1997). Mutagenesis combined with activity measurements and thermodynamic experiments like, for example, melting (Conn et al., 1998), can be used in the absence of direct structural data.

Simulations of RNA folding pathways may be used as theoretical support for alternative secondary structures, which are often formed during transcription (Gultyaev et al., 1995b, 1997, 1998). As seen above, metastable foldings may be experimentally verified by thorough analysis of functional measurements, mutagenesis data, and/or structural probing under different conditions to resolve different conformations. One of the few direct methods to visualize the presence of a population of conformers is temperature gradient gel electrophoresis (TGGE), which has been successfully applied to viroid RNAs (e.g., Loss et al., 1991; Repsilber et al., 1999). If a kinetic trap captures a molecule for a sufficiently long time, the presence of a metastable structure may be proven by denaturation/renaturation cycles (Pan et al., 1997; Nagel et al., 1999). In some cases, refolding kinetics may be directly monitored by structural or functional assays (Biebricher & Luce, 1992; Poot et al., 1997; Nagel et al., 1999).

COUPLED COVARIATIONS AND PHYLOGENETICS

RNA sequences are often used to infer phylogenetic relations between species (e.g., Woese & Pace, 1993). However, accurate estimates of evolutionary distances and phylogeny reconstruction require certain assumptions on the mechanisms of compensatory substitutions under secondary structure constraints. Intuitively, the process of compensatory substitution in RNA is seen as a two-step process, with temporary lowering of the fitness due to the introduction of a mismatch, followed by a second mutation restoring the pairing. In particular, the relatively high thermodynamic stability of G-U mismatches suggests that transitions between A-U and G-C pairs occur mainly through G-U intermediates. Indeed, this has been shown to be the dominant process of compensatory changes in highly divergent domains of the LSU rRNA of *Drosophila* (Rousset et al., 1991). Such a mechanism implies that the rate of evolution is determined mainly by the frequencies of single substitutions. In principle, coupled covariations may proceed similarly, for example, from a U-A-U combination through U-G-U to C-G-C.

However, such a simple model probably cannot always be applied to reconstruct phylogeny. In a population of RNA molecules, compensatory changes may happen very fast, because mismatches do not always decrease the mean fitness of the population (Stephan, 1996; Higgs, 1998). Therefore covariations appear to be instantaneous on an evolutionary scale, so that the evolution of some sequences is better understood if covariations are explained as if they arose from single genetic events (Tillier & Collins, 1998). The frequency of these events may vary significantly depending on secondary structure constraints: it is estimated that their relative contribution, as compared to the pace of single substitutions, is much higher in slowly changing core rRNA regions than in rapidly evolving domains (Tillier & Collins, 1998). Presumably, the frequencies of triple compensatory changes are lower when compared with simple covariations. Nevertheless, properly taking into account the interdependence of such substitutions may improve phylogenetic reconstructions, especially those based on RNA sequences under strong structural constraints.

ACKNOWLEDGMENTS

We thank Prof. C. Pleij and Dr. M. de Smit for their comments on the manuscript. The work in the K. Gerdes laboratory was supported by the Center for Interaction, Structure, Function and Engineering of Macromolecules (CISFEM). The work was also supported by a short-term grant to A.P.G. from The Plasmid Foundation. A.P.G. was further supported by The Netherlands Foundation for Chemical Research (SON) and Foundation for Life Sciences (SLW) with financial aid from The Netherlands Organisation for Scientific Research (NWO), and in part by the European Commission (project BIO4-98-0189). We also thank an anonymous referee for very useful remarks on examples of covariations.

Received March 31, 1999; returned for revision April 27, 2000; revised manuscript received July 28, 2000

REFERENCES

Ambros S, Hernandez C, Desvignes JC, Flores R. 1998. Genomic structure of three phenotypically different isolates of peach latent mosaic viroid: Implications of the existence of constraints limiting the heterogeneity of viroid quasispecies. *J Virol 72*:7397–7406.

Biebricher CK, Luce R. 1992. In vitro recombination and terminal elongation of RNA by Qβ replicase. *EMBO J 11*:5129–5135.

Brion P, Westhof E. 1997. Hierarchy and dynamics of RNA folding. *Annu Rev Biophys Biomol Struct 26*:113–137.

Cao Y, Woodson SA. 1998. Destabilizing effect of an rRNA stem-loop on attenuator hairpin in the 5′ exon of the *Tetrahymena* pre-rRNA. *RNA 4*:901–914.

Conn GL, Gutell RR, Draper DE. 1998. A functional ribosomal RNA tertiary structure involves a base triple interaction. *Biochemistry 37*:11980–11988.

Costa M, Deme E, Jacquier A, Michel F. 1997. Multiple tertiary interactions involving domain II of group II self-splicing introns. *J Mol Biol 267*:520–536.

Downloaded from www.rnajournal.org on February 14, 2006

Costa M, Michel F. 1995. Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J 14*:1276–1285.
Davies RW, Waring RB, Ray JA, Brown TA, Scazzocchio C. 1982. Making end meet: A model for RNA splicing in fungal mitochondria. *Nature 300*:719–724.
De Smit MH. 1998. Translational control by mRNA structure in Eubacteria: Molecular biology and physical chemistry. In: Simons R, Grunberg-Manago M, eds. *RNA structure and function*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 495–540.
De Smit MH, van Duin J. 1990. Secondary structure of the ribosome binding site determines translational efficiency: A quantitative analysis. *Proc Natl Acad Sci USA 87*:7668–7672.
Di Serio F, Daros JA, Ragozzino A, Flores R. 1997. A 451-nucleotide circular RNA from cherry with hammerhead ribozymes in its strands of both polarities. *J Virol 71*:6603–6610.
Fast NM, Roger AJ, Richardson CA, Doolittle WF. 1998. U2 and U6 snRNA genes in the microsporidian *Nosema locustae*: Evidence for a functional spliceosome. *Nucleic Acids Res 26*:3202–3207.
Flores R, Di Serio F, Hernandez C. 1997. Viroids: The noncoding genomes. *Semin Virol 8*:65–73.
Franch T, Gerdes K. 1996. Programmed cell death in bacteria: Translational repression by mRNA end-pairing. *Mol Microbiol 21*:1049–1060.
Franch T, Gultyaev AP, Gerdes K. 1997. Programmed cell death by *hok/sok* of plasmid R1: Processing at the *hok* mRNA 3′end triggers structural rearrangements that allow translation and antisense RNA binding. *J Mol Biol 273*:38–51.
Gautheret D, Gutell RR. 1997. Inferring the conformation of RNA base pairs and triples from patterns of sequence variations. *Nucleic Acids Res 25*:1559–1564.
Gerdes K, Gultyaev AP, Franch T, Pedersen K, Mikkelsen ND. 1997. Antisense RNA—Regulated programmed cell death. *Annu Rev Genet 31*:1–31.
Gerdes K, Poulsen LK, Thisted T, Nielsen AK, Martinussen J, Hove Andreasen P. 1990. The *hok* killer gene family in gram-negative bacteria. *New Biologist 2*:946–956.
Gultyaev AP, Franch T, Gerdes K. 1997. Programmed cell death by *hok/sok* of plasmid R1: Coupled nucleotide covariations reveal a phylogenetically conserved folding pathway in the *hok* family of mRNAs. *J Mol Biol 273*:26–37.
Gultyaev AP, van Batenburg FHD, Pleij CWA. 1995a. The computer simulation of RNA folding pathways using a genetic algorithm. *J Mol Biol 250*:37–51.
Gultyaev AP, van Batenburg, FHD, Pleij CWA. 1995b. The influence of a metastable structure in plasmid primer RNA on antisense RNA binding kinetics. *Nucleic Acids Res 23*:3718–3725.
Gultyaev AP, van Batenburg FHD, Pleij CWA. 1998. Dynamic competition between alternative structures in viroid RNAs simulated by an RNA folding algorithm. *J Mol Biol 276*:43–55.
Gutell RR. 1993. Comparative studies of RNA: Inferring higher-order structure from patterns of sequence variation. *Curr Opin Struct Biol 3*:313–322.
Gutell RR, Damberger SH 1996. Comparative sequence analysis of experiments performed during evolution. In: Green R, Schroeder R, eds. *Ribosomal RNA and Group I Introns*. Georgetown, Texas: R.G. Landes Company. pp 15–33.
Gutell RR, Power A, Hertz GZ, Putz EJ, Stormo GD. 1992. Identifying constraints on the higher-order structure of RNA: Continued development and application of comparative sequence analysis methods. *Nucleic Acids Res 20*:5785–5795.
Higgs PG. 1998. Compensatory neutral mutations and the evolution of RNA. *Genetica 102–103*:91–101.
Hughes JMX. 1996. Functional base-pairing interaction between highly conserved elements of U3 small nucleolar RNA and the small ribosomal subunit RNA. *J Mol Biol 259*:645–654.
James BD, Olsen GJ, Liu J, Pace NR. 1988. The secondary structure of ribonuclease P RNA, the catalytic element of a ribonucleoprotein enzyme. *Cell 52*:19–26.
Lodmell JS, Dahlberg AE. 1997. A conformational switch in *Escherichia coli* 16S ribosomal RNA during decoding of messenger RNA. *Science 277*:1262–1267.
Loss P, Schmitz M, Steger G, Riesner D. 1991. Formation of a thermodynamically metastable structure containing hairpin II is critical for infectivity of potato spindle tuber viroid RNA. *EMBO J 10*:719–727.
Ma CK, Kolesnikow T, Rayner JC, Simons EL, Yim H, Simons RW. 1994. Control of translation by mRNA secondary structure: The importance of the kinetics of structure formation. *Mol Microbiol 14*:1033–1047.
Matysiak M, Wrzesinski J, Ciesiolka J. 1999. Sequential folding of the genomic ribozyme of the hepatitis delta virus: Structural analysis of RNA transcription intermediates. *J Mol Biol 291*:283–294.
Mereau A, Fournier R, Gregoire A, Mougin A, Fabrizio P, Lührmann R, Branlant C. 1997. An *in vivo* and *in vitro* structure–function analysis of the *Saccharomyces cerevisiae* U3A snoRNP: Protein–RNA contacts and base-pair interaction with the pre-ribosomal RNA. *J Mol Biol 273*:552–571.
Michel F, Hanna M, Green R, Bartel D, Szostak JW. 1989. The guanosine binding site of the *Tetrahymena* ribozyme. *Nature 342*:391–395.
Michel F, Ellington AD, Couture S, Szostak JW. 1990. Phylogenetic and genetic evidence for base-triples in the catalytic domain of group I introns. *Nature 347*:578–580.
Michel F, Westhof E. 1990. Modeling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J Mol Biol 216*:585–610.
Nagel JHA, Gultyaev AP, Gerdes K, Pleij CWA. 1999. Metastable structures and refolding kinetics in *hok* mRNA of plasmid R1. *RNA 5*:1408–1419.
Neefs JM, De Wachter R. 1990. A proposal for the secondary structure of a variable area of eukaryotic small ribosomal subunit RNA involving the existence of a pseudoknot. *Nucleic Acids Res 18*:5695–5704.
Nielsen AK, Gerdes K. 1995. Mechanism of post-segregational killing by *hok*-homologue *pnd* of plasmid R483: Two translational control elements in the *pnd* mRNA. *J Mol Biol 249*:270–282.
Nilsen TW. 1998. RNA–RNA interactions in nuclear pre-mRNA splicing. In: Simons R, Grunberg-Manago M, eds. *RNA structure and function*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 279–307.
Novick RP, Iordanescu S, Projan SJ, Kornblum J, Edelman I. 1989. PT181 plasmid replication is regulated by a countertranscript-driven transcriptional attenuator. *Cell 59*:395–404.
Olsthoorn RCL, Mertens S, Brederode FT, Bol JF. 1999. A conformational switch at the 3′ end of a plant virus RNA regulates viral replication. *EMBO J 18*:4856–4864.
Pan J, Thirumalai D, Woodson SA. 1997. Folding of RNA involves parallel pathways. *J Mol Biol 273*:7–13.
Pedersen K, Gerdes K. 1999. Multiple *hok* genes on the chromosome of *Escherichia coli*. *Mol Microbiol 32*:1090–1102.
Poot RA, Tsareva NV, Boni IV, van Duin J. 1997. RNA folding kinetics regulates translation of phage MS2 maturation gene. *Proc Natl Acad Sci USA 94*:10110–10115.
Pouwels PH, van Luijk N, Leer RJ, Posno M. 1994. Control of replication of the *Lactobacillus pentosus* plasmid p353-2: Evidence for a mechanism involving transcriptional attenuation of the gene coding for the replication protein. *Mol Gen Genet 242*:614–622.
Qu F, Heinrich C, Loss P, Steger G, Tien P, Riesner D. 1993. Multiple pathways of reversion in viroids for conservation of structural elements. *EMBO J 12*:2129–2139.
Repsilber D, Wiese S, Rachen M, Schröder AW, Riesner D, Steger G. 1999. Formation of metastable RNA structures by sequential folding during transcription: Time-resolved structural analysis of potato spindle tuber viroid (−)-stranded RNA by temperature-gradient gel electrophoresis. *RNA 5*:574–584.
Rousset F, Pelandakis M, Solignac M. 1991. Evolution of compensatory substitutions through G.U intermediate state in *Drosophila* rRNA. *Proc Natl Acad Sci USA 88*:10032–10036.
Russell R, Herschlag D. 1999. New pathways in folding of the *Tetrahymena* group I RNA enzyme. *J Mol Biol 291*:1155–1167.
Semrad K, Schroeder R. 1998. A ribosomal function is necessary for efficient splicing of the T4 phage *thymidylate synthase* intron *in vivo*. *Genes & Dev 12*:1327–1337.
Sharma K, Tollervey D. 1999. Base pairing between U3 small nucleolar RNA and the 5′ end of 18S rRNA is required for pre-rRNA processing. *Mol Cell Biol 19*:6012–6019.
Staley JP, Guthrie C. 1998. Mechanical devices of the spliceosome: Motors, clocks, springs, and things. *Cell 92*:315–326.

Downloaded from www.rnajournal.org on February 14, 2006

Stephan W. 1996. The rate of compensatory evolution. *Genetics 144*:419–426.

Thirumalai D, Woodson SA. 1996. Kinetics of folding of proteins and RNA. *Acc Chem Res 29*:433–439.

Thisted T, Nielsen AK, Gerdes K. 1994. Mechanism of post-segregational killing: Translation of Hok, SrnB and Pnd mRNAs of plasmids R1, F and R483 is activated by 3′-end processing. *EMBO J 13*:1950–1959.

Thisted T, Sorensen NS, Gerdes K. 1995. Mechanism of post-segregational killing: Secondary structure analysis of the entire *hok* mRNA from plasmid R1 suggests a fold-back structure that prevents translation and antisense RNA binding. *J Mol Biol 247*:859–873.

Tillier ERM, Collins RA. 1998. High apparent rate of simultaneous compensatory base-pair substitutions in ribosomal RNA. *Genetics 148*:1993–2002.

Tomizawa J. 1993. Evolution of functional structures of RNA. In: Gesteland RF, Atkins JF, eds. *The RNA world.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 419–445.

Treiber DK, Rook MS, Zarrinkar PP, Williamson JR. 1998. Kinetic intermediates trapped by native interactions in RNA folding. *Science 279*:1943–1946.

Wassenegger M, Heimes S, Sanger HL. 1994. An infectious viroid RNA replicon evolved from an *in vitro*-generated non-infectious viroid deletion mutant via a complementary deletion *in vivo*. *EMBO J 13*:6172–6177.

Williams KP, Bartel DP. 1996. Phylogenetic analysis of tmRNA secondary structure. *RNA 2*:1306–1310.

Woese CR, Pace NR. 1993. Probing RNA structure, function, and history by comparative analysis. In: Gesteland RF, Atkins JF, eds. *The RNA world.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 91–117.

Woodson SA. 1998. Ironing out the kinks: Splicing and translation in bacteria. *Genes & Dev 12*:1243–1247.

Woodson SA, Cech TR. 1991. Alternative secondary structures in the 5′ exon affect both forward and reverse self-splicing of the *Tetrahymena* intervening sequence RNA. *Biochemistry 30*:2042–2050.

Woodson SA, Emerick VL. 1993. An alternative helix in the 26S rRNA promotes excision and integration of the *Tetrahymena* intervening sequence. *Mol Cell Biol 13*:1137–1145.

Zarrinkar PP, Wang J, Williamson JR. 1996. Slow folding kinetics of RNase P RNA. *RNA 2*:564–573.

Zarrinkar PP, Williamson JR. 1996. The kinetic folding pathway of the *Tetrahymena* ribozyme reveals possible similarities between RNA and protein folding. *Nature Struct Biol 3*:432–438.