Nucleic Acids Research, 1996, Vol. 24, No. 3 **441–449**

Extended interactions between the primer $tRNA_i^{Met}$ and genomic RNA of the yeast Ty1 retrotransposon

S. Friant, T. Heyman[1], M. L. Wilhelm and F. X. Wilhelm*

Unité Propre de Recherche 9002 du Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, 67084 Strasbourg Cedex, France and [1]Unité de Recherche Associée 1342 du Centre National de la Recherche Scientifique, Institut Curie-Biologie, Centre Universitaire, 91405 Orsay Cedex, France

Received October 25, 1995; Revised and Accepted December 5, 1995

ABSTRACT

Reverse transcription of the yeast Ty1 retrotransposon is primed by $tRNA_i^{Met}$ base paired to the primer binding site near the 5′-end of Ty1 genomic RNA. To understand the molecular basis of the $tRNA_i^{Met}$–Ty1 RNA interaction the secondary structure of the binary complex was analysed. Enzymatic probes were used to test the conformation of $tRNA_i^{Met}$ and of Ty1 RNA in the free form and in the complex. A secondary structure model of the $tRNA_i^{Met}$–Ty1 RNA complex consistent with the probing data was constructed with the help of a computer program. The model shows that besides interactions between the primer binding site and the last 10 nt at the 3′-end of $tRNA_i^{Met}$, three short regions of Ty1 RNA named boxes 0, 1 and 2.1 interact with the T and D stems and loops of $tRNA_i^{Met}$. Mutations were made in the boxes or in the complementary sequences of $tRNA_i^{Met}$ to study the contribution of these sequences to formation of the complex. We find that interaction with at least one of the two boxes 0 or 1 is absolutely required for efficient annealing of the two RNAs. Sequence comparison showing that the primary sequence of the boxes is strictly conserved in Ty1 and Ty2 elements and previously published *in vivo* results underline the functional importance of the primary sequence of the boxes and suggest that extended interactions between genomic Ty1 RNA and the primer $tRNA_i^{Met}$ play a role in the reverse transcription pathway.

INTRODUCTION

The Ty transposable elements are a family of retrotransposons in the yeast *Saccharomyces cerevisiae* (1). They are structurally and functionally similar to several animal retroviruses. Like retroviruses they alternate their genetic material between RNA and DNA (2,3). The genomic Ty RNA, which forms a substrate for reverse transcriptase, is packaged inside virus-like particles (VLPs) together with a specific host-encoded tRNA used as a primer to initiate reverse transcription. For the best characterized yeast retrotransposon Ty1 (4) the primer selected from the pool of cellular tRNAs is initiator methionine tRNA ($tRNA_i^{Met}$). Initiation of reverse transcription involves the formation of a binary complex between the primer $tRNA_i^{Met}$ and Ty1 genomic RNA (5,6). The last 10 nt of the acceptor stem of $tRNA_i^{Met}$ anneal to the complementary sequences of the so-called minus strand primer binding site (PBS) located immediatly adjacent to the 5′ long terminal repeat (LTR) of Ty1 RNA. Reverse transcription is initiated from the 3′ hydroxyl group of the primer $tRNA_i^{Met}$ annealed at the PBS. Our recent results showing that three short sequences of Ty1 RNA (boxes 1, 2.1 and 2.2; see Fig. 1) complementary to the T and D stems and loops of the primer $tRNA_i^{Met}$ are absolutely required for transposition of the Ty1 element (7) suggest that the interactions between the primer $tRNA_i^{Met}$ and genomic Ty1 RNA are not limited to base pairing between the 10 nt of the acceptor stem of $tRNA_i^{Met}$ and the complementary PBS. The existence of extended contacts of the primer tRNA with sequences outside the PBS have been found not only for Ty1, but have also been observed in several retroviruses. It has been proposed that interaction between seven bases of the T loop of the $tRNA^{Trp}$ primer and sequences in the U5 region of the 5′ LTR may be required in an early step of reverse transcription by Rous sarcoma virus (RSV) reverse transcriptase (8,9) and may be necessary for annealing of the primer tRNA to the viral RNA (10,11). Sequence and structure comparisons of the 5′ region of several retroviral RNAs indicate that similar interactions may exist in other retroviruses (12). For HIV-1 recent enzymatic and chemical probing of the conformation of the HIV-1 RNA–$tRNA_3^{Lys}$ complex reveals a compact structure in which most of the anticodon loop, the 3′ strand of the anticodon stem and the 5′ part of the variable loop of $tRNA_3^{Lys}$ interact with viral sequences 12–39 nt upstream of the PBS (13). It has been suggested that these interactions produce a specific orientation of the RNA secondary structure preferentially recognized by reverse transcriptase or may be involved in annealing and encapsidation of the primer tRNA.

Here we report a structural probing study of the interaction between the Ty1 RNA template and the $tRNA_i^{Met}$ primer. In the proposed secondary structure model of the $tRNA_i^{Met}$–Ty1 RNA complex boxes 1 and 2.1 of Ty1 RNA form Watson–Crick base pairs with the primer $tRNA_i^{Met}$. Box 2.2 is not annealed to the primer $tRNA_i^{Met}$ but we find that another region of Ty1 RNA, named box 0, interacts with the primer $tRNA_i^{Met}$. By introducing specific nucleotide changes in Ty1 RNA or in $tRNA_i^{Met}$ that

* To whom correspondence should be addressed

disrupt some of the Watson–Crick base pairs between the two RNAs the role of interactions between the boxes and the primer $tRNA_i^{Met}$ on formation of the binary complex was further investigated. We show that proper complementarity between boxes 0 and 1 of Ty1 RNA and $tRNA_i^{Met}$ is essential for annealing of the two RNAs and that the role of box 2.1 is not as crucial for formation of the complex.

The *in vitro* structural study reported here, sequence comparison of the Ty1 and Ty2 elements and previously published *in vivo* results underline the functional importance of the primary sequence of the boxes and suggest that extended interactions between genomic Ty1 RNA and the primer $tRNA_i^{Met}$ play a role in the reverse transcription pathway.

MATERIALS AND METHODS

Plasmid construction

Standard procedures were used for restriction enzyme digestion and plasmid construction (14). *Escherichia coli* strain MV1190 was used for plasmid amplification. The DNA sequences of all constructed plasmids were confirmed by the dideoxy sequencing method (15).

Plasmid pSF4, containing nt 241–817 of the Ty1 H3 element under the control of the bacteriophage T7 promoter, was obtained by cloning a 0.58 kb *Xho*I–*Hpa*I fragment of plasmid pGTy1H3 (4), kindly provided by J. D. Boeke, into plasmid pSL1180 (Pharmacia 27-4384) and inserting the bacteriophage T7 promoter by site-directed mutagenesis (16) with the following oligodeoxynucleotide 5′-CCACTAGTACTTCTCATCGATTAATACGACTCACTAAGGAGAACTTCTAG-3′. After digestion of pSF4 with *Pvu*II or *Pfl*MI and *in vitro* transcription using T7 RNA polymerase we obtained the two Ty1 RNAs 1–237 and 1–297, starting at position +1 of the Ty1 retrotransposon RNA and containing no additional sequence derived from the plasmid.

Plasmids pSF1, pSF2 and pSF12, with mutations in boxes 1, 2.1 or 1 and 2.1, were obtained by subcloning a 0.362 kb *Acc*I fragment of the mutant elements previously described (7) into the *Acc*I site of pSF4. Plasmids were digested with *Pvu*II and transcribed *in vitro* with T7 RNA polymerase.

The plasmid $ptRNA^{Met}$ (17), kindly provided by B. Senger, contains a synthetic gene corresponding to the yeast initiator $tRNA^{Met}$ sequence with the upstream consensus promoter of T7 RNA polymerase. Mutants $tRNA_i^{Met}$ B1, B2, B01 and B012 were constructed using the method of Kunkel (16). Mutations B1, B2 and B01 were created with oligodeoxynucleotides B1 (5′-CCCTGATGTTCTGGGGTCGAAACCGAGCGTA-3′), B2 (5′-GGCGCCGTGGCGCGGTACTTGCGCGCAGGGCTC-3′) and B01 (5′-CCCTGATGTTCTGGGGTCGAAACAGCGGCGCTA-3′). Mutation B012 was created using the two mutagenic oligodeoxynucleotides B2 and B01. The plasmids were digested with *Bst*NI before transcription by T7 RNA polymerase to generate the mature tRNAs.

RNA synthesis and purification

T7 RNA polymerase was purified from the overproducing strain BL21/pAR1219 (18). Transcripts were prepared in a 500 µl reaction mixture containing 40 mM Tris–HCl, pH 8, 10 mM $MgCl_2$, 1 mM spermidine, 5 mM DTE, 4 mM each NTP, 20 mM GMP, 0.01% Triton X-100, 60 U RNasin, 50 µg digested template DNA and 500 U T7 RNA polymerase. The transcription medium was incubated for 1 h at 37°C. The transcripts were purified on 8% polyacrylamide gels, eluted for 16 h at 4°C into 10 mM Tris–HCl, pH 8, 250 mM NaCl, 1 mM EDTA, 0.1% SDS, precipitated with ethanol and resuspended in water prior to use. RNAs were checked for purity and integrity on 8% denaturing polyacrylamide baby gels.

The $tRNA_i^{Met}$–Ty1 RNA complex: determination of the optimal hybridization conditions

Annealing of Ty1 RNA and $tRNA_i^{Met}$ was performed as follows. The two RNAs dissolved in water were heated for 2 min at 90°C, chilled for 2 min on ice and incubated for 30 min at the hybridization temperature in 50 mM Tris–HCl, pH 8, 50 mM KCl. After hybridization, samples were chilled on ice and incubated for 15 min at 20°C in 50 mM Tris–HCl, pH 8, 300 mM KCl, 5 mM $MgCl_2$. The samples were loaded on 2% MetaPhor agarose gels after addition of loading buffer containing glycerol. Ethidium bromide (0.5 µg/ml) was included in the gels to visualize the RNA bands. Electrophoresis was at 20°C in 90 mM Tris–borate buffer, pH 8.3, at 7 V/cm for 3–4 h.

A series of parameters (temperature, ionic strength and concentration of $tRNA_i^{Met}$) were tested to optimize formation of the complex. The effect of temperature was studied in the range 30–75°C. We found that the optimal temperature for hybridization was 50°C. The influence of ionic strength on hybridization was investigated in the presence of increasing KCl concentration (50–300 mM). $tRNA_i^{Met}$–Ty1 RNA complex formation was better at low salt concentrations (50–100 mM KCl). At high KCl concentration (300 mM) binary complex formation was inefficient, but once the two RNAs were annealed the ionic strength could be increased and magnesium ions could be added to stabilize the RNA structure. Ty1 RNA fragments were annealed with increasing $tRNA_i^{Met}$ concentrations to determine the amount of tRNA necessary to hybridized all Ty1 RNA. We found that a three times excess of $tRNA_i^{Met}$ allowed all Ty1 RNA to be hybridized.

Enzymatic probing of $tRNA_i^{Met}$

$tRNA_i^{Met}$ 5′- and 3′-end-labelling. $tRNA_i^{Met}$ dephosphorylated with bacterial alkaline phosphatase was 5′-end-labelled using T4 polynucleotide kinase and [γ-^{32}P]ATP. $tRNA_i^{Met}$ was 3′-end-labelled using tRNA nucleotidyl transferase and [α-^{32}P]ATP. Labelled tRNAs were purified by electrophoresis on a 8% denaturing polyacrylamide gel.

Enzymatic cleavage. A standard assay contained 0.01 µg 3′- or 5′-end-labelled $tRNA_i^{Met}$ (50 000 c.p.m.), free or hybridized with a 10-fold excess of Ty1 RNA 1–237. For each reaction a control was treated in parallel omitting the enzyme. After addition of 1 µg carrier *E.coli* tRNA samples were treated with 0.01 U/µg RNase T1, 0.06 U/µg RNase T2 or 0.035 U/µg RNase V1. Incubation was at 20°C for 5, 10 or 15 min. After phenol/chloroform extraction and ethanol precipitation tRNAs were loaded on 8% polyacrylamide denaturing gels. In order to discriminate between primary and secondary RNase cuts all RNase probing experiments were conducted on 3′- and 5′-labelled $tRNA_i^{Met}$ and only cuts that were observed in both types of experiments were taken into account (19,20).

Probing of Ty1 RNA

A standard assay contained 0.03 µg Ty1 RNA (1–237 or 1–297), free or hybridized with 0.1 µg $tRNA_i^{Met}$, incubated in 50 mM Tris–HCl, pH 8, 300 mM KCl, 5 mM $MgCl_2$. After addition of 1 µg carrier *E.coli* tRNA samples were treated with 0.015 U/µg

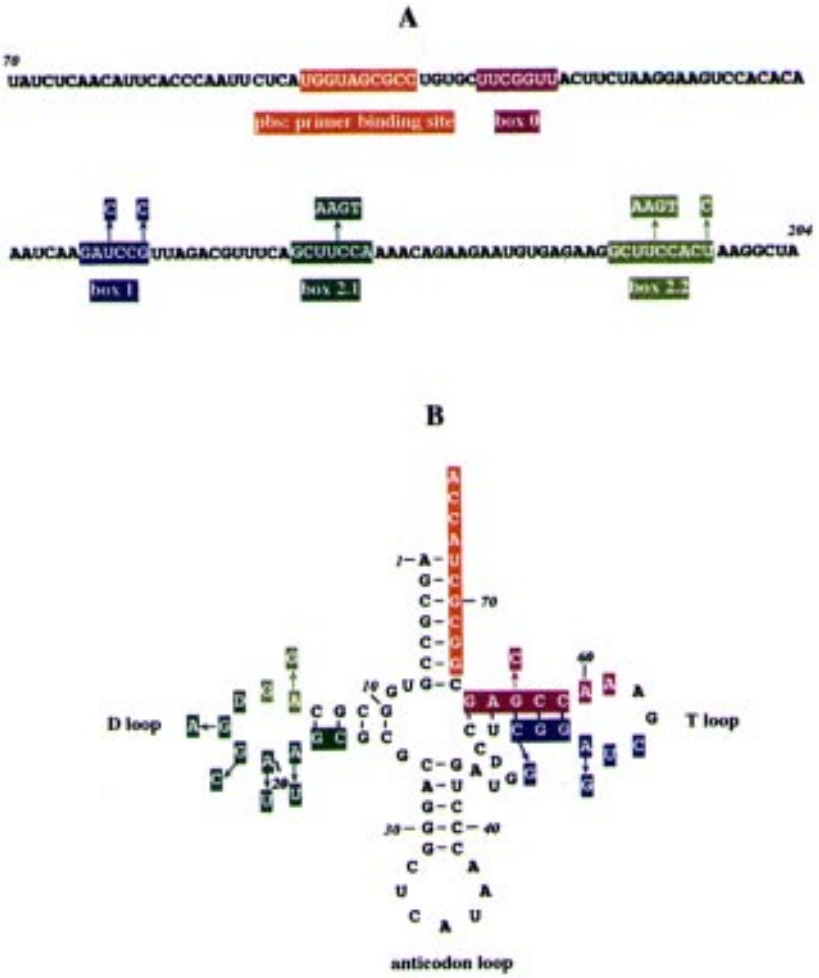

Figure 1. (**A**) Sequence of the Ty1 RNA including the PBS and the four boxes (0, 1, 2.1 and 2.2) complementary to the T and D stems and loops of $tRNA_i^{Met}$. Boxes 2.1 and 2.2 comprise 7 or 9 nt complementary to the same region of the tRNA. Mutations in the boxes that disrupt complementarity with $tRNA_i^{Met}$ are indicated. These mutations do not change the coding sequence for the TYA protein. (**B**) Cloverleaf structure of yeast $tRNA_i^{Met}$. The regions of $tRNA_i^{Met}$ complementary to the PBS and to the boxes are indicated. Mutations in $tRNA_i^{Met}$ which restore complementarity with mutant Ty1 RNAs are indicated. Mutant tRNA B1, C51 is changed to G51 and A54 to G54; mutant tRNA B2, A14 is changed to G14, G18 to A18, G19 to C19, A20 to U20 and A21 to U21; mutant tRNA B01, C51 is changed to G51, A54 to G54 and G63 to C63; mutant B012, A14 is changed to G14, G18 to A18, G19 to C19, A20 to U20, A21 to U21, C51 to G51, A54 to G54 and G63 to C63.

RNase T1, 0.08 U/µg RNase T2 or 0.05 U/µg RNase V1. For each reaction a control was treated in parallel omitting the enzyme. Incubation was at 20°C for 5 or 10 min. All reactions were stopped with phenol/chloroform extraction and ethanol precipitation. The A(N1) and C(N3) Watson–Crick positions of free Ty1 RNA 1–297 were also probed with dimethylsulfate (DMS). After addition of 1 µg carrier tRNA 0.03 µg Ty1 RNA 1–297 were treated for 10 min with 0, 0.5, 1 or 1.5 µl DMS diluted 20 times in ethanol. Reactions were stopped by ethanol precipitation. Detection of cleavage or modification was done using primer extension with reverse transcriptase. Three different primers complementary to residues 116–134, 184–201 and 234–251 of Ty1 RNA were used. Primers were 5′-end-labelled using T4 polynucleotide kinase and [γ-^{32}P]ATP. For annealing the labelled primer (100 000 c.p.m.) was added to the hydrolysed or modified template and heated for 2 min at 90°C. The mixture was then returned to room temperature and allowed to anneal for 10 min. Elongation of the primer by avian myeloblastosis virus reverse transcriptase and analysis of the generated cDNA fragments were performed as described (21).

RNA secondary structure analysis

RNA structure predictions were performed using the RNAFold program (version 2) of Zucker and co-workers (22,23). The possible alternative foldings in the range 10% less stable than the computed lowest free energy structure were predicted. The different foldings were screened and secondary structure motifs consistent with the probing data were selected. The folding program was then allowed to run again, forcing integration of these secondary structure elements. The secondary structure model was obtained after several cycles of refinement.

RESULTS

The Ty1 RNA sequence (4) extending from nt 70 to 204 is shown in Figure 1. It includes the minus strand PBS and four short sequences (boxes 0, 1, 2.1 and 2.2) complementary to the $tRNA_i^{Met}$ T and D stems and loops. In a previous work we showed that the three boxes 1, 2.1 and 2.2 contribute to packaging of the primer

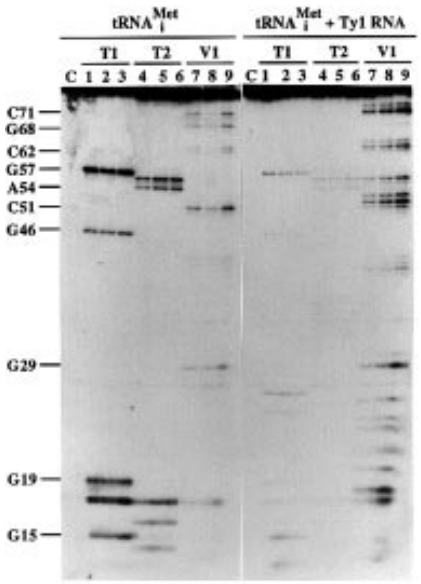

Figure 2. Enzymatic probing of 5′-end-labelled tRNA$_i^{Met}$ in the free form or hybridized to Ty1 RNA. tRNA$_i^{Met}$ was probed with RNase T1 specific for single-stranded G (lanes 1–3), RNase T2 specific for single-stranded RNA (lanes 4–6) or with RNase V1 specific for double-stranded RNA (lanes 7–9). Incubation was for 5 (lanes 1, 4 and 7), 10 (lanes 2, 5 and 8) or 15 min (lanes 3, 6 and 9). Controls were incubated for 15 min without enzyme (lane C).

tRNA$_i^{Met}$ into the VLPs and are essential for transposition of the Ty1 element (7). To understand the molecular basis of tRNA$_i^{Met}$–Ty1 RNA interaction Ty1 RNA fragments and tRNA$_i^{Met}$ were prepared by *in vitro* transcription and the secondary structure of the two RNAs in both the free state and the binary complex was investigated. Mutations were made in the boxes of Ty1 RNA or in the complementary sequences of tRNA$_i^{Met}$ (Fig. 1) to study the contribution of these sequences to formation of the tRNA$_i^{Met}$–Ty1 RNA complex.

Probing of the tRNA$_i^{Met}$

RNase T1 (specific for single-stranded G), RNase T2 (specific for single-stranded regions) and RNase V1 (specific for double-stranded and stacked regions) were used to determine the regions of tRNA$_i^{Met}$ involved in formation of the tRNA$_i^{Met}$–Ty1 RNA complex. Positions of enzymatic cleavage within the primer tRNA$_i^{Met}$ were identified directly using 3′- or 5′-end-labelled tRNA$_i^{Met}$. A probing gel of 5′-end-labelled tRNA$_i^{Met}$ in the free form and in the Ty1 RNA–tRNA$_i^{Met}$ complex is shown in Figure 2. Differences in reactivity of the tRNA$_i^{Met}$ in the free form compared with the bound state are observed at the 3′-end of the tRNA and in the T and D stems and loops. The appearance of strong RNase V1 cuts at the 3′-end of hybridized tRNA$_i^{Met}$ (nt C71, U72 and A73) is consistent with annealing of this region to the PBS. In the T loop residues 54–56, which were RNase T2 sensitive in free tRNA$_i^{Met}$, became unreactive to this enzyme in the complex. Conversely, nt 56 became accessible to RNase V1 in hybridized tRNA$_i^{Met}$. Nucleotides 51–53 of the T stem, which are poorly cleaved by RNase V1 in free tRNA$_i^{Met}$, probably because of the compact tertiary structure of tRNA$_i^{Met}$, are strongly cleaved by this enzyme in the complex. The 3′ part of the T stem (nt 61–65) appears to be double-stranded in the primer–template complex, since nt 61–63 are strongly reactive to RNase V1 (Fig. 2). Thus most of the 3′ part of tRNA$_i^{Met}$ (residues 51–76), with the exception of G57 and A58, are base paired in the binary complex. In the secondary structure model of the tRNA$_i^{Met}$–Ty1 RNA complex (Fig. 4B) we propose that, in addition to interaction between the 10 nt at the 3′-end of tRNA$_i^{Met}$ with the PBS, intermolecular interactions occur between nt 51–56 of tRNA$_i^{Met}$ and nt 144–149 of Ty1 RNA (box 1) and nt 59–65 of tRNA$_i^{Met}$ and nt 110–116 of Ty1 RNA (box 0).

Reactivity changes were also observed in the D loop. In hybridized tRNA$_i^{Met}$ nt U16, G18 and G19 are no longer cleaved by single-strand-specific nucleases. In contrast nt G18–G22 are cut by double-strand-specific nuclease RNase V1. This can be explained by the existence of base pairing between box 2.1 of Ty1 RNA and the complementary nucleotides of tRNA$_i^{Met}$. Other reactivity changes occuring upon formation of the complex involve nt G24–A28 and C11–C13, which become sensitive to RNase V1, and nucleotide G46, which is no longer cleaved by RNase T1. A possible explanation for these observations is an intramolecular rearrangement of tRNA$_i^{Met}$ resulting in base pairing of nt G24–G26 with nt U45–U47 to form an extended anticodon stem and base pairing of nt G3–C6 with nt G10–C13 to form a new stem–loop at the 5′-end of tRNA$_i^{Met}$ (Fig. 4B).

Probing of the viral RNA

In vitro-transcribed Ty1 RNA fragments spanning nt 1–237 or 1–297 were used to analyse the structure of Ty1 RNA in the free form and in the binary complex. Double-stranded and single-stranded regions of RNA were assigned using specific enzymatic and chemical probes under conditions that maintain the secondary structure of the RNA (see Materials and Methods). Examples of enzymatic probing gels of Ty1 RNA in the free form and in the binary complex are shown in Figure 3. The enzymatic and chemical reactivities obtained are compiled in Table 1. These data were used to derive a secondary structure model of Ty1 RNA in the free form (Fig. 4A) and in the binary complex (Fig. 4B). As shown in Figure 4, Ty1 RNA is composed of a variable domain and of regions whose structures remain unchanged upon formation of the tRNA$_i^{Met}$–Ty1 RNA complex.

The unchanged domain comprises nt 72–94, 117–130 and 169–203 (Table 1). A stem interrupted by a bulged loop (nt 83–88) can be formed by base pairing nt 72–92 to nt 173–186 and two stem–loops can be formed by nt 117–130 and 188–203. Interestingly box 2.2 is included in the nt 188–203 stem–loop, which remains unchanged upon formation of the tRNA$_i^{Met}$–RNA Ty1 complex, indicating that this box does not interact with tRNA$_i^{Met}$ in the complex.

The variable domain comprising the PBS and boxes 0, 1 and 2.1 spans nt 95–168. In the free form of Ty1 RNA the PBS (nt 95–104) is base paired with box 2.1 (nt 162–168); this helix is interrupted by an internal U–U mismatch which could form a non-canonical U–U pair (24). Annealing of the PBS with box 2.1 in free Ty1 RNA is corroborated by the reactivity pattern of the Ty1 RNA box 2.1 mutant in which nt 164–167 (UUCC) are changed to AAGU. In this mutant box 2.1 and the PBS can no longer base pair and we find indeed that nt 161–165 of box 2.1 and nt 99–102 of the PBS are single-stranded (Fig. 3C and D).

In the tRNA$_i^{Met}$–Ty1 RNA complex the 10 nt at the 3′-end of the tRNA are base paired with the PBS. Accordingly, all nucleotides of the PBS become reactive to the double-strand-specific nuclease RNase V1 (Fig. 3D). Annealing of the PBS with tRNA$_i^{Met}$ is expected to disrupt base pairing with box 2.1. However, box 2.1 does

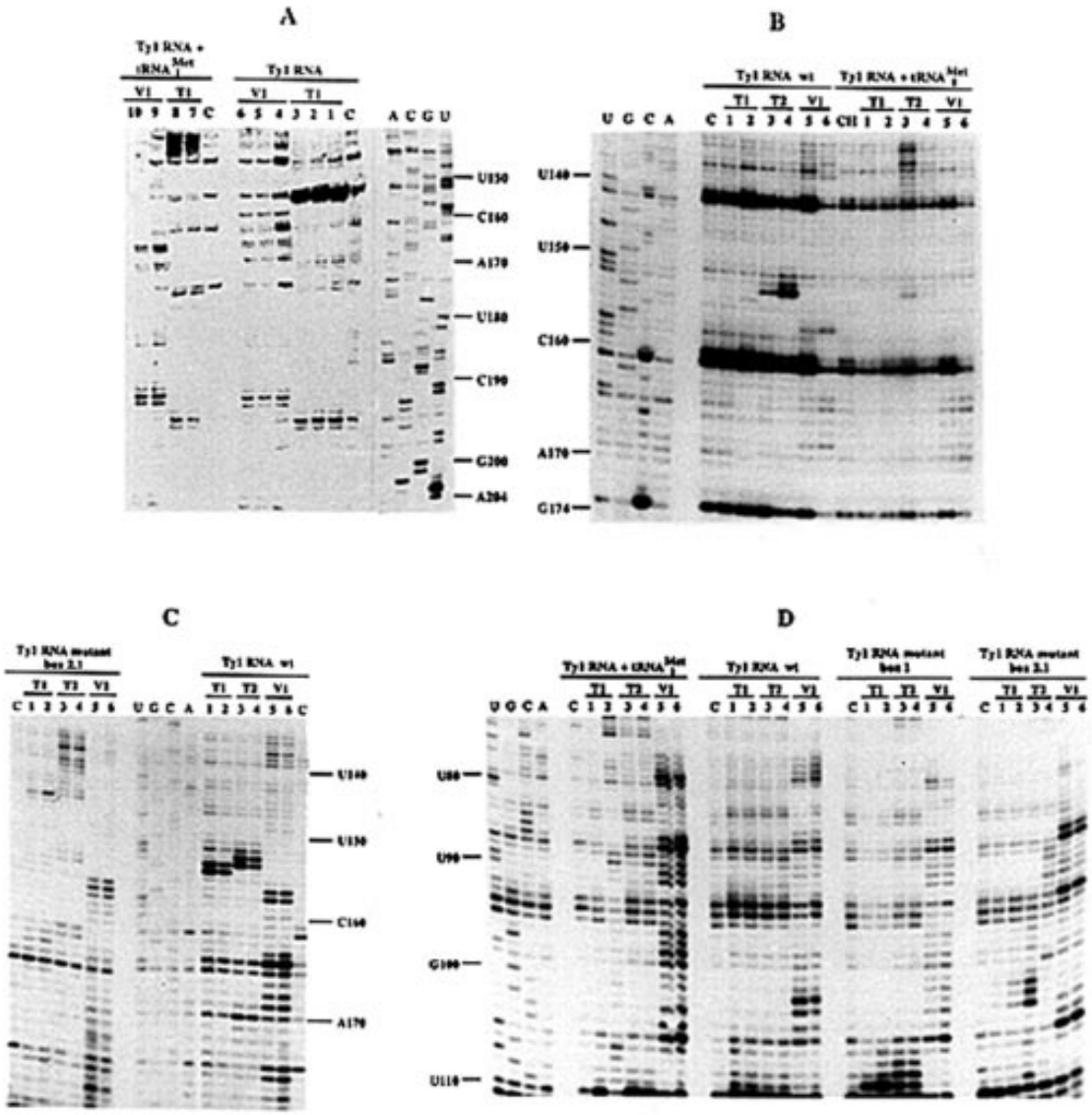

Figure 3. Enzymatic probing of Ty1 RNA. *In vitro*-transcribed Ty1 RNAs in the free form or hybridized to $tRNA_i^{Met}$ were treated with single-strand-specific nucleases (RNase T1 or RNase T2) or with double-strand-specific nuclease (RNase V1) for 5 (**A**, lanes 1, 4, 7 and 9, and **B**–**D**, lanes 1, 3 and 5), 10 (**A**, lanes 2, 5, 8 and 10, and **B**–**D**, lanes 2, 4 and 6) or 15 min (**A**, lanes 3 and 6). Lanes C and CH are controls without enzyme. Enzymatic cuts were detected by primer extension with reverse transcriptase. Dideoxy sequencing (A, C, G and U) of the same RNA fragment was used to identify the positions of the enzymatic cuts. Electrophoresis was on 8% polyacrylamide–7 M urea gels. Numbers at the side of the gels represent Ty1 RNA coordinates. In (A) the polyacrylamide gel was scanned with a Fuji Bio-Imager BAS2000.

not become single stranded in the complex, since RNase V1 cuts are still observed in this region (Fig. 3B). Intermolecular interactions between box 2.1 and residues 16–23 in the D stem and loop of $tRNA_i^{Met}$ account for the RNase V1 cuts at nt 162–167 of box 2.1 in the $tRNA_i^{Met}$–Ty1 RNA complex and are supported by probing of $tRNA_i^{Met}$ (see above).

Other structural rearrangements occuring upon formation of the $tRNA_i^{Met}$–Ty1 RNA complex are observed for nt 105–108, which are base paired to nt 134–137 in free Ty1 RNA but become reactive to single-strand-specific nucleases in the complex. Nucleotides 110–116 (box 0), which are partially single stranded in the free RNA, become susceptible to RNase V1 in the complex. By looking for alternative base pairing of nt 110–116 we have found that an intermolecular base pairing between these nucleotides and nt 59–65 in the 3′ part of the T stem and loop of $tRNA_i^{Met}$ is consistent with our experimental data.

The last reactivity changes observed are those occuring in the upstream and downstream regions of box 1. In the binary complex nt 131–141 upstream of box 1 are strongly cut by RNase T2 and nt 151–153 downstream of box 1 become less reactive to this enzyme. These structural rearrangements are consistent with an intermolecular base pairing between box 1 and nt 51–56 of the T stem and loop of $tRNA_i^{Met}$.

Mutations altering the nucleotide complementarity between the boxes and $tRNA_i^{Met}$ inhibit formation of the binary complex

The mutations indicated in Figure 1 were made in boxes 0, 1, 2.1 and 2.2 of Ty1 RNA or in the complementary sequence of $tRNA_i^{Met}$ to study the contribution of the boxes to formation of the binary complex. The two RNAs were annealed as described in Materials

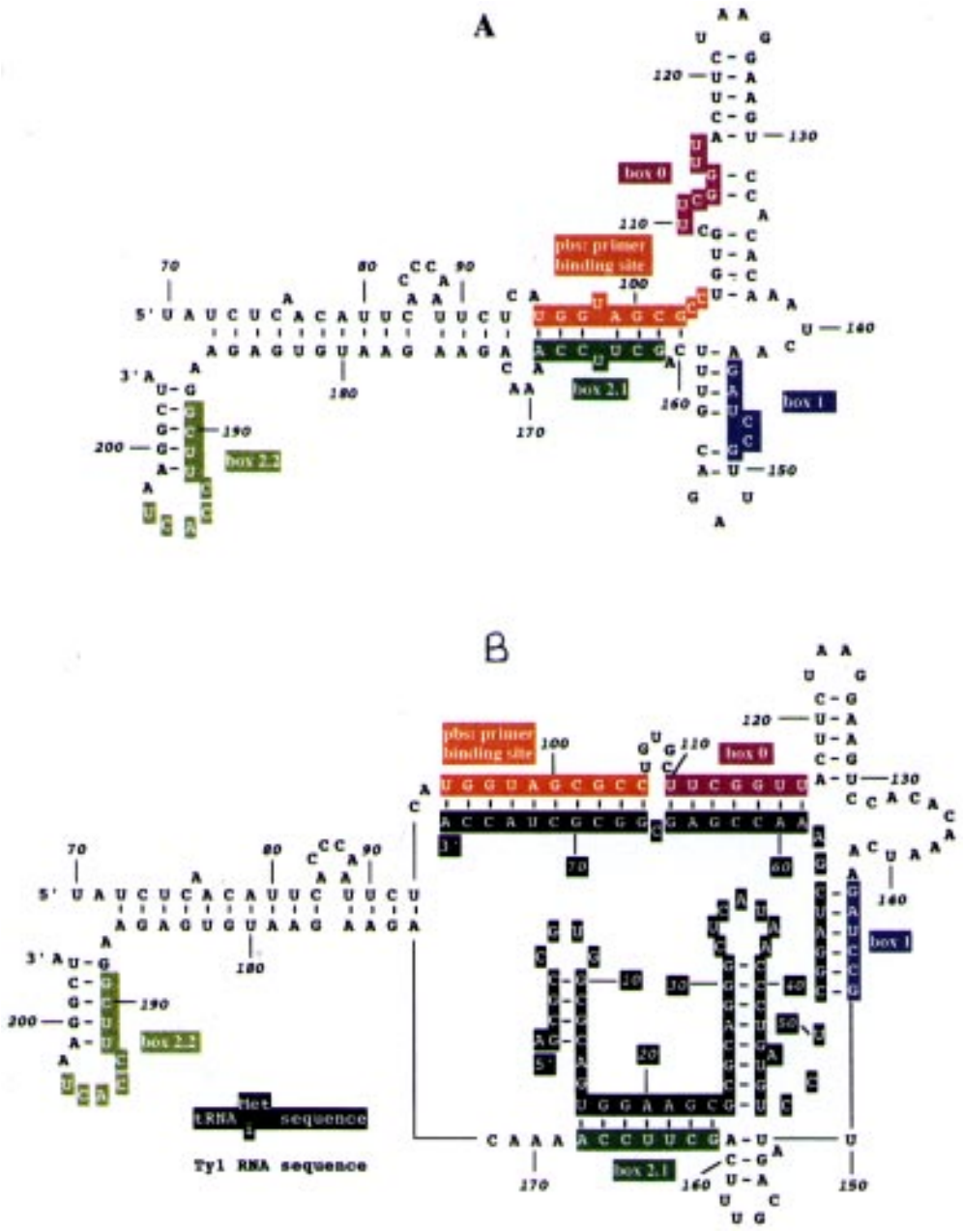

Figure 4. (**A**) Secondary structure model of Ty1 RNA in the free form. The PBS and the boxes are coloured as in Figure 1. (**B**) Secondary structure model of the $tRNA_i^{Met}$–Ty1 RNA complex. The tRNA sequence is drawn in black.

and Methods and analysed by agarose gel electrophoresis. We find that mutations in one of the boxes have only a slight effect on hybridization. The effect of altering the complementarity with box 1 or box 2.1 is shown in Figure 5. When two boxes are mutated hybridization is reduced by 50–70% provided one of the two boxes is box 2.1, whereas complex formation is abolished when box 1 and box 0 are mutated. This suggests that interaction with at least one of the two boxes 0 or 1 is absolutely required for efficient annealing of the two RNAs and that the role of box 2.1 is not as crucial for formation of the $tRNA_i^{Met}$–Ty1 RNA complex. The importance of Watson–Crick base pairing between the Ty1 RNA boxes and $tRNA_i^{Met}$ was confirmed by compensatory mutations which restored formation of the complex. In one example shown in Figure 5 annealing of $tRNA_i^{Met}$ bearing mutations in the T arm and loop (tRNA B01) is partially restored by compensatory mutations in box 1 of Ty1 RNA. These results clearly show that the regions of complementarity between $tRNA_i^{Met}$ and the boxes of Ty1 RNA are important for formation of extended interactions between the primer $tRNA_i^{Met}$ and Ty1 RNA and are necessary to stabilize the complex between the two RNAs.

Table 1. Reactivities of nt 70–204 of Ty1 RNA in the free form and in the binary complex with single-strand-specific nucleases (RNase T1 or RNase T2) or with a double-strand-specific nuclease (RNase V1)

Nucleotide	Free		Complex		Nucleotide	Free		Complex		Nucleotide	Free		Complex	
	ss	ds	ss	ds		ss	ds	ss	ds		ss	ds	ss	ds
U70	nd	nd			U115	+	–	–	±	C160	nd	nd		
A71	+	–			U116	–	–	–	±	A161	–	–	–	±
U72	–	–			A117	–	+			G162	–	–	–	±
C73	–	–			C118	–	+			C163	–	+		
U74	–	+			U119	–	+			U164	–	+		
C75	–	+			U120	–	–			U165	–	–	–	+
A76	–	+			C121	nd	nd			C166	–	–	–	±
A77	–	++			U122	nd	nd			C167	–	+		
C78	–	±			A123	+	–			A168	nd	nd		
A79	–	–	–	±	A124	+	–			A169	–	–		
U80	–	–	–	±	G125	–	–			A170	–	–		
U81	–	–	–	±	G126	nd	nd			A171	–	–		
C82	nd	nd			A127	–	±			C172	nd	nd		
A83	+	–			A128	–	–			A173	–	+		
C84	+	–			G129	–	–			G174	–	±		
C85	nd	nd			U130	–	+	–	–	A175	–	+		
C86	nd	nd			C131	–	+	–	–	A176	–	±		
A87	++	–			C132	–	+	+	–	G177	–	–		
A88	++	–			A133	–	+	+	–	A178	–	–		
U89	–	±	–	+	C134	–	+	+	–	A179	–	±		
U90	–	–	–	+	A135	–	±	+	–	U180	–	±		
C91	–	–			C136	–	+	+	–	G181	–	+		
U92	nd	nd			A137	–	+	+	–	U182	–	+		
C93	nd	nd			A138	–	–	+	–	G183	–	+		
A94	+	–			A139	–	–	+	–	A184	–	–		
U95	–	–	–	+	U140	nd	nd			G185	–	–		
G96	–	–	–	+	C141	nd	nd			A186	–	–		
G97	nd	nd	–	+	A142	–	–			A187	–	–		
U98	nd	nd	–	+	A143	–	–			G188	–	–		
A99	–	+			G144	–	–	–	+	G189	–	+		
G100	–	+			A145	–	–	–	–	C190	–	+		
C101	–	++			U146	–	–	–	+	U191	–	+		
G102	–	++			C147	–	–	–	+	U192	–	+		
C103	–	±	–	+	C148	–	–	–	+	C193	±	–		
C104	nd	nd	–	±	G149	–	–	–	±	C194	+	–		
U105	–	–			U150	nd	nd			A195	+	–		
G106	–	–	±	–	U151	++	–	–	–	C196	±	–		
U107	–	–	+	–	A152	++	–	+	–	U197	nd	nd		
G108	–	–	±	–	G153	++	–	–	–	A198	–	–		
C109	nd	nd			A154	–	–			A199	–	–		
U110	nd	nd			C155	–	–			G200	–	±		
U111	+	–	–	+	G156	–	+	–	–	G201	–	±		
C112	–	–	–	+	U157	nd	nd			C202	–	+		
G113	–	–			U158	–	+	–	–	U203	–	+		
G114	–	–	–	±	U159	nd	nd			A204	–	–		

The reactivities to the chemical probe DMS, specific for the A(N1) and C(N3) Watson–Crick positions, were also checked for nt 70–110. The data are consistent with several independent experiments. Four levels of reactivity were estimated: ++, highly reactive; +, moderately reactive; ±, marginally reactive; –, unreactive. nd, not determined, the reactivity could not be determined because a non-specific stop of the reverse transcriptase was observed in the control experiment. ss refers to reactivity to single-strand-specific probes (DMS, RNases T1 or/and RNase T2) and ds refers to reactivity to the double-strand-specific RNase V1. Reactivity of Ty1 RNA in the complex is indicated only when different from reactivity of Ty1 RNA in the free state. Some nucleotides localized in a compact tertiary structure are inaccessible to single-strand and double-strand nucleases. For these nucleotides we have no evidence that they are in double-stranded or single-stranded regions.

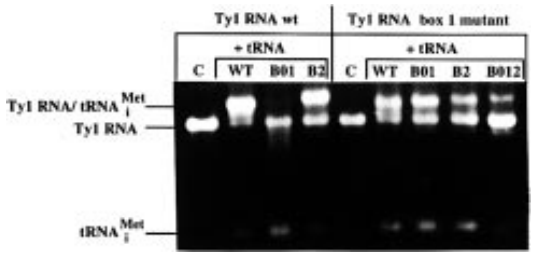

Figure 5. Effect of mutations in the boxes of Ty1 RNA and in $tRNA_i^{Met}$ on formation of the $tRNA_i^{Met}$–Ty1 RNA complex. Wild-type or box 1 mutant Ty1 RNA extending from nt 1 to 237 was annealed to wild-type or mutant $tRNA_i^{Met}$ and analysed on an agarose gel. tRNA WT, wild-type $tRNA_i^{Met}$; tRNA B1, $tRNA_i^{Met}$ mutated in the 5′ part of the T stem and loop; tRNA B2, $tRNA_i^{Met}$ mutated in the D loop; tRNA B01, $tRNA_i^{Met}$ mutated in the 5′ and 3′ parts of the T stem and loop; tRNA B012, $tRNA_i^{Met}$ mutated in the 5′ and 3′ parts of the T stem and loop and in the D loop.

CONCLUSION

The avalaible experimental data on the complex formed by retroviral genomic RNA and primer tRNA used to initiate reverse transcription suggest that extended contacts take place between the tRNA and the viral RNA (9–13,25). Here we present a structural probing study of the retrovirus-like yeast retrotransposon Ty1 RNA and of the $tRNA_i^{Met}$–Ty1 RNA complex showing that the interaction between the primer $tRNA_i^{Met}$ and Ty1 RNA is not limited to base pairing of the 10 nt at the 3′-end of $tRNA_i^{Met}$ with the PBS, but that three regions of Ty1 RNA, named boxes 0, 1 and 2.1, interact with the T and D stems and loops of $tRNA_i^{Met}$. The importance of these boxes is supported by a sequence comparison of Ty1 and Ty2 elements (26–29). Indeed, it is striking that the primary sequence of the boxes is strictly conserved in all elements and that the whole domain comprising the PBS and the boxes is very conserved (Fig. 6). This points to the importance of the primary sequence of this domain for promoting interaction between $tRNA_i^{Met}$ and Ty1 RNA and can be explained by the functional importance of the extended interactions between the two RNAs. The sequence of the structural domain spanning nt 72–94 and 169–203, which is not rearranged upon formation of the $tRNA_i^{Met}$–Ty1 RNA complex, is much less conserved. However, the overall secondary structure of this domain can be maintained in the different Ty1 and Ty2 elements by realignment of the primary sequences or compensatory substitutions, suggesting that the secondary structure of this region might also be functionally important.

The structural model of the $tRNA_i^{Met}$–Ty1 RNA complex is presented in Figure 4. For intermolecular base pairing between the two RNAs to occur the tertiary structure of $tRNA_i^{Met}$ must be unfolded and intramolecular base pairs of $tRNA_i^{Met}$ must be opened up. However, once the binary complex is formed the number of base paired $tRNA_i^{Met}$ residues is increased (54 residues are base paired in the complex compared with 42 in free $tRNA_i^{Met}$). Hence the $tRNA_i^{Met}$–Ty1 RNA complex obtained by hybridization at high temperature followed by renaturation adopts a stable thermodynamic conformation. We do not know which factors are involved in formation of the binary complex *in vivo*, but by analogy with retroviruses it is possible that Ty1 reverse transcriptase or the gag-like nucleocapsid protein TYA may help to unwind the two RNAs and promote formation of a thermodynamically stable conformation (30–32).

Until now it has not been possible to probe the structure of the $tRNA_i^{Met}$–Ty1 RNA initiation complex *in vivo*, but some experimental evidence has suggested that extended $tRNA_i^{Met}$–Ty1 RNA interactions play a role *in vivo*. As mentioned above, we have shown that mutations in the boxes result in defective transposition of the Ty1 element (7). In keeping with

Figure 6. Secondary structure model and phylogenetic analysis of Ty1 RNA. The sequence/structure of the Ty1 H3 element is shown and compared with 14 Ty elements. Out of 14 elements only nine have different sequences, as indicated in the Figure. Nucleotide changes occuring in these elements are indicated by arrows and the number of the particular element.

our results, Keeney *et al.* (33) have shown that mutations in the T arm and loop of $tRNA_i^{Met}$ have dramatic effects on transposition. For two double mutants in which A60/A54 were changed to C60/T54 or T60/C54 the transposition frequency was reduced to the background level. These mutations would disrupt base pairing between bases 60 and 54 of $tRNA_i^{Met}$ and U115 of box 1 and U146 of box 0 respectively. Weakening of the interaction between $tRNA_i^{Met}$ and Ty1 RNA could explain the transposition-defective phenotype of these mutants. This is consistent with our finding that annealing of $tRNA_i^{Met}$ with Ty1 RNA is abolished in a box 0/box 1 double mutant. By using hybrid *S.cerevisiae*/*Arabidopsis thaliana* and *S.cerevisiae*/*Schizosaccharomyces pombe* $tRNA_i^{Met}$ as primers Keeney *et al.* (33) found that additional determinants for Ty1 transposition are clustered in the D arm. Interestingly, in the structural model of the $tRNA_i^{Met}$–Ty1 RNA complex 2 nt of the 3′ part of the D arm of tRNA interact with box 2.1 of Ty1 RNA. These 2 nt are similar in the *S.cerevisiae* and *S.cerevisiae*/*A.thaliana* hybrid $tRNA_i^{Met}$, but differ in the *S.cerevisiae*/*S.pombe* tRNA hybrid, which is not able to support transposition. Examination of the Ty1 sequence shows that interaction between box 2.1 of Ty1 RNA and the D stem and loop of the *S.cerevisiae*/*S.pombe* $tRNA_i^{Met}$ hybrid would be destabilized or disrupted in a complex between this hybrid $tRNA_i^{Met}$ and Ty1 RNA. As a consequence, the secondary structure of the binary complex would be altered and we speculate that it would no longer be recognized by Ty1 reverse transcriptase, suggesting that the secondary structure of the $tRNA_i^{Met}$–Ty1 RNA complex plays an important role in the process of reverse transcription.

The detailed *in vitro* structural study reported here reveals that Ty1 RNA forms a specific secondary structure and that extended interactions occur between the primer $tRNA_i^{Met}$ and Ty1 RNA. The *in vitro* results and previously published *in vivo* studies (7,33) suggest that extended interactions between the primer tRNA and the genomic RNA would be necessary not only to stabilize the complex, but also to form a duplex secondary structure specifically recognized by reverse transcriptase. The fact that similar interactions and a specific secondary structure exist in several retroviral RNAs suggest that these viruses and the Ty1 retrotransposon probably utilize similar mechanisms to initiate reverse transcription. Experiments are now in progress to determine the steps of the priming process affected by the secondary structure of the $tRNA_i^{Met}$–Ty1 RNA complex.

ACKNOWLEDGEMENTS

We thank R. Marquet, F. Michel and P. Romby for helpful discussions, G. Keith for 3′-end-labelling of the $tRNA_i^{Met}$ and B. Senger for kindly providing the $ptRNA^{Met}$ plasmid. We are grateful to Professors G. Dirheimer and B. Ehresmann for their constant interest in our work. This work was supported in part by grants from the Association pour la Recherche contre le Cancer (ARC) and from the Ligue Nationale contre le Cancer, Comité Départemental du Haut-Rhin.

REFERENCES

1 Boeke,J.D. and Sandmeyer,S.B. (1991) In *The Molecular and Cellular Biology of the Yeast Saccharomyces: Genome Dynamics, Protein Synthesis, and Energetics.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, Vol. I, pp. 193–261.
2 Varmus,H. (1988) *Science*, **240**, 1427–1435.
3 Boeke,J.D., Garfinkel,D.J., Styles,C.A. and Fink,G.R (1985) *Cell*, **40**, 491–500.
4 Boeke,J.D., Eichinger,D., Castrillon,D. and Fink,G.R (1988) *Mol. Cell. Biol.*, **8**, 1432–1442.
5 Chapman,K.B., Byström,A.S. and Boeke,J.D. (1992) *Proc. Natl. Acad. Sci USA*, **89**, 3236–3240.
6 Pochart,P., Agoutin,B., Fix,C., Keith,G. and Heyman,T. (1993) *Nucleic Acids Res.*, **21**, 1517–1521.
7 Wilhelm,M., Wilhelm,F.X., Keith,G., Agoutin,B. and Heyman,T. (1994) *Nucleic Acids Res.*, **22**, 4560–4565.
8 Haseltine,W.A., Kleid,D.G., Panet,A., Rothenberg,E. and Baltimore,D. (1976) *J. Mol. Biol.*, **106**, 897–907.
9 Cordell,B., Swanstrom,R., Goodman,H.M. and Bishop,J.M. (1979) *J. Biol. Chem.*, **254**, 1866–1874.
10 Aiyar,A., Cobrinik,D., Ge,Z., Kung,H.-J. and Leis,J. (1992) *J. Virol.*, **66,** 2464–2472.
11 Aiyar,A., Cobrinik,D., Ge,Z. and Leis,J. (1994) *J. Virol.*, **68,** 611–618.
12 Leis,J., Aiyar,A. and Cobrinik,D. (1993) In *Reverse Transcriptase.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 33–47.
13 Isel,C., Ehresmann,C., Keith,G., Ehresmann,B. and Marquet,R. (1995) *J. Mol. Biol.*, **247,** 236–250.
14 Maniatis,T., Fritsch,E.F. and Sambroock,J. (1982) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
15 Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci USA*, **252**, 1682–1689.
16 Kunkel,T.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 488–492.
17 Despons,L., Senger,B., Fasiolo,F. and Walter,P. (1992) *J. Mol. Biol.*, **225,** 897–907.
18 Davanlo,P., Rosenberg,A.H., Dunn,J.J. and Studier,W.F. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2035–2039.
19 Peattie,D. and Gilbert,W. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 4679–4682.
20 Krol,A. and Carbon,P. (1989) *Methods Enzymol.*, **180**, 212–227.
21 Ehresmann,C., Baudin,F., Mougel,M., Romby,P., Ebel,J.P. and Ehresmann,B. (1987) *Nucleic Acids Res.*, **15,** 9109–9128.
22 Zucker,M. (1989) *Science*, **244**, 48–52.
23 Jaeger,J., Turner,D.H. and Zucker,M. (1990) *Methods Enzymol.*, **183**, 281–306.
24 Wimberly,B., Varani,G. and Tinoco,I. (1993) *Biochemistry*, **32**, 1078–1087.
25 Kohlstaedt,L.A. and Steitz,T.A. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 9652–9656.
26 Warmington,J.R., Waring,R.B., Newlon,C.S., Indge,K.J. and Oliver,S.G. (1985) *Nucleic Acids Res.*, **13,** 6679–6693.
27 Fulton,A.M., Rathjen,P.D., Kingsman,S.M. and Kingsman,A.J. (1985) *Nucleic Acids Res.*, **13,** 6679–6693.
28 Clare,J. and Farabaugh,P. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 2829–2833.
29 Stucka,R., Lochmüller,H. and Feldmann,H.(1989) *Nucleic Acids Res.*, **17,** 4993–5001.
30 Barat,C., Lullien,V., Schatz,O., Keith,G., Nugeyre,M.T., Grüninger-Leitch,F., Barr-Sinoussi,F., LeGrice,S.F.J. and Darlix,J.L. (1989) *EMBO J.*, **8**, 3279–3285.
31 Barat,C., Schatz,O., Le Grice,S. and Darlix,J.L. (1993) *J. Mol. Biol.*, **231**, 185–190.
32 Khan,R. and Giedroc,D.P. (1992) *J. Biol. Chem.*, **267**, 6689–6695.
33 Keeney,J.B., Chapman,K.B., Lauermann,V., Voytas,D.F., Aström,S.U., von Pawel-Rammingen,U., Byström,A. and Boeke,J.D. (1995) *Mol. Cell. Biol.*, **15**, 217–226.