R2 retrotransposition on assembled nucleosomes depends on the translational position of the target site

Junqiang Ye, Zungyoon Yang¹,
Jeffrey J. Hayes¹ and Thomas H. Eickbush²

University of Rochester, Department of Biology, Rochester, NY 14627
and ¹University of Rochester Medical Center, Department of
Biochemistry and Biophysics, Rochester, NY 14642, USA

²Corresponding author
e-mail: eick@mail.rochester.edu

R2 retrotransposons insert into the 28S rRNA genes of insects. Integration occurs by specific cleavage of the target site and utilization of the released DNA end to prime reverse transcription of the RNA transcript. Specificity of the protein to the target site is dependent upon nucleotide sequence recognition extending from 35 bp upstream to 15 bp downstream of the cleavage site. In this report, we show that sequence recognition and cleavage by the R2 protein can occur while the target site is assembled into nucleosomes. Reconstitution of DNA fragments containing the 28S gene sequence into a set of nucleosomes with different translational frames revealed that the R2 site adopted the same rotational orientation with respect to the histone octamer. Binding and cleavage by the R2 protein were most efficient when the upstream binding site for the R2 protein was near a nucleosome end. Interaction of the R2 protein with the nucleosome disrupted the histone:DNA contacts in the 50 bp region directly bound by R2, but did not modify the remainder of the nucleosome structure.

Keywords: accessibility/chromatin/nucleosome/
retrotransposons/rRNA genes

Introduction

R2 elements are among the best-studied non-long terminal repeat (non-LTR) retrotransposons (Eickbush, 2002). These elements, along with another distinct non-LTR retrotransposon (R1), insert specifically into the tandemly repeated 28S rRNA genes of most insects (Figure 1A). R2 encodes a single open reading frame (R2 protein), which has both reverse transcriptase and endonuclease activities. *In vitro* studies using purified protein have established the initial steps in the mechanism of R2 retrotransposition (Luan *et al.*, 1993). The protein first cleaves the target DNA at a specific location on one strand, and then cDNA is polymerized directly onto the newly exposed 3' hydroxyl group using the R2 RNA as template, a process termed target primed reverse transcription (TPRT). Second-strand cleavage of the target DNA occurs after TPRT. This TPRT mechanism serves as the likely model for the insertion of other non-LTR retrotransposons (Eickbush, 2002; Moran and Gilbert, 2002).

While our *in vitro* studies of R2 TPRT have been conducted with naked DNA, all eukaryotic DNA *in vivo* is assembled into chromatin (Wolffe, 1998; Kornberg and Lorch, 1999). The basic structural subunit of chromatin, the nucleosome core, is composed of 147 bp of DNA wrapped around an octamer of two each of the four core histone proteins with neighboring nucleosome cores connected by an ~50 bp linker DNA (Luger *et al.*, 1997). The tight association of core histones with genomic DNA and the highly bent DNA structure of DNA in the nucleosome impose a challenge for enzymes employed in DNA replication, transcription and recombination. Many of these enzymes are assisted in accessing DNA within chromatin by the activities of ATP-dependent chromatin remodeling complexes (Wu, 1997; Becker and Horz, 2002) and co-factors that carry out post-translational modification of core histone tails (Grunstein, 1997; Struhl, 1998). These activities alter nucleosome and chromatin structures such that the DNA is more accessible to DNA-dependent activities.

The chromatin structure of the rRNA genes from many organisms has been extensively investigated (Lohr, 1983; Palen and Cech, 1984; Sogo *et al.*, 1984). Many early studies indicated that while the upstream promoter/spacer region is organized into unusual structures, the transcribed regions of these genes are organized into nucleosomal arrays. In the only insect studied, *Drosophila melanogaster*, the nucleosomes covering the rRNA genes appear to be in multiple translational phases both in early development and in tissue culture cells (Kolchinskii *et al.*, 1984; Udvardy *et al.*, 1984). More detailed recent studies by Sogo and co-workers have suggested that the rDNA units of higher eukaryotes can be divided into two chromatin structures (Conconi *et al.*, 1989, 1992; Stancheva *et al.*, 1997). The transcriptionally inactive rDNA units are packaged into typical nucleosome arrays, while transcriptionally active rDNA units are nucleosome free. Such findings correlate with electron micrographs of actively transcribing *D.melanogaster* rDNA units, in which each rDNA unit is either in an extensively transcribed or completely inactive state (Jamrich and Miller, 1984). Presumably, the R1 and/or R2 inserted units fall into the latter category as few long (i.e. inserted) units were observed to be transcribed, and only low levels of R1 or R2 transcripts are detected on northern blots (Long and Dawid, 1979; Kidd and Glover, 1981).

Given these results, it would be predicted that new R2 insertions would predominantly occur in uninserted rDNA units. However, when R2 protein was injected into *D.melanogaster* embryos to monitor R2 TPRT *in vivo* (Eickbush *et al.*, 2000), 37% of the R2 TPRT reactions occurred into rDNA units already containing an R1 insertion (45% insertion into R1-containing units was expected if the insertions were entirely random).

Fig. 1. The R2 target site and DNA fragments used to reconstitute nucleosomes. (A) Diagram of a typical rRNA gene unit of insects. The 18S, 5.8S and 28S rRNA genes are indicated by the black boxes. The 3.6 kb R2 element encodes a single open reading frame with a centrally located reverse transcriptase (RT) domain, C-terminal endonuclease (ENDO) domain and an N-terminal DNA binding (DB) domain. Most insects also contain a 5.1 kb R1 non-LTR retrotransposon which inserts 74 bp downstream of the R2 insertion site. Insects contain hundreds of these tandemly repeated rDNA units, a variable fraction of which are inserted with R1 and/or R2 elements (Jakubczak *et al.*, 1991). (B) Location of the different 147 bp DNA fragments used to reconstitute nucleosomes. Black and gray boxes represent the 35 bp upstream and 15 bp downstream binding sites of R2, respectively. Primers used to PCR amplify these fragments are given in Table I.

Therefore, either two different chromatin structures do not exist for *D.melanogaster* rDNA units *in vivo*, or the ability of the R2 protein to conduct TPRT is not dramatically affected by different chromatin conformations. In this report, we attempt to directly address the question of how chromatin structure affects the TPRT reaction. Nucleosomes were reconstituted using purified histones and five overlapping 28S rRNA gene fragments in which the R2 target site was placed in different translational positions within the nucleosome. We found that the ability of the R2 protein to access its target site was highly dependent upon the translational position of the nucleosomes; however, all nucleosomes were able, at some level, to support a TPRT reaction.

Results

In vitro nucleosome reconstitution

To determine whether the R2 protein can access its 28S rRNA gene target site when that site is assembled into a nucleosome, we reconstituted nucleosomes using five overlapping 147 bp fragments containing the R2 target sequence (Figure 1B). Based on deletion studies of the target sequence, the R2 protein has an asymmetric binding site (Xiong and Eickbush, 1988). Specific binding and cleavage require the 35 bp DNA sequence upstream of the insertion site (upstream binding site) to 15 bp downstream of this site (downstream binding site). The R2 protein can be shown to interact with DNA at various positions throughout this 50 bp region (S.Christensen and

Fig. 2. Sucrose gradient purification of reconstituted nucleosomes. The results shown are bottom strand-labeled 39 END nucleosomes; however, all nucleosome reconstitutions displayed similar profiles on sucrose gradients. After centrifugation, fractions were collected (lane 1 represents the bottom of the gradient) and the products analyzed in 0.7% agarose gels. Lane D, free DNA control; lane N, total nucleosomes from the reconstitution. Lanes 4 and 5 (labeled with triangles) were used directly for cleavage and footprinting experiments.

T.Eickbush, in preparation). All DNA fragments used for nucleosome reconstitution contained complete upstream and downstream binding sites with the 50 bp recognition sequence situated at different positions within the 147 bp fragments. Each DNA fragment or corresponding nucleosome is denoted by the distance of the initial R2 nick site on the bottom strand relative to the upstream end of the fragment. Thus, 39 END has the upstream binding site closest to the upstream end of the DNA fragment, while fragment 130 END has the downstream binding site closest to the downstream end of the fragment (Figure 1B).

The R2 protein used in this study was derived from the *Bombyx mori* R2 element (R2Bm) open reading frame expressed in *Escherichia coli* (Xiong and Eickbush, 1988). Protein purification and enzyme assays were conducted under conditions similar to previous assays of R2 cleavage and TPRT reactions (Luan *et al.*, 1993). The DNA substrates used for the reactions were derived from the *D.melanogaster* rDNA sequence to enable direct comparison to previous and future *in vivo* experiments (Eickbush *et al.*, 2000). The 28S rRNA gene sequence surrounding the R2 target site is highly conserved in all arthropods, with only 7 bp substitutions between the *D.melanogaster* and *B.mori* sequences in the 238 bp region used in this study. Only one of the seven substitutions exists in the R2 binding region (see Figure 3C), and no difference has been detected in the ability of the *B.mori* R2 protein to bind or conduct the TPRT reaction with the 28S gene from these two species.

DNA fragments were generated by PCR amplification with one primer ^{32}P -end labeled in order that each DNA was labeled at the 5' end of only one strand. Nucleosomes were reconstituted by salt dialysis with the reconstitution efficiency checked on 0.7% agarose gels (Hayes and Lee, 1997). In a typical reaction, ~80% of the labeled DNA fragments were assembled into nucleosomes. Reconstituted nucleosomes were then separated from free DNA by centrifugation through 5–20% sucrose gradients, as shown in Figure 2. Because sucrose does not affect R2 activity, all binding, cleavage and TPRT experiments were carried out with nucleosomes directly after 3- to 5-fold dilution of fractions from the sucrose gradient.

Mapping of reconstituted nucleosomes containing the R2 target site

Both DNase I and hydroxyl radical footprinting were used to map the various reconstituted nucleosomes. Figure 3A compares the DNase I and hydroxyl radical cleavage patterns of 49 END nucleosomes and free 49 END DNA in

Fig. 3. DNase I and hydroxyl radical mapping of reconstituted nucleosomes. (A) Mapping of 49 END DNA and nucleosome. Lane 1, naked DNA; lane 2, DNA digested by DNase I; lane 3, reconstituted nucleosome digested by DNase I; lane 4, naked DNA cleaved by hydroxyl radical; lane 5, nucleosome cleaved by hydroxyl radical; lane 6, DNA cleaved by R2 to reveal the insertion site. Procedures for DNase I and hydroxyl radical mapping are described in the Materials and methods. (B) Hydroxyl radical mapping of bottom strand-labeled 39 END, 49 END, 74 END, 79 END and 130 END nucleosomes. For each fragment, free DNA (D) and/or nucleosome (N) substrates were exposed to hydroxyl radical and the nicked products resolved by 8% PAGE alongside a DNA sequence G track (G) of the same fragment. Arrows indicate the site of R2 cleavage. (C) Schematic representation of the rotational phase of the DNA around the R2 target site while the DNA is assembled into nucleosomes. The height of each bar above or below each nucleotide indicates the accessibility of hydroxyl radical to that site. The cleavage sites of R2 and *Bst98*I on both strands are indicated by arrows. Two positions of difference between the *D.melanogaster* and *B.mori* 28S rRNA genes are circled. In *B.mori*, the top strand contains a C at position +43 and an A at position +32.

which the bottom strand was radioactively end-labeled (Figure 1B). Free DNA cleaved by DNase I or hydroxyl radical yielded a roughly uniform cleavage at all positions (lanes 2 and 4). Cleavage of the nucleosome, on the other hand, showed a characteristic 10 bp repeating pattern (lanes 3 and 5), reflecting the periodic interaction of the DNA backbone with the histone surface. The location of the R2 cleavage site on the bottom strand of 49 END DNA is shown in lane 6.

The 10 bp periodic cleavage pattern establishes the rotational phase of the DNA bound to the histone octamer. As shown in Figure 3B, hydroxyl radical mapping revealed that nucleosomes reconstituted with 39 END, 49 END, 74 END and 79 END DNA all showed a similar

rotational orientation. This can be seen by comparison to the G cleavage lane in each panel or by the arrow next to each panel, which indicates the location of the R2 cleavage site on the bottom DNA strand. In each nucleosome, the R2 cleavage site is positioned between strong and weak hydroxyl radical cleavages. The ability of the 28S gene segment to assume a preferred rotational phase around the histone octamer is best demonstrated in the comparison between the 74 END and 79 END nucleosomes. Even when the sequence frame encompassed by these two DNA fragments is shifted by 5 bp, they still bear the same rotational phase with respect to the target site. In the case of the 130 END nucleosome, there appeared to be several preferred rotational phases for the nucleosomes assembled

with this DNA fragment, such that the individual cleavage patterns add ‘destructively’, resulting in no discernible periodic cleavage pattern.

The combined hydroxyl radical mapping data of the top and bottom strand labeled nucleosomes (top strand mapping data not shown) were quantitated and are summarized in Figure 3C. The R2 cleavage sites on both strands lie at the junction between the strongest and weakest points of hydroxyl radical cleavage. Based on the inability of the 130 END DNA to assume this same rotational phase, the DNA sequences that direct this preferred positioning on the surface of the histone octamer extend further downstream than upstream of the R2 cleavage site.

R2 cleavage of nucleosomes

To determine how efficiently the R2 protein cleaves the target site when assembled into nucleosomes, competition experiments were conducted with free DNA and assembled nucleosomes (Figure 4A). In each reaction, the naked DNA fragments included in the reaction were different from those used for nucleosome reconstitution so that cleavage of the two substrates could be independently scored by the different lengths of the DNA products. All cleavage reactions were conducted in the presence of RNA corresponding to the 3′ untranslated region (UTR) of the R2 element, because previous experiments have shown that the presence of this RNA improves the specificity of R2 cleavage (Yang and Eickbush, 1998).

In the competition between 74 END nucleosomes and 49 END free DNA, we found that R2 digestion of the nucleosome DNA occurred at a considerably lower level than that of free DNA (Figure 4B). Based on the extrapolated initial rates of product formation, inhibition of R2 cleavage by the nucleosome was ~50-fold. In the competition between 39 END nucleosomes and 49 END DNA, inhibition of cleavage by nucleosome formation was less extensive (Figure 4C). The initial rate of digestion of 39 END nucleosome was only 6-fold slower than the rate of digestion of free DNA, with the final accumulation of cleaved products ~40% that of free DNA.

To eliminate the possibility that the nucleosome cleavage conducted by the R2 protein was the result of partial dissociation of the nucleosomes or heterogeneity in translational positioning, we also tested the ability of the restriction enzyme *Bst98I* to cleave the nucleosomes. The activity of various restriction enzymes has been shown to be dramatically inhibited by nucleosome structure (Polach and Widom, 1995; Widom, 1998). *Bst98I* cleaves the 28S gene sequence 2 bp from the R2 cleavage site on the bottom strand and 4 bp from the R2 site on the top strand (see Figure 3C). As shown in Figure 4D and E, digestion of free DNA by *Bst98I* was completed within 2 min, while digestion of either the 74 END or 39 END nucleosomes was minimal. The inhibition of the reaction rate of *Bst98I* cleavage by nucleosome structure was calculated to be >2000-fold, which is within the range found for nucleosome inhibition of many restriction enzymes (Polach and Widom, 1995). The near complete inhibition of our purified nucleosome preparations to *Bst98I* cleavage under the same conditions as the R2 digestion suggests that the R2 protein is able to recognize and cleave the target site when it is assembled into a nucleosome.

Fig. 4. Accessibility of the R2 cleavage site as measured by R2 or *Bst98I* cleavage. (A) Diagram of the experimental approach using 74 END nucleosome and 39 END DNA as an example. Equal amounts of DNA as reconstituted nucleosome and naked DNA with the cleavage site in a different position were digested with the R2 protein or the restriction enzyme *Bst98I* for increasing periods of time. See Figure 3C for the location of the R2 and *Bst98I* cleavage sites. The products were then resolved by denaturing 8% PAGE. (B) R2 digestion of 74 END nucleosomes and 49 END DNA. (C) R2 digestion of 39 END nucleosomes and 49 END DNA. (D) *Bst98I* digestion of 74 END nucleosomes and 39 END DNA. (E) *Bst98I* digestion of 39 END nucleosomes and 49 END DNA. In each panel, the top figure shows the denaturing gel, while the bottom figure shows the amount of cleavage determined by ImageQuant.

Competition experiments were conducted next to determine the relative reactivity of the various nucleosomes to R2. Equal amounts of labeled 39 END, 74 END and 130 END nucleosomes were mixed and digested with the R2 protein in the presence of 3′ UTR RNA. As shown in Figure 5A, the 39 END nucleosome is more readily cleaved by the R2 protein than the other two nucleosomes. Two controls were conducted for this reaction. First, again as a control for the stability of the nucleosomes, additional aliquots of the nucleosomes were digested with *Bst98I* under the same conditions of the R2 digests (Figure 5A). *Bst98I* was able to cleave the 39 END nucleosome somewhat better than the 74 END nucleosome, but no better than the 130 END nucleosome. The second control was to determine whether the higher rate of R2 cleavage of the 39 END nucleosome was simply because the cleavage site was near a DNA end. As shown in Figure 5B, in a

Fig. 5. Competition experiments in which the target site has been placed at different translational positions on the nucleosome. (A) Digestion of 39 END, 74 END and 130 END nucleosomes mixtures by R2 and *Bst98I*. (B) Digestion of 39 END, 74 END and 130 END DNA mixtures by R2 and *Bst98I*. In order to better observe the initial rates of cleavage, R2 protein and *Bst98I* concentrations used to digest the naked DNA in (B) were only 1/20 and 1/100, respectively, of that used to digest nucleosomes in (A). Symbols for each substrate digested with either R2 or *Bst98I* are illustrated in the panel to the right of (A).

competition involving free DNA, neither the R2 protein nor *Bst98I* exhibited a significant preference for any of the three DNA substrates. Note that the concentrations of *Bst98I* and R2 in the free DNA reactions (Figure 5B) were considerably lower than with the nucleosome reactions (Figure 5A) in order to slow the rate of product accumulation.

We conclude that the R2 protein can readily cleave its DNA target site when reconstituted into a nucleosome if the upstream binding site for the protein is near the upstream end of the nucleosome. Nucleosomes are not good substrates for R2 cleavage when the target site is near the middle of the nucleosome or near the downstream end of the nucleosome.

R2 protein binding to nucleosomes

The high levels of 39 END nucleosome cleavage by R2 suggested that we should be able to isolate complexes of the R2 protein bound to nucleosomes. The R2 protein remains bound to the substrate after cleavage of the bottom

DNA strand (Yang and Eickbush, 1998). This complex remains stable until second-strand cleavage, which occurs only slowly when TPRT is inhibited by the absence of dNTPs. Therefore, we examined the R2 binding to nucleosomes under standard DNA cleavage conditions by gel mobility shift assays. In each experiment, two concentrations of R2 protein were pre-incubated with nucleosomes or naked DNA substrates for 20 min at 37°C before separation of the products on native polyacrylamide gels. Binding was conducted in both the presence and absence of 3' UTR RNA. As shown in Figure 6A, in the absence of RNA, both concentrations of R2 protein bound to the DNA with most of the DNA unable to enter the gel (lanes 2 and 3). Greater than 80% of the DNA substrate is nicked under these conditions (data not shown) and a small fraction of DNA is cleaved on both strands, giving rise to the shorter cleavage product, which is released from the R2 protein. At these same concentrations of R2 protein, lower levels of the 39 END nucleosomes are bound (Figure 6A, lanes 5 and 6). As with the free DNA, the majority of the bound nucleosome remained within the wells of the gel. We note that the reconstituted nucleosomes migrated as two bands on these native gels. The two bands probably represent two translational positions of the DNA in the nucleosomes separated by 10 bp.

Interestingly, binding of the R2 protein to the 39 END nucleosome was increased in the presence of the 3' UTR RNA (compare lanes 5 and 6 with 8 and 9). Moreover, migration of the bound DNA and nucleosomes through the gel was also changed. While some material remained within the wells, the predominant shifted band now enters the gel. Because the shifted nucleosome band migrates faster than the predominant shifted DNA band (compare lanes 8 and 9 and 11 and 12), we hypothesize that the R2-nucleosome complex has a more compact structure than the structure formed by the R2 protein on naked DNA. The nature of the broad range of shifted bands formed with the naked DNA substrate will be the subject of a separate study (S.Christensen and T.Eickbush, in preparation).

Next, we compared in these gel shift assays the ability of R2 to bind the five nucleosomes with the R2 target site in different positions. The results with the 39 END, 49 END and 74 END nucleosomes are shown in Figure 5B. While the R2 protein bound the 49 END nucleosomes somewhat less efficiently than the 39 END nucleosome (lanes 2 and 3, and 5 and 6 in Figure 6B), it was unable to bind a significant fraction of the 74 END nucleosome (lanes 8 and 9). The R2 protein was also unable to shift a significant fraction of the 79 END and 130 END nucleosomes (data not shown). These results suggest that the R2 protein is readily able to bind the nucleosome when its upstream binding site is near a nucleosome end. That binding is dramatically reduced when the target site is moved towards the middle or to the opposite end of the nucleosome. These findings are in agreement with the cleavage experiments in Figure 5.

Footprint of the R2-bound nucleosome

To further investigate the interaction between the R2 protein and its target site assembled into a nucleosome, we footprinted the complex using DNase I and hydroxyl radical mapping. Because cleavage of the target site on the

Fig. 6. R2 protein binding to DNA and nucleosomes as monitored by gel mobility shift assays. **(A)** 39 END DNA (lanes 1–3 and 10–12) and nucleosome (lanes 4–9) were incubated with 10 and 20 ng of R2 protein (for DNA) or 20 and 40 ng of the R2 protein (for nucleosomes) in the absence (lanes 1–6) or presence (lanes 7–12) of R2 3' UTR RNA. Autoradiograms of the complexes resolved by 5% native PAGE are shown. Both the nucleosomes and free DNA in these experiments were derived from sucrose gradients as in Figure 2, so that they both contained the same non-specific carrier DNA used in the nucleosome reconstitution. The nature of the shifted complexes is illustrated on the left side of the gel. A low level of double-stranded (ds) DNA cleavage as well as DNA–R2–RNA complex is seen in the absence of R2 RNA because of co-purification of a small amount of RNA with the R2 protein. **(B)** Gel mobility shift assay for R2 with 39 END, 49 END and 74 END nucleosomes in the presence of R2 3' UTR RNA. All reaction conditions are as in (A).

bottom strand eliminates our ability to footprint the DNA upstream of the cleavage site, we also used an endonuclease mutant (*Endo*[−]) R2 protein for these studies. This mutant R2 protein has a single amino acid substitution (Asp to Ala) in the active site of the endonuclease domain (Yang *et al.*, 1999). All other activities of the mutant enzyme, including DNA binding and reverse transcription, are similar to the wild-type enzyme. Similar complexes were characterized with both the 39 END and 49 END

nucleosomes; however, only the data derived from the 49 END nucleosome are shown here as the greater length of the DNA upstream of the target site allowed better resolution of cleavage bands in the upstream binding site from the full-length DNA. The nucleosomes, labeled on their bottom strand, were incubated with the R2 protein for 20 min at 37°C. The nucleosomes were then lightly treated with either DNase I or hydroxyl radicals and loaded onto native PAGE gel (Figure 7, top panels). Eluted DNA bands from this native gel were then separated on an 8% denaturing gel (Figure 7, bottom panels). All numbered bands from the native gels correspond to the individual lanes in the denaturing gels.

In the case of the hydroxyl radical mapping (Figure 7A), the DNA downstream of the target site retained the typical 10 bp repeated structure of a nucleosome when the R2 protein had bound and cleaved, or simply bound, the target sequence (compare lanes 2 and 3, and 4 and 5). However, the DNA regions immediately surrounding the target site in the *Endo*[−] mutant can be seen to be protected from hydroxyl radical cleavage (lane 4). The 10 bp repeating structure also appeared to be slightly shifted away from the cleavage site for two turns of the DNA helix downstream (~20 bp) and at least three turns of the helix upstream (~30 bp) of the cleavage site (lanes 2 and 4). Thus, the majority of the structure of the nucleosome appears to be unaffected by R2 protein binding and cleavage, except in the region actively involved in R2 protein binding.

As shown in Figure 7B, bound R2 protein had a greater impact on the DNase I footprint of the nucleosome. The DNA from 20 bp downstream to 30 bp upstream of the cleavage site is extensively protected from DNase I cleavage (compare lanes 6 and 7, and 8 and 9). Unlike the hydroxyl radical footprint, there are also modifications of the DNase I footprint 50–90 bp downstream of the cleavage site. For comparison, we also show the DNase I footprint of the R2 protein on free DNA (lanes 11 and 13). While the DNase I protection upstream of the cleavage site is similar on nucleosomes and free DNA, protection is not as great on the DNA substrate in the region downstream of the cleavage site (compare lanes 7 and 9, and 11 and 13). Most significantly, there is no protection of the naked DNA 50–90 bp downstream of the cleavage site. We suggest that this protection well downstream of the cleavage site corresponds to the large R2 protein (at least two subunits of a 120 kDa protein) partially interfering with DNase I cleavage of the DNA helix, as it makes its second superhelical turn within the same nucleosome. Because the smaller hydroxyl radicals are able to obtain similar access to this region as in an unbound nucleosome (compare lanes 1 and 2), this downstream protection does not appear to involve direct contact of the R2 protein with this downstream DNA.

We conclude from the combined DNase I and hydroxyl radical mapping of the R2 protein–nucleosome complexes that, after R2 binding and lower strand cleavage, most of the DNA in the nucleosome remains wrapped about the histone core. There is, however, disruption of the DNA path on the surface of the histone core in the region from 20 bp downstream to 30 bp upstream of the cleavage site. A more detailed analysis of this region of the nucleosome will be the subject of a subsequent report.

Fig. 7. DNase I and hydroxyl radical footprinting of the R2 protein bound to the 49 END nucleosome. (A) Hydroxyl radical footprint. Top: gel mobility shift assay. Both wild-type (wt) R2 and an endonuclease mutant (Endo⁻) were used to shift nucleosomes (see Yang *et al.*, 1999). After pre-incubating the R2 protein with its substrate, the complexes were treated with hydroxyl radical and resolved on 5% native PAGE gels. Bottom: eluted bands from the native gel were analyzed in 8% denaturing gels. The lane number corresponds to the eluted band from the native gel. D, naked DNA; G, DNA sequence G tract. The R2 cleavage site is labeled by an arrow. (B) The same sets of experiments as in (A), except the complex was digested with DNase I. Altered DNase I sensitivity of the DNA compared with the nucleosome well downstream of the cleavage site is indicated by circles to the left of the denaturing gel.

R2 can conduct TPRT on a nucleosome substrate

Finally, all five nucleosomes were tested for their ability to support a complete TPRT reaction (i.e. utilization of the cleaved lower DNA strand as a primer for reverse transcription of the 3' UTR RNA). The TPRT reaction can be conducted by simply adding dNTPs to our cleavage reactions. All nucleosomes were originally found to support a low level of the TPRT reaction. However, because in our *in vitro* reactions only a small fraction of the cleaved DNA is actually used as primer for TPRT, it was necessary to rule out the possibility that the low levels of TPRT products observed were derived from low levels of DNA, or disrupted nucleosomes, contaminating our nucleosome preparations. We therefore conducted the experiment illustrated in Figure 8A. The nucleosome preparations under the conditions of our TPRT reaction were pre-digested with *Bst*98I to cleave all target sites not

assembled into nucleosomes. The R2 protein was then added to the mixture to allow cleavage and TPRT. To demonstrate the efficiency with which *Bst*98I pre-cleavage of the target site eliminated TPRT, 39 END and 74 END DNA were also subjected to this experiment.

As shown in Figure 8B and C, when either 39 END or 74 END DNA was pre-digested with *Bst*98I, the ability of this substrate to support TPRT was completely eliminated (compare lanes 3 and 5 in each panel). However, pre-digestion of 39 END or 74 END nucleosomes with *Bst*98I only slightly inhibited the ability of these nucleosomes to serve as substrates for TPRT (compare lanes 8 and 10 of each panel). Thus, the R2 protein can conduct TPRT on a nucleosome substrate. Indeed, low levels of TPRT can even be observed on the 74 END nucleosome, where it was difficult to observe significant levels of protein binding (Figure 6B).

Fig. 8. A complete TPRT reaction can occur on nucleosome substrates. (A) Diagram of the experimental approach. Purified nucleosomes or naked DNA are pre-digested with *Bst98I* under conditions of a TPRT reaction for 10 min to cleave all free DNA. R2 protein is then added and the reaction is allowed to proceed for another 50 min. (B) 39 END DNA and 39 END nucleosomes. DNA (lanes 1–5) and 39 END nucleosome (lanes 6–10). Digestion by *Bst98I* alone (lanes 2 and 7), R2 alone (lanes 3 and 8), both *Bst98I* and R2 (lanes 4 and 9) or *Bst98I* for 10 min and then R2 for 50 min (lanes 5 and 10). After incubation, DNA was extracted and resolved on an 8% denaturing gel. (C) The same experiments as in (B), but using 74 END DNA and 74 END nucleosomes.

These experiments also revealed information about the efficiency with which the cleaved target site is used for TPRT. In the case of the free 39 END and 74 END DNA, only ~5% of the DNA that is cleaved is further used as substrate in the TPRT reaction. In the case of the 39 END nucleosome, again ~5% of the cleaved DNA is used in a TPRT reaction. Surprisingly, with the 74 END nucleosome, nearly 15% of the cleaved DNA was used for TPRT. Thus, it appears that the limiting step in the reaction of the R2 protein with a nucleosome is the binding of the protein to the target site. Once the protein is bound to the nucleosome, cleavage of the lower strand of DNA and TPRT proceeds with at least the same efficiency as on free DNA.

Discussion

The DNA sequences of genes can contain intrinsic structures that energetically favor a preferred translational and rotational phase when assembled with core histones into nucleosomes. This positioning has been extensively documented for many sequences (Simpson, 1991) and can have profound affects on the activity of *trans*-acting factors *in vivo* (Simpson, 1990). In this report we have shown that the 28S rRNA gene sequence surrounding the R2 target site also has an intrinsic preference for rotational positioning. The 39, 49, 74 and 79 END DNA fragments were all assembled into nucleosomes in which the R2 target site adopts the same rotational orientation with respect to the histone octamer surface. In this preferred orientation, the R2 cleavage sites on both DNA strands are located at the junction between regions of high and low hydroxyl radical cleavage, suggesting they are not fully accessible to the R2 protein on the nucleosome surface. The finding that the 130 END DNA fragment was not

assembled into a homogeneous rotational phase indicates that the region of intrinsic DNA structure may extend further downstream of the target site than upstream of this site.

Beyond rotational positioning, our results indicate that the translational position of the nucleosome on the 28S gene sequence is a key factor in the ability of the R2 protein to gain access to its target site. All evidence suggests that the R2 protein only interacts with a region extending from 35 bp upstream of the insertion site to 15 bp downstream of this site. DNA replacement studies (Xiong and Eickbush, 1988), protein footprinting studies using DNase I, hydroxyl radical and missing nucleoside (S.Christensen and T.Eickbush, in preparation; see also Figure 7B for the DNase I pattern), and the total elimination of flanking DNA from the DNA substrate (Figure 5B) all suggest that only the region from –35 to +15 is involved in DNA binding, cleavage and TPRT.

This 50 bp target site is most accessible to the R2 protein when the upstream binding site is located near one end of a nucleosome. Access is significantly reduced when the target site is moved to the center or to the other end of the nucleosome. However, R2 binding and cleavage occur throughout the nucleosome to extents greater than that predicted by a simple histone competition model for binding nucleosomal DNA (Polach and Widom, 1995). The greater accessibility of the 39 END and 49 END nucleosomes can not be explained by aberrant nucleosomes, or nucleosomes with modified structures. Both methods used to study nucleosome structure, hydroxyl radical and restriction enzyme sensitivity, suggest that the 28S sequence fragments used in this study are assembled into typical nucleosomes (Figures 3, 4, 5 and 7).

The interaction of the R2 protein with its target site on assembled nucleosomes can be compared with the inter-

action of the glucocorticoid receptor (GR) with its 6 bp binding site. So long as the major groove of these sequences is facing away from the nucleosome surface, then GR can readily bind this sequence when the nucleosome is in various translational positions (Perlmann, 1992). However, if the 6 bp recognition sequence is in a rotational position in which it is facing towards the nucleosome surface, then GR can not bind if the sequence is near the dyad axis of the nucleosome, but can bind if the sequence is nearer the ends of the nucleosome (Li and Wrangé, 1995). The ability of GR to gain access to a hidden recognition sequence at the ends of nucleosomes was suggested to be a result of the less tight association of core histone and DNA at these positions (Simpson, 1991).

We suggest that in our reconstituted, rotationally positioned nucleosome, not all aspects of the large R2 recognition sequence are available to the R2 protein, i.e. facing outward from the nucleosome. When the target site is near the end of the nucleosome, however, the weaker binding to the histones allows the R2 protein to gain access to its entire recognition sequence. The 49 END nucleosome allows efficient R2 protein binding and cleavage, while the 130 END nucleosome, in which the cleavage site is only 17 bp from the end of the nucleosome, does not permit R2 protein binding. Therefore, the DNA region that must be weakly bound to the histone octamer for R2 accessibility is highly asymmetric with reference to the cleavage site, with the critical region located well upstream of this site.

A direct correlation was found between nucleosome binding by the R2 protein, as revealed by mobility shifts, and the ability of the protein to cleave the target site (compare Figures 5A and 6B). Furthermore, once the target site is cleaved, the efficiency with which it is used as a primer for TPRT is similar if not higher in a nucleosome compared with naked DNA. Thus, once the R2 protein has gained complete access to its target site, cleavage and TPRT are not inhibited. The integrity of the nucleosome has not been disrupted by R2 protein binding, however. Starting two helical turns downstream of the insertion site, the R2-cleaved 39 and 49 END nucleosomes have the typical periodic pattern of hydroxyl radical cleavage, suggesting that they have remained wrapped about the core histones. The protection from DNase I of the DNA 50–90 bp downstream of the target site is also consistent with the large R2 protein partially shielding the DNA, which, after wrapping around the core histone octamer, is now on the same side of the nucleosome as the target site. More detailed studies of the footprint of the R2 protein on free DNA and on the nucleosome will be the subject of future reports.

Mobile elements and chromatin structure

The integration machinery of a mobile element faces many of the same obstacles faced by the cell's transcriptional apparatus. Both must gain access to the DNA when it is assembled into nucleosomes and higher order structures. However, unlike the specificity required of the transcriptional apparatus, mobile elements that insert throughout the genome can simply be opportunistic and insert wherever sites become accessible. Many transposable elements that appear to insert at random locations in the

genome frequently insert within or near transcriptionally active regions (reviewed in Sandmeyer *et al.*, 1990). At the local level, the insertion sites of such mobile element generalists can be greatly influenced by nucleosome positions. For example, *in vitro* studies of insertion site preference of the human immunodeficiency virus have shown that insertions are stimulated in areas of a nucleosome where the DNA is distorted (Pruss *et al.*, 1994a,b). In a similar manner, the cleavage site for the mammalian non-LTR retrotransposon, L1, is changed when the target DNA is assembled into nucleosomes (Cost *et al.*, 2001).

On the other hand, some transposable elements insert into specific locations in the genome and are thus likely to require a more elaborate adaptation to chromatin structures. The best-studied examples are the LTR retrotransposons of yeast, which preferentially insert upstream of tRNA genes. The specificity of the elements for these sites is controlled by an interaction of the element-encoded integrase machinery with the polymerase III transcriptional apparatus (Kirchner *et al.*, 1995; Devine and Boeke, 1996). Also present in yeast are Ty5 elements, which exclusively insert near telomeres and the silent mating loci by means of an interaction of the Ty5 integrase with Sir3p and Sir4p (Zhu *et al.*, 1999). Retrotransposons in many other organisms are likely to insert by interacting with specific components of chromatin because they encode integrases with apparent chromodomains (Malik and Eickbush, 1999a), insert near tRNA genes (Szafranski *et al.*, 1999) or insert near telomeres (Higashiyama *et al.*, 1997).

The final group of mobile element insertions are those that are directed by the DNA sequence itself. These include elements that insert into specific sequences of the 28S gene, spliced leader exons or telomeric sequences (Eickbush, 2002; Takahashi and Fujiwara, 2002). In the case of R2, the region of the 28S rRNA encoded by its target sequence plays an integral role in the interaction of the large and small ribosomal subunits, which is why the nucleotide sequence of this region has remained virtually unchanged in all eukaryotes (Yusupov *et al.*, 2001). From the insect's point of view, each insertion of an R2 element inactivates another rDNA transcription unit, bringing it one step closer to insufficient numbers of units. Thus, one would predict that the insect hosts, unable to change the target sequence because of functional constraints on the encoded rRNA, would attempt to make this target site inaccessible to the R2 element by means of chromatin structure.

Phylogenetic analysis suggests that R2 elements have been inserting into the same 28S gene target site for hundreds of millions of years (Burke *et al.*, 1998, 1999; Malik *et al.*, 1999), yet no aspect of their evolution or their activity would suggest a model of a mutually beneficial relationship (see discussions in Malik and Eickbush, 1999b; Perez-Gonzalez and Eickbush, 2001; Eickbush, 2002). Indeed, R2 is not the only element that has evolved specificity for this target site. Four other non-LTR retrotransposons (R1, R3, R4 and R5) have been identified in arthropods, nematodes and flatworms that specifically insert into the 28S gene within 75 bp of the R2 target site (Jakubczak *et al.*, 1991; Burke *et al.*, 1995; W.Burke, D.Singh and T.Eickbush, in preparation). Why so many

Table I. Primers used in this study

Name	Sequence
39 End Up	5'-CGCGGGTCAACGGCGGGAGTAACCTA-3'
39 End Down	5'-CAAGCCCCGTTCCCTTGGCTGTGGTT-3'
49 End Up	5'-TTCAAGTAAGCGCGGGTCAACGG-3'
49 End Down	5'-CCCTTGGCTGTGGTTTCGCTAGA-3'
74 End Up	5'-TGCTCTGAATGTCAAAGTGAAGAAATTC-3'
74 End Down	5'-GTAGATAGGGACAGTAGGAATCTCGTTAA-TCC-3'
79 End Up	5'-CCCAGTGCTCTGAATGTCAAAGTGA-3'
79 End Down	5'-TAGGGACAGTAGGAATCTCGTTAATCC-ATT-3'
130 End Up	5'-AAACAAAGCATTGTGATGGCCCTAG-3'
130 End Down	5'-GACGAGGCATTGGCTACCTTAAGA-3'
<i>Sacup</i>	5'-AGCGAGCTCGCCTGCATGGAGCTGCCATT-GGTACAGATCT-3'
<i>Pst</i> down	5'-AGCCTGCAGAGGCGTTCAGGCATAATCC-AACGGACGTAGC-3'
AB2N	5'-CGGGATCCTAATACGACTCACTATAGGCC-TTGACAGTAGTCCAGCG-3'
AB2B	5'-TTTTCATCGCCGGATCATC-3'

elements have chosen this small region for insertion remains a mystery.

In summary, the experiments described here have indicated that R2 can conduct the TPRT reaction when the 28S gene target site is assembled into nucleosomes in various translational frames. This may explain our *in vivo* findings that R2 can insert into both R1-inserted (inactive) and uninserted (active) rDNA units (Eickbush *et al.*, 2000). To address the question of target site accessibility more specifically, it will be necessary to determine the chromatin structure of this specific region of the 28S gene *in vivo*. The accessibility of this site is critical to the success of R2 and other site-specific mobile elements, and may be involved in the transcription of the rDNA unit itself.

Materials and methods

Protein, RNA and DNA purification

R2 protein was purified as described previously (Luan *et al.*, 1993) except that 0.4–1.2 M NaCl gradients were used to elute the protein from the DNA cellulose column. DNA cellulose column fractions having R2 activity were pooled and dialyzed against 25 mM Tris–HCl pH 7.5, 400 mM NaCl, 50% glycerol at 4°C for 1 h. The final R2 storage solution was adjusted to 25 mM Tris–HCl pH 7.5, 400 mM NaCl, 1 mM EDTA, 50% glycerol, 0.1% Triton X-100 and 0.2 mg/ml BSA. R2 protein concentrations were determined on SDS–polyacrylamide gels stained with Sypro Orange using BSA as a protein concentration standard.

A 2 kb fragment of *D.melanogaster* 28S rDNA containing the R2 target site was PCR amplified from genomic DNA with primers *Sacup* and *Pst*down (Table I) and cloned into the *SacI* and *PstI* sites of pUC18 (named pDM108). From this plasmid, DNA substrates for nucleosome assembly were amplified with ³²P-5' end-labeled primers, which annealed to specific positions upstream and downstream of the R2 target site (Table I). End-labeling reactions were performed as described in the Fermentas Inc. protocol in 20 µl volume containing 200 ng of primer and 25 µCi of [³²P]ATP (3000 Ci/mmol; New England Nuclear). Polynucleotide kinase was inactivated by heating to 90°C for 15 min. Labeled substrate DNA was further purified by 8% PAGE, eluted with 0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA pH 8.0, 0.1% SDS, precipitated with 3 vols of ethanol and dissolved in 10 mM Tris pH 8.0, 1 mM EDTA (TE solution).

RNA corresponding to the 3' UTR of the *B.mori* R2 element was made in an *in vitro* transcription system. DNA containing the 247 bp 3' UTR

sequence was PCR amplified with primer AB2N containing T7 promoter sequence and AB2B (Table I) from plasmid pBmR2-250A4 (Luan and Eickbush, 1995). Half a microgram of this DNA was transcribed with T7 RNA polymerase (Promega) in the presence of 1 mM rNTPs. After incubation at 37°C for 1.5 h, the template DNA was destroyed by the addition of 12 U of DNase I (Ambion) and further incubated for 40 min. Samples were phenol/chloroform extracted, ethanol precipitated and the RNA pellet dissolved in ddH₂O.

Nucleosome reconstitution and sucrose gradient separation

Nucleosomes were reconstituted by standard salt dialysis (Hayes and Lee, 1997). Briefly, 100 ng of labeled target DNA were mixed with 3 µg of sonicated calf thymus DNA and ~9 µg of core histone proteins purified from chicken erythrocytes in 2 M NaCl solution at 37°C for 30 min. Samples were then dialyzed at 4°C against 10 mM Tris–HCl pH 8.0, 10 mM β-mercaptoethanol containing 1.2, 1.0, 0.8 and 0.6 M NaCl for 2 h at each NaCl concentration, and finally against TE overnight. The quality of reconstituted nucleosomes was monitored by 0.7% agarose–nucleoprotein gel electrophoresis or 5% PAGE in 45 mM Tris–borate, 1 mM EDTA pH 8.0 (0.5× TBE) buffer. Reconstituted nucleosomes were purified on a 5–20% sucrose gradient in 10 mM Tris pH 8.0 and centrifuged at 34 000 r.p.m. at 4°C for 18 h. Fractions were collected from the bottom of tubes, and the nucleoprotein content of each fraction was monitored on 0.7% agarose gels.

DNase I and hydroxyl radical mapping of nucleosome

For DNase I mapping, 100 000 c.p.m. of the purified nucleosomes, containing ~5 ng of DNA, were made 2 mM MgCl₂ and treated with 0.2 U of DNase I (Ambion, TX) at room temperature for 3 min. Reactions were quenched by adding 1/10 vol. of 50 mM EDTA pH 8.0, 50% glycerol. The samples were loaded on native 5% polyacrylamide gels in 0.5× TBE, and the nucleosome bands were cut out after wet gel X-ray exposure. Ten-fold less DNase I was used to digest the same amount of free DNA. Hydroxyl radical mapping was conducted as described previously (Hayes and Lee, 1997). Working solutions of 1 mM Fe²⁺ and 2 mM EDTA, 10 mM sodium ascorbate and 0.12% H₂O₂ were made immediately before each reaction. Nucleosome samples (7/10 of final volume) were mixed with 1/10 vols of the above three solutions and allowed to react at room temperature for 2 min. Reaction was quenched by the addition of glycerol to a final concentration of 5%, loaded onto native gels, the nucleosome bands excised and the DNA eluted. Cleavage of free DNA was conducted in a similar manner to that of the nucleosomes, except that the samples were precipitated directly instead of after gel isolation. Cleaved DNA samples were separated on 8% denaturing polyacrylamide gels and dried; a Molecular Dynamics PhosphorImager was used to collect the data and ImageQuant v.1.2 software was used for quantitation.

R2 and Bst98I cleavage kinetics

For each time point, nucleosomes containing 0.5 ng of labeled DNA and equivalent amounts of naked DNA were incubated at 37°C with 20 ng of R2 or 5 U of *Bst*98I (Promega) in 10 mM Tris–HCl pH 8.0, 100 mM NaCl, 2 mM MgCl₂, 1 mM DTT, 0.1 mg/ml BSA with 500 ng of R2 3' UTR RNA. Samples were removed at various times and cleavage stopped by the addition of EDTA and SDS to a final concentration of 5 mM and 0.02%, respectively. Aliquots were extracted and precipitated before loading onto 8% denaturing PAGE gels. The gels were dried and the extent of cleavage was determined with the PhosphorImager. Competition experiments between nucleosomes reconstituted with different templates were conducted in the same manner as the single nucleosome/DNA competitions using 0.33 ng each of the labeled nucleosomes or naked DNAs. In the case of the free DNA competition, only 1 ng of R2 or 0.05 U of *Bst*98I was used for each time point.

Gel mobility shift assays

Samples of nucleosomes containing 0.5 ng of labeled DNA or the same amount of free DNA were incubated at 37°C for 20 min with 20–40 ng of R2 protein in 10 mM Tris–HCl pH 8.0, 50 mM NaCl, 2 mM MgCl₂, 1 mM DTT, 0.1 mg/ml BSA and 5% glycerol buffer with or without 500 ng of R2 3' UTR RNA. In the case of the DNase I footprinting, after incubation with the R2 protein, the nucleosomes were treated with 0.2 U of DNase I (Ambion) for 3 min. The reaction was quenched by the addition of 1/10 vol. of 50 mM EDTA pH 8.0. The nucleosome or DNA samples were separated by 5% PAGE run in 0.5× TBE at room temperature. Electrophoresis was either for 45 min at 100 V on small (5 × 10 cm) gels or for 2 h and 15 min at 150 V on large (17 × 17 cm) gels. Gels were either dried for PhosphorImager analysis or the shifted bands after wet gel

autoradiography were excised, the DNA eluted and separated by 8% denaturing PAGE.

TPRT reaction by sequential addition of Bst98I and R2

One nanogram of labeled free DNA or assembled nucleosome was mixed with 5 U of *Bst*98I in 50 mM NaCl, 10 mM Tris-HCl, pH 8.0, 2 mM MgCl₂, 1 mM DTT, 300 ng of R2 3' UTR RNA and 0.1 mg/ml BSA with 25 µM dNTPs at 37°C for 10 min. Twenty nanograms of R2 protein were then added and the incubation continued for another 50 min. Samples were treated with RNase A, and phenol/chloroform extracted before separation by 8% denaturing PAGE.

Acknowledgements

We thank Shawn Christensen for unpublished results and comments on the manuscript, and Danna Eickbush for comments on the manuscript. This work was supported by NIH grant GM42790 to T.H.E. and by GM52426 to J.J.H.

References

- Becker,P.B. and Horz,W. (2002) ATP-dependent nucleosome remodeling. *Annu. Rev. Biochem.*, **71**, 247–273.
- Burke,W.D., Muller,F. and Eickbush,T.H. (1995) R4, a non-LTR retrotransposon specific to the large subunit rRNA genes of nematodes. *Nucleic Acids Res.*, **23**, 4628–4634.
- Burke,W.D., Malik,H.S., Lathe,W.C.,III and Eickbush,T.H. (1998) Are retrotransposons long-term hitchhikers? *Nature*, **392**, 141–142.
- Burke,W.D., Malik,H.S., Jones,J.P. and Eickbush,T.H. (1999) The domain structure and retrotransposition mechanism of R2 elements are conserved throughout arthropods. *Mol. Biol. Evol.*, **16**, 502–511.
- Conconi,A., Widmer,R.M., Koller,T. and Sogo,J.M. (1989) Two different chromatin structures coexist in ribosomal RNA genes throughout the cell cycle. *Cell*, **57**, 753–761.
- Conconi,A., Sogo,J.M. and Ryan,C.A. (1992) Ribosomal gene clusters are uniquely proportioned between open and closed chromatin structures in both tomato leaf cells and exponentially growing suspension cultures. *Proc. Natl Acad. Sci. USA*, **89**, 5256–5260.
- Cost,G.J., Golding,A., Schlissel,M.S. and Boeke,J.D. (2001) Target DNA chromatinization modulates nicking by L1 endonuclease. *Nucleic Acids Res.*, **29**, 573–577.
- Devine,S.E. and Boeke,J.D. (1996) Integration of the yeast retrotransposon Ty1 is targeted to regions upstream of genes transcribed by RNA polymerase III. *Genes Dev.*, **10**, 620–633.
- Eickbush,T.H. (2002) R2 and related site-specific non-LTR retrotransposons. In Craig,N.L., Craigie,R., Gellert,M. and Lambowitz,A.M. (eds), *Mobile DNA II*. American Society of Microbiology Press, Washington, DC, pp. 813–833.
- Eickbush,D.G., Luan,D.D. and Eickbush,T.H. (2000) Integration of *Bombyx mori* R2 sequences into the 28S ribosomal RNA genes of *Drosophila melanogaster*. *Mol. Cell. Biol.*, **20**, 213–223.
- Grunstein,M. (1997) Histone acetylation in chromatin structure and transcription. *Nature*, **389**, 349–352.
- Hayes,J.J. and Lee,K.M. (1997) *In vitro* reconstitution and analysis of mononucleosomes containing defined DNAs and proteins. *Methods*, **12**, 2–9.
- Higashiyama,T., Noutoshi,Y., Fujie,M. and Yamada,T. (1997) Zepp, a LINE-like retrotransposon accumulated in the *Chlorella* telomeric region. *EMBO J.*, **16**, 3715–3723.
- Jakubczak,J.L., Burke,W.D. and Eickbush,T.H. (1991) Retrotransposable elements R1 and R2 interrupt the rRNA genes of most insects. *Proc. Natl Acad. Sci. USA*, **88**, 3295–3299.
- Jamrich,M. and Miller,O.L. (1984) The rare transcripts of interrupted rDNA genes in *Drosophila melanogaster* are processed or degraded during synthesis. *EMBO J.*, **3**, 1541–1545.
- Kidd,S.J. and Glover,D.M. (1981) *Drosophila melanogaster* ribosomal DNA containing type II insertions is variable transcribed in different strains and tissues. *J. Mol. Biol.*, **151**, 645–662.
- Kirchner,J., Connolly,C.M. and Sandmeyer,S.B. (1995) Requirement of RNA polymerase III transcription factors for *in vitro* position-specific integration of a retrovirus like element. *Science*, **267**, 1488–1491.
- Kolchinskii,A.M., Bashakidze,R.P., Preobrazhenskaya,O.V., Karpov,V.L. and Mirzabekov,A.D. (1984) Chromatin structure of ribosomal genes of *Drosophila melanogaster*. Random arrangement of nucleosomes on DNA and peculiarities in the organization of the nontranscribed spacer. *Mol. Biol. (Mosk.)*, **18**, 1141–1150.
- Kornberg,R.D. and Lorch,Y. (1999) Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell*, **98**, 285–294.
- Li,Q. and Wrangé,O. (1995) Accessibility of a glucocorticoid response element in a nucleosome depends on its rotational positioning. *Mol. Cell. Biol.*, **15**, 4375–4384.
- Lohr,D. (1983) Chromatin structure differs between the coding and upstream flanking sequences of the yeast 35S ribosomal genes. *Biochemistry*, **22**, 927–934.
- Long,E.O. and Dawid,I.B. (1979) Expression of ribosomal DNA insertions in *Drosophila melanogaster*. *Cell*, **18**, 1185–1196.
- Luan,D.D. and Eickbush,T.H. (1995) RNA template requirements for target DNA-primed reverse transcription by the R2 retrotransposable element. *Mol. Cell. Biol.*, **15**, 3882–3891.
- Luan,D.D., Korman,M.H., Jakubczak,J.L. and Eickbush,T.H. (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell*, **72**, 595–605.
- Luger,K., Mader,A.W., Richmond,R.K., Sargent,D.F. and Richmond,T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, **389**, 251–260.
- Malik,H.S. and Eickbush,T.H. (1999a) Modular evolution of the integrase domain in the Ty3/Gypsy class of LTR retrotransposons. *J. Virol.*, **73**, 5186–5190.
- Malik,H.S. and Eickbush,T.H. (1999b) Retrotransposable elements R1 and R2 in the rDNA units of *Drosophila mercatorum*: abnormal abdomen revisited. *Genetics*, **151**, 653–665.
- Malik,H.S., Burke,W.D. and Eickbush,T.H. (1999) The age and evolution of non-LTR retrotransposable elements. *Mol. Biol. Evol.*, **16**, 793–805.
- Moran,J.V. and Gilbert,N. (2002) Mammalian LINE-1 retrotransposons and related elements. In Craig,N.L., Craigie,R., Gellert,M. and Lambowitz,A.M. (eds), *Mobile DNA II*. American Society of Microbiology Press, Washington, DC, pp. 836–869.
- Palen,T.E. and Cech,T.R. (1984) Chromatin structure at the replication origins and transcription-initiation regions of the ribosomal RNA genes of *Tetrahymena*. *Cell*, **36**, 933–942.
- Perez-Gonzalez,C. and Eickbush,T.H. (2001) Dynamics of R1 and R2 elements in the rDNA locus of *Drosophila simulans*. *Genetics*, **158**, 1557–1567.
- Perlmann,T. (1992) Glucocorticoid receptor DNA-binding specificity is increased by the organization of DNA in nucleosomes. *Proc. Natl Acad. Sci. USA*, **89**, 3884–3888.
- Polach,K.J. and Widom,J. (1995) Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. *J. Mol. Biol.*, **254**, 130–149.
- Pruss,D., Bushman,F.D. and Wolffe,A.P. (1994a) Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. *Proc. Natl Acad. Sci. USA*, **91**, 5913–5917.
- Pruss,D., Reeves,R., Bushman,F.D. and Wolffe,A.P. (1994b) The influence of DNA and nucleosome structure on integration events directed by HIV integrase. *J. Biol. Chem.*, **269**, 25031–25041.
- Sandmeyer,S.B., Hansen,L.J. and Chalker,D.L. (1990) Integration specificity of retrotransposons and retroviruses. *Annu. Rev. Genet.*, **24**, 491–518.
- Simpson,R.T. (1990) Nucleosome positioning can affect the function of a *cis*-acting DNA element *in vivo*. *Nature*, **343**, 387–389.
- Simpson,R.T. (1991) Nucleosome positioning: occurrence, mechanisms, and functional consequences. *Prog. Nucleic Acid Res. Mol. Biol.*, **40**, 143–184.
- Sogo,J.M., Ness,P.J., Widmer,R.M., Parish,R.W. and Koller,T. (1984) Psoralen-crosslinking of DNA as a probe for the structure of active nucleolar chromatin. *J. Mol. Biol.*, **178**, 897–919.
- Stancheva,I., Lucchini,R., Koller,T. and Sogo,J.M. (1997) Chromatin structure and methylation of rat rRNA genes studied by formaldehyde fixation and psoralen cross-linking. *Nucleic Acids Res.*, **25**, 1727–1735.
- Struhl,K. (1998) Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev.*, **12**, 599–606.
- Szafarski,K., Glockner,G., Dingermann,T., Dannat,K., Noegel,A.A., Eichinger,L., Rosenthal,A. and Winckler,T. (1999) Non-LTR retrotransposons with unique integration preferences downstream of *Dictyostelium discoideum* tRNA genes. *Mol. Gen. Genet.*, **262**, 772–780.

- Takahashi,H. and Fujiwara,H. (2002) Transplantation of target site specificity by swapping the endonuclease domains of two LINEs. *EMBO J.*, **21**, 408–417.
- Udvardy,A., Louis,C, Han,S. and Schedl,P. (1984) Ribosomal RNA genes of *Drosophila melanogaster* have a novel chromatin structure. *J. Mol. Biol.*, **175**, 113–130.
- Widom,J. (1998) Structure, dynamics, and function of chromatin *in vitro*. *Annu. Rev. Biophys. Biomol. Struct.*, **27**, 285–327.
- Wolffe,A.P. (1998) *Chromatin Structure and Function*, 3rd edn. Academic Press, San Diego, CA.
- Wu,C. (1997) Chromatin remodeling and the control of gene expression. *J. Biol. Chem.*, **272**, 28171–28174.
- Xiong,Y. and Eickbush,T.H. (1988) Functional expression of a sequence-specific endonuclease encoded by the retrotransposon R2Bm. *Cell*, **55**, 235–246.
- Yang,J. and Eickbush,T.H. (1998) RNA-induced changes in the activity of the endonuclease encoded by the R2 retrotransposable element. *Mol. Cell. Biol.*, **18**, 3455–3465.
- Yang,J., Malik,H.S. and Eickbush,T.H. (1999) Identification of the endonuclease domain encoded by R2 and other site-specific, non-long terminal repeat retrotransposable elements. *Proc. Natl Acad. Sci. USA*, **96**, 7847–7852.
- Yusupov,M.M., Yusupova,G.Z., Baucom,A., Lieberman,K., Earnest,T.N., Cate,J.H. and Noller,H.F. (2001) Crystal structure of the ribosome at 5.5 Å resolution. *Science*, **292**, 883–896.
- Zhu,Y., Zou,S., Wright,D.A. and Voytas,D.F. (1999) Tagging chromatin with retrotransposons: target specificity of the *Saccharomyces* Ty5 retrotransposon changes with the chromosomal localization of Sir3p and Sir4p. *Genes Dev.*, **13**, 2738–2749.

Received August 8, 2002; revised October 18, 2002;
accepted October 23, 2002