Hybridization of RNA to double-stranded DNA: Formation of R-loops

(gene isolation/gene mapping/electron microscopy/restriction endonucleases)

MARJORIE THOMAS, RAYMOND L. WHITE, AND RONALD W. DAVIS

Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305

Communicated by Norman Davidson, April 26, 1976

ABSTRACT RNA can hybridize to double-stranded DNA in the presence of 70% formamide by displacing the identical DNA strand. The resulting structure, called an R-loop, is formed in formamide probably because of the greater thermodynamic stability of the RNA-DNA hybrid when it is near the denaturation temperature of duplex DNA. The rate of R-loop formation is maximal at the temperature at which half of the duplex DNA is irreversibly converted to single-stranded DNA (the strand separation temperature or \( t_{\text{m}} \)). In the present study, the rate of R-loop formation was dependent on the concentration of RNA and was observed in the presence of 50% formamide. The rate of hybridization of RNA was minimized by raising the RNA concentration to 100% (the RNA-DNA hybrid is less stable in the presence of 100% formamide than in the presence of 50% formamide). The rate of hybridization of RNA was minimized by raising the RNA concentration to 100% (the RNA-DNA hybrid is less stable in the presence of 100% formamide than in the presence of 50% formamide).

Materials and Methods

Reagents. \( \lambda \)gt-Sc1109 DNA and yeast total rRNA were generously supplied by R. Kramer. Formamide was purified by crystallization at \(-3^\circ\) in a salt-ice bath. The mixture was stirred with a motor-driven stainless steel paddle to the consistency of soft ice cream. A 250-mL polycarbonate centrifuge tube was cut in half and a nylon screen cemented between the halves. The formamide crystals were recovered by centrifugation onto the nylon screen. A hole, drilled just below the screen, was used to remove the sedimented liquid. The formamide was stored at \(-20^\circ\) for hybridization reactions. Small glass test tubes (5 × 60 mm) were siliconized by treatment with a 1% solution of dimethyl dichlorosilane dissolved in benzene. The tubes were then dried in an oven of 200° and extensively washed with water.

R-loop Formation Buffer. A buffered formamide solution was first prepared by mixing 0.42 ml of formamide, 50 \( \mu \)l of 1 M Pipes [piperazine-\( N,N' \)-bis(2-ethanesulfonic acid)] \( \text{Na}_4 \), at pH 7.8, 12 \( \mu \)l of 0.5 M \( \text{Na}_3 \text{EDTA} \), and 18 \( \mu \)l of \( \text{H}_2\text{O} \). Fifty microliters of this solution were delivered with Teflon or polyethylene tubing to the bottom of a 5- × 60-mm siliconized glass test tube. Ten microliters of a solution containing approximately 30 \( \mu \)g/ml each of RNA and DNA, 0.1 M NaCl, and 0.05 M Tris-HCl at pH 7.5, was then added. The final cation concentration is approximately 0.17 M. The solution was covered with paraffin oil and the tube sealed with Parafilm. The test tubes were incubated in a Haake constant temperature bath filled with ethylene glycol.

Determination of Strand Separation Temperature \( (t_{\text{m}}) \). DNA cleaved with EcoRI endonuclease was placed in the R-loop formation buffer as described above. Samples of 4 \( \mu \)l each were taken at 1° intervals starting at 45°. Five minutes were allowed for equilibration between intervals. Each sample was diluted 20-fold into 0.5 M ammonium acetate at pH 8, 0.01 M \( \text{Na}_3 \text{EDTA} \), and 75 \( \mu \)g/ml of cytochrome c at 0°. Samples were mounted for electron microscopy by the aqueous drop method. Drops of 25 \( \mu \)l of the above DNA plus cytochrome solutions were placed on a polished teflon bar. A parlodion-coated microscope grid was touched to the side of each drop and was stained with uranyl acetate (2). At the strand separation temperature \( (t_{\text{m}}) \) of a specific EcoRI DNA fragment, the duplex strands are converted into collapsed single-strand bushes. The \( t_{\text{m}} \) of the RNA-DNA hybrid was determined in an identical manner, except that RNA was first hybridized to its EcoRI-cleaved complementary DNA strand at 47° in the R-loop formation buffer.

Our initial work on R-loop formation was complicated by DNA degradation and nonreproducible reaction rates. We believe this was the consequence of changes in the reaction constituents and their concentrations during the course of the reaction. The following revisions have resulted in our ability to obtain reproducible R-loop formation rates with no detectable DNA degradation when assayed by electron microscopy:

1. Recrystallized formamide was used; (2) the reactions were conducted under oil in sealed siliconized glass test tubes, which prevented evaporation and solvent condensation on the walls; (3) the temperature was maintained to within \( \pm 0.2^\circ \); and (4) the decomposition of the RNA and the pH change of the formamide solution was minimized by using Pipes buffer at pH 7.8.

RESULTS

Model system for R-loop formation

The study of R-loop formation has been greatly facilitated by the use of a simple model system. A viable hybrid DNA molecule (\( \lambda \)gt-Sc1109) containing the 2.6 kb (kilobase pair) repetitive...
Table 1. Strand separation temperatures \( (t_{ss}) \) of EcoRI segments of \( \lambda \text{gt-Sc1109} \) DNA

<table>
<thead>
<tr>
<th>EcoRI segment</th>
<th>( t_{ss} ) °C</th>
<th>Approximate % G+C (5, 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda B )</td>
<td>45</td>
<td>37</td>
</tr>
<tr>
<td>rDNA</td>
<td>49</td>
<td>47</td>
</tr>
<tr>
<td>( \lambda \text{gt right} )</td>
<td>50</td>
<td>47</td>
</tr>
<tr>
<td>( \lambda \text{gt left} )</td>
<td>55</td>
<td>57</td>
</tr>
<tr>
<td>rDNA-rRNA</td>
<td>58</td>
<td>47</td>
</tr>
</tbody>
</table>

EcoRI DNA fragment from \( S. \text{ceres} \) (yeast) covalently inserted into the middle of bacteriophage \( \lambda \) DNA was used. The inserted DNA fragment hybridizes to yeast \( 26S \) rRNA, which can easily be prepared from yeast cells in large quantities. This \( \lambda \)-yeast hybrid is one of several isolated from large pools of \( \lambda \)-yeast hybrids by \textit{in situ} hybridization of RNA to individual plaques (Kramer, Cameron, and Davis, manuscript in preparation). As shown below, the yeast DNA fragment coding for rRNA (rDNA) is about \( \frac{3}{4} \) of the length of the \( 26S \) rRNA and hybridizes at one end of the \( 26S \) rRNA.

Strand separation temperatures of the \( \lambda \text{gt-Sc EcoRI} \) DNA fragments

The strand separation temperatures of the EcoRI DNA fragments from \( \lambda \text{gt-Sc1109} \) were determined under the R-loop reaction conditions (70% formamide, 0.1 M Pipes, 0.01 M Na3EDTA). Samples were taken at 1°C intervals starting at 45°C and examined in the electron microscope as described in Materials and Methods. The temperature at which half of the duplex DNA is irreversibly converted to single-stranded DNA is defined as the strand separation temperature \( (t_{ss}) \) of this fragment. The strand separation of the rDNA fragment occurred within a 1°C interval, while the strand separation of the left and right end fragments of bacteriophage \( \lambda \) occurred with partial denaturation within some of the molecules over a 4°C range. Table 1 shows the \( t_{ss} \) of the EcoRI fragments rDNA, \( \lambda B \), \( \lambda \text{gt left arm} \), and \( \lambda \text{gt right arm} \) (3), and their approximate base compositions (4, 5). The \( t_{ss} \) is conceptually different from the melting temperature \( (t_m) \) where \( \frac{1}{2} \) of the base pairs are no longer paired.

R-loop formation to saturation

An initial attempt to obtain a uniform population of DNA

![R-loops](image.png)

**FIG. 1.** R-loops were made by heating 5 \( \mu \text{g/ml} \) of \( \lambda \text{gt-Sc1109} \) DNA and 5 \( \mu \text{g/ml} \) total rRNA in 70% vol/vol formamide, 0.1 M Pipes at pH 7.8, and 0.01 M Na3EDTA at 47°C for 20 hr. The reaction was performed under oil in a sealed, siliconized glass tube. All 500 RNA molecules examined contained an R-loop similar to those shown. The sample was mounted for electron microscopy by the formamide technique (2). Grids were stained with uranyl acetate and shadowed with Pt/Pd.
molecules containing R-loops was made by heating a mixture of λgt-Sc1109 DNA at 5 μg/ml and total rRNA at 5 μg/ml at 47°C for 20 hr. As shown in Fig. 1, clear R-loops were formed. Each of 500 DNA molecules examined contained an R-loop. Therefore, it is possible to saturate DNA for the formation of R-loops. Measurements of the length of the DNA-RNA portion of the R-loops indicate that the rDNA fragment is homologous to about 1/3 of the length of the 26S rRNA. Single-stranded tails of RNA were only observed on one end of each R-loop, indicating that the rDNA fragment is homologous to one end of the 26S rRNA molecule.

**Determination of the rate of R-loop formation**

The reaction conditions used here were selected to give a reaction rate which could be readily measured. Much higher reaction rates can be achieved by increasing the RNA and salt concentrations. The course of R-loop formation under various conditions was followed by examination of samples in the electron microscope. The fraction of DNA molecules containing an R-loop was scored. The size of the R-loop was not considered. A total of 100–200 molecules were scored at each time point and 2–4 time points were taken for each condition. The accuracy of scoring in the electron microscope is greatest when 1/2 of the molecules have reacted. Therefore, time points near the 1/2 reaction time were scored and, in this work, the rate will be expressed as the time for 1/2 reaction since this value is directly measured.

For the reaction DNA + RNA → DNA·RNA in which RNA is in molar excess so that its concentration does not change during the reaction, a plot of the log of the fraction of unreacted DNA as a function of time should give a straight line. This is supported by the data in Fig. 2 at two different RNA concentrations.

**Rate of R-loop formation as a function of temperature**

The time course of R-loop formation was followed at a series of temperatures using a 20-fold molar excess of RNA. Some of these time courses are shown in Fig. 3. The 1/2 reaction time was determined from such plots for each temperature. A plot of the log of the 1/2 reaction time as a function of temperature is shown in Fig. 4. The rate of R-loop formation is very dependent upon the temperature of the reaction. This plot also demonstrates that the temperature (t_m) at which maximum reaction rate occurs is very close to the t_m (49°C) of the DNA. The temperature at which the largest and most uniform-appearing R-loops were formed was from t_m -2°C to t_m +1°C. At t_m -9°C, most of the R-loops were quite small and were formed from small RNA fragments. At t_m +2°C, denatured fragments of λgt-Sc1109 DNA were observed.

**Rate of R-loop formation as a function of RNA concentration and temperature**

For a simple bimolecular reaction, the rate of R-loop formation should be directly proportional to the RNA concentration. This prediction is verified for the reaction at t_m +1°C as shown in Table 2. However, the rate at t_m -6°C, as shown in Fig. 2 and Table 2, is increased by a factor of only 1.5 when the RNA concentration is increased by a factor of 10. This behavior is also found at t_m -9°C. Between t_m -6°C and t_m +1°C, the reaction rate converts from one virtually independent of RNA concentration to one directly proportional to RNA concentration.
Rate of R-loop formation as a function of ionic strength

The rate of DNA renaturation increases about 10-fold when the sodium ion concentration is increased from 0.17 M to 0.67 M (6). The rate of R-loop formation was determined for this same increase in salt concentration. The addition of 0.5 M NaCl to the reaction mixture should increase the $t_{ss}$ of DNA by 9° (6). The $\frac{1}{2}$ reaction time at 56° ($t_{ss} = 2°$) in 0.5 M NaCl, 0.1 M Pipes at pH 7.8, 0.01 M Na$_2$EDTA with 5 µg/ml of DNA and 5 µg/ml of RNA is about 3.5 min. This represents approximately a 15-fold increase in reaction rate over the same reaction in the absence of NaCl.

Stability of R-loops

The large R-loops used in this work, once formed, are quite stable. They are quantitatively retained after storage for several days at 5° in the formamide reaction mixture. Two samples, stored for 2 wk after dialysis into 1 M NaCl--0.01 M Tris-HCl at pH 7.5--0.01 M Na$_2$EDTA or 0.1 M NaCl--0.01 M Tris-HCl at pH 7.5--0.01 M Na$_2$EDTA, that originally contained 100% R-loops, were reduced to 99% and 98%, respectively. These dialyzed R-loop samples were also partially resistant to heat. About 25% of the R-loops remained in the sample containing 1 M NaCl after heating to 70° ($t_{m} = 30°$) for 5 min, and about 75% of the R-loops remained in the sample containing 0.1 M NaCl after heating to 50° ($t_{m} = 30°$) for the same length of time.

Since such stability of the R-loops was observed, their stability during and after restriction endonuclease cleavage of R-loops containing DNA at 37° was tested. Fig. 5 shows one example of R-loops which were easily observed after cleavage of the DNA with EcoRI endonuclease.

Removal of R-loops

Upon removal of formamide, R-loops are thermodynamically unstable at $t_{m} = 30°$ and are displaced by branch migration (7). However, this displacement of large R-loops by identical DNA strands is slow. Therefore, destruction of the RNA seems necessary for the elimination of R-loops. Two methods were successfully used. (1) The RNA was hydrolyzed by addition of 0.2 M NaOH and incubated at 37° for 10 min. Denaturation of the DNA was prevented by addition of MgSO$_4$ to a final concentration of 0.01 M prior to addition of alkali. After incubation, the solution was neutralized with Tris-HCl and the Mg ion was removed by increasing the EDTA concentrations to 0.02 M. Some DNA aggregation was observed in the electron microscope with this method. (2) The R-loops are quite sensitive to RNase. Incubation for 5 min at 37° in 0.1 M NaCl, 0.05 M Tris-HCl at pH 7.5, 10^{-4} M EDTA with 10 µg/ml of RNase A resulted in the total loss of R-loops.

DISCUSSION

The temperature for maximum rate ($t_{max}$) of R-loop formation is within one degree of the $t_{ss}$ of the DNA. From the values for $t_{ss}$ of three λDNA EcoRI fragments with different known base compositions (Table 1), the following equation relating $t_{max}$ to base composition can be derived for the conditions used in these experiments:

$$t_{max} = 26.5 + 0.50 (%G+C).$$

If one assumes that the effect of salt concentration on the denaturation of DNA in formamide is the same as in aqueous so-
lution, then from Schildkraut's equation (8) in 70% formamide:
\[ t_{\text{max}} = 39 + 0.50 \times (\% \text{G+C}) + 16.6 \log \text{Na}^+ \]

Also from Schildkraut's equation, the effect of formamide can be expressed as:
\[ t_{\text{max}} = 81.5 + 0.50 \times (\% \text{G+C}) + 16.6 \log \text{Na}^+ - 0.60 \times \text{formamide} \]

These equations have not been rigorously tested and may not apply if the conditions are significantly varied from those used in this work.

R-loops are not thermodynamically stable at \( t_m -30^\circ \) after the formamide is removed since the RNA is slowly displaced by the identical DNA strand through branch migration (7). At low temperatures (\( t_m -70^\circ \)), in the absence of formamide, the stability may result from random base pairing in the single-stranded regions which block branch migration. At higher temperatures (\( t_m -30^\circ \)), branch migration commences and the RNA is found to be partially displaced. However, if displacement occurs at both ends, the RNA must thread through the R-loop once for each turn of duplex RNA-DNA hybrid removed. As the RNA is displaced, the loop becomes progressively smaller, and, concurrently, the displacement may become progressively slower. Consistent with this notion are the observations that, after heating in the absence of formamide, most R-loops are very small and that the rate of displacement at \( t_m -30^\circ \) is slower in low ionic strength buffer where electrostatic repulsion between the strands is greater. As evidenced by the RNase sensitivity of the R-loops, if the RNA is destroyed as it is displaced, these blocks to the branch migration are removed.

The formation of R-loops is probably equally complicated. The most likely mechanism for their formation is the generation of denatured regions in the duplex DNA and nucleation of RNA within these denatured regions. The size and frequency of these denatured regions would depend upon the temperature, being small and infrequent below the \( t_m \) of the DNA. If the size of a denatured region in the DNA is small relative to the RNA, then it seems unlikely that irreversible nucleation can occur easily in the center of the RNA because the RNA would have to thread through a very small loop. Nucleation in this case probably occurs most readily at the end of the RNA. Thus, if there are nonhomologous sequences on the end of the RNA, as is the case in this study, the rate of R-loop formation may be reduced.

The observation that the rate of R-loop formation is virtually independent of the RNA concentration at temperatures below \( t_m -5^\circ \) is difficult to understand. One possible explanation is that the rate-limiting step under these conditions is the formation of denatured regions. Another possible explanation is that the rate-limiting step under these conditions is subsequent to the nucleation event. The addition of base pairs to the nucleation complex may be slow due to steric interference in winding. The nucleation complexes that do not proceed to full R-loop formation dissociate before or during mounting for electron microscopy. The rate limiting step above \( t_m \) should be the rate of RNA nucleation. Therefore, above \( t_m \) the rate of R-loop formation should be directly proportional to the RNA concentration as is found in this study. Also, as is found in this study, the rate of R-loop formation above the \( t_m \) of the DNA should decrease because of the decreased stability of the nucleation events. With low RNA concentrations at temperatures below \( t_m \), the rate-limiting step could be the RNA nucleation. Consistent with this suggestion is the observation that, at low RNA concentration between \( t_m -6^\circ \) and \( t_m \), the rate of R-loop formation is fairly independent of temperature.

It is not clear if very small R-loops can be generated since our initial attempts to form R-loops with yeast 55 RNA have not been successful. Small R-loops have an additional complication in that the RNA is probably rapidly displaced.

The experiments presented here indicate that the R-loop formation reaction is quite complex. There are additional variables that were not quantitatively investigated. For example, the secondary structure in the RNA which is not totally removed under these conditions could affect the R-loop formation rate. Also, the size of the RNA may affect the rate since it was observed that the ratio of R-loops derived from small RNA to those derived from large RNA was greater at temperatures below \( t_m \) than at temperatures above \( t_m \).

The base sequence and composition may also affect the rate of formation and stability of R-loops. R-loops are sufficiently stable to withstand dialysis into a buffer of low ionic strength (0.05 M Tris-Cl at pH 7.5, 0.1 M NaCl) and cleavage with an EcoRI restriction endonuclease. This should greatly facilitate the mapping of a number of rRNA and mRNA sequences in DNA. The end of the R-loop can be rather precisely mapped relative to nearby restriction sites.

The use of R-loops should facilitate the physical isolation of specific genes. For example, R-loops might be formed on unfraccionated DNA with use of a specific mRNA. The resulting preparation containing duplex DNA and R-loops on the specific gene sequence could be sedimented to equilibrium in a CsCl density gradient. The R-loop should give the DNA containing the specific gene sequence a density shift; this would result in considerable purification.