The EMBO Journal vol.9 no.11 pp.3767-3775, 1990

Control of replication of plasmid R1: structures and sequences of the antisense RNA, CopA, required for its binding to the target RNA, CopT

Christine Persson[1], E.Gerhart H.Wagner and Kurt Nordström

Department of Microbiology, Biomedical Center, University of Uppsala, Box 581, S-751 23 Uppsala, Sweden

[1]To whom correspondence should be addressed

Communicated by K.Nordström

The replication frequency of plasmid R1 is determined by the availability of the RepA protein, which acts at the origin of replication to promote initiation. Synthesis of RepA is negatively regulated both at the transcriptional and post-transcriptional levels. Post-transcriptional control is exerted through the action of an antisense RNA, CopA RNA. The target of CopA RNA, CopT RNA, is located in the leader region of the RepA mRNA. Binding between CopA and CopT inhibits *repA* expression. We have previously presented an *in vitro* analysis of the binding reaction between CopA and CopT RNAs. In this communication, we extend the *in vitro* analysis by determining the regions of CopA required for binding, and also demonstrate that binding occurs in at least two steps. The first step is the formation of an initial, transient complex; stem-loop II is the structure in CopA necessary and sufficient for this step. The subsequent step(s), resulting in the formation of a complete duplex, requires a stretch of single-stranded nucleotides located 5′ to stem-loop II in CopA, and its counterpart in CopT. We show that the single-stranded region can be positioned on either side of stem-loop II provided that there is a complementary stretch of nucleotides in CopT, indicating that the second step(s) is not sequence-specific. Furthermore, the effects of salt concentration and temperature on the binding reaction indicate that duplex formation occurs through a mechanism of gradual intra-strand breaking and inter-strand formation of hydrogen bonds.

Key words: antisense RNA/plasmid R1/replication control/RNA binding domains/stem-loop structures

Introduction

Antisense RNAs are small, untranslated RNA molecules that regulate gene expression and DNA replication by several different mechanisms. They are termed antisense because they are (at least partially) complementary to the RNA molecule that is the target of regulation. This implies that, in most cases, the antisense and target RNAs are encoded by the same DNA segment, but transcribed in opposite directions.

Antisense RNAs have been shown or proposed to exert an effect in plasmids, phages, transposon Tn*10*, and chromosomal genes of *Escherichia coli* (see reviews by Green *et al.*, 1986; Simons and Kleckner, 1988). Regulation of DNA replication by antisense RNA occurs in plasmid ColE1 and its relatives, through inhibition of primer formation (Tomizawa *et al.*, 1981), and in plasmid R6K, where replication from one of the origins is inhibited by antisense RNA (Patel and Bastia, 1987). Antisense RNA also regulates replication of the IncFII plasmids (Light and Molin, 1983), of plasmid R1162 (Kim and Meyer, 1986) and of plasmid pT181 (Kumar and Novick, 1985), but in these cases regulation is exerted on the mRNAs of the replication initiation proteins. In addition, antisense RNA is involved in regulation of the *hok/sok* killing function of plasmid R1 (Gerdes *et al.*, 1988) and in control of the conjugal transfer of plasmids of the IncFI and IncFII groups (Dempsey, 1989). Phage development switches can be regulated, at least in part, by antisense RNAs. The antisense RNA of P22 regulates the synthesis of the antirepressor protein (Liao *et al.*, 1987), and in phage lambda two different antisense RNAs regulate the synthesis of the Q gene product (Ho and Rosenberg, 1985; Hoopes and McClure, 1985) and of the cII protein (Krinke and Wulff, 1987), respectively. In transposon Tn*10* the rate of tranposition is regulated through action of an antisense RNA on the mRNA of the transposase (Simons and Kleckner, 1983). Finally, it has been proposed that antisense RNAs may regulate expression of three chromosomal genes: *ompF* (Mizuno *et al.*, 1984), *crp* (Okamoto and Freundlich 1986), and *sulA* (Cole and Honoré, 1989).

Plasmid R1 is a low copy number plasmid and belongs to the IncFII group. The basic replicon, shown in Figure 1, is a contiguous DNA segment of 2 kb containing all genetic information necessary for replication and copy number control (see reviews by Nordström *et al.*, 1984, and Rownd *et al.*, 1985). The replication frequency is tightly controlled by the availability of a rate-limiting protein, RepA, which acts at the origin to promote relication. The synthesis of RepA protein is negatively controlled at the transcriptional and post-transcriptional levels. Post-transcriptional control is exerted by binding of the antisense RNA, CopA, to its target, CopT, located in the leader region of the RepA mRNA. CopA RNA is transcribed from the same DNA as the mRNA leader, but in the opposite direction, and is thus completely complementary to CopT. The mechanism by which the interaction between CopA RNA and CopT RNA results in inhibition of *repA* expression might involve cleavage of the RNA duplex by RNase III (Blomberg *et al.*, 1990).

We have previously presented an *in vitro* analysis of the interaction between CopA RNA and CopT RNA. We showed that the specificity of the interaction between CopA and CopT resides in loop II of CopA and the corresponding CopT loop, since one single base change in these loops resulting in a mismatch at the RNA-RNA level virtually abolished binding *in vitro* (Persson *et al.*, 1988). This result is in accordance with the results of *in vivo* incompatibilty tests performed with cells harbouring both the wild-type plasmid and a plasmid carrying the same mutation (Brady *et al.*, 1983). In addition, we showed that the rate of binding

© Oxford University Press

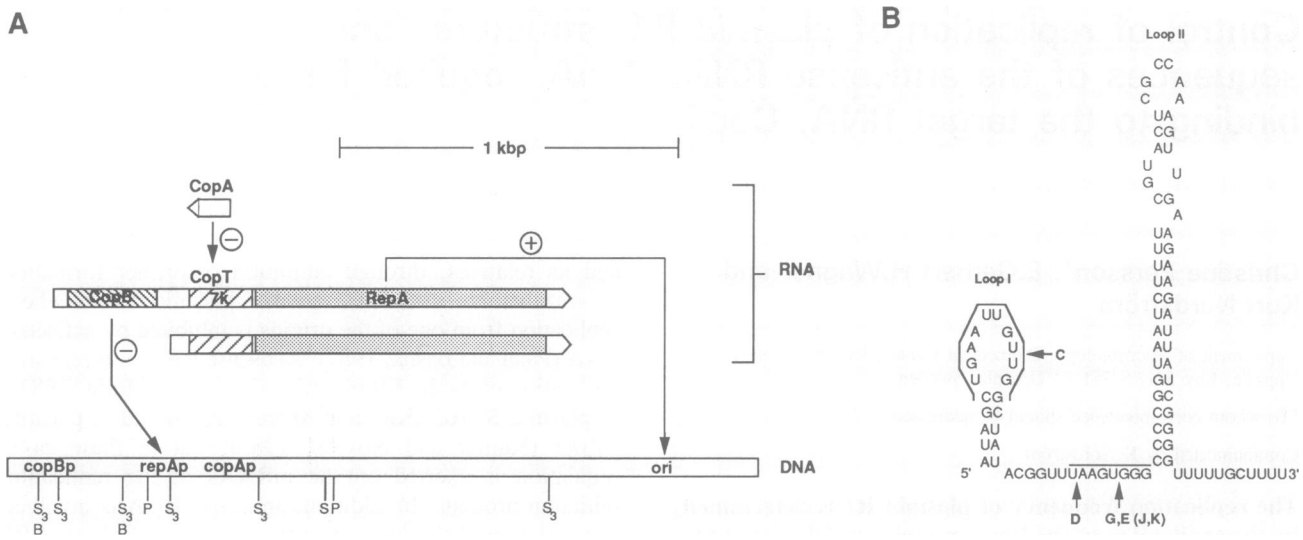

Fig. 1. Basic replicon of plasmid R1. (**A**) The transcription and translation patterns and the regulatory loops involved in replication control. The translated regions are indicated as hatched areas, the positions of the promoters and of the origin of replication are indicated between the lines representing the DNA. The locations of the *Sau*3AI (S_3), *Bgl*II (B), *Pst*I (P), and *Sal*I (S) cleavage sites are shown. The rightmost *Bgl*II site corresponds to position 1 in Ryder *et al.* (1982). (**B**) Secondary structure of CopA RNA (Wagner and Nordström, 1986). The hybridization sites of the oligonucleotides used to cleave CopA (see Materials and methods) are shown as lines above the sequence. The letters refer to the molecules presented in Figure 2, with the arrows indicating the position where cleavage has occurred to yield each molecule. Species D appears spontaneously in CopA preparations (Wagner and Nordström, 1986).

between the RNAs was affected by the nucleotide sequence in the CopA/CopT loop II. Furthermore, the fact that the majority of copy number mutants of R1 is either altered in the primary sequence of the loops or has altered loop sizes points to the loops as being central in mediating the interaction between the molecules (Rosen *et al.*, 1980; Danbara *et al.*, 1981; Stougaard *et al.*, 1981; Brady *et al.*, 1983; Givskov and Molin, 1984). In plasmid ColE1, it has been shown that the three loops in the antisense RNA (RNAI) are involved in an initial 'kissing' reaction with the target RNA (RNAII). The subsequent formation of a stable duplex requires an interaction between the 5′ tail of RNAI and the complementary region of RNAII (Tomizawa, 1984).

In this communication we have analysed which regions of the CopA molecule are essential for duplex formation to occur. The data indicate that stem−loop II and a single-stranded region located 5′ to stem−loop II are required. The results also show that formation of the RNA duplex takes place in at least two steps. The first step(s) is the formation of a 'kissing intermediate' involving the loops II of CopA and CopT and the upper parts of the corresponding stems. This interaction is a prerequisite for complete pairing. The second step, resulting in the formation of the full duplex, requires a single-stranded stretch of nucleotides located 5′ to stem−loop II in the CopA molecule. This second interaction between CopA and CopT is not sequence-specific.

Results

Determination of regions of CopA essential for binding to CopT in vitro

To determine what parts of the CopA RNA molecule are required for binding to CopT *in vitro* we prepared different segments of CopA RNA (see Materials and methods) and tested them in binding experiments. A molecule consisting of only the 5′-most 31 nucleotides of CopA was inactive in binding CopT (species G in Figure 2). This result was expected, since the 5′ fragment lacks loop II, which has been implicated in determining rate and specificity of the binding reaction *in vitro* (Persson *et al.*, 1988). On the other hand, the corresponding 3′ segment, lacking the 5′-most 31 nucleotides (species E in Figure 2), bound CopT with a rate about 50-fold lower than full-length CopA RNA (Figures 2 and 3E). The 5′-most stem−loop (loop I) and almost all of the single-stranded region connecting stem−loops I and II are deleted in this molecule, indicating that stem−loop II is not sufficient for effective duplex formation. This was corroborated by testing a CopA derivative consisting of only stem−loop II ($CopA_F$−species F in Figure 2), which bound CopT with an even lower rate. Since the 5′ region of CopA appeared to be important for efficient duplex formation, we tested two RNAs with increasingly longer 5′ tails. One species lacked 27 nucleotides at the 5′ end (species D in Figure 2), and the other was shortened by 13 nucleotides (species C in Figure 2); these RNAs had wild-type 3′ ends. The $CopA_D$ species showed a higher binding rate than $CopA_E$, but still lower than full-length CopA, whereas the rate of binding of $CopA_C$ to CopT was comparable to that of intact CopA (Figures 2 and 3C). A control experiment was performed in parallel where CopA RNA with its 5′-triphosphate deleted was tested (species B in Figure 2), since the CopA segments we used lacked 5′-triphosphate. The presence or absence of triphosphate had no effect on binding rates (Figure 2). These results indicate that the single-stranded region on the 5′ side of stem−loop II of CopA is important in determining the rate of binding. The data furthermore suggest that the single-stranded region must be of some length, and that an intact stem−loop I structure is dispensable for binding.

We also analysed whether an extended CopA molecule was functionally active in binding CopT RNA. Such a molecule could, in principle, adopt an alternative configuration, precluding the formation of essential structures. A longer CopA RNA species of about 200 nucleotides, which

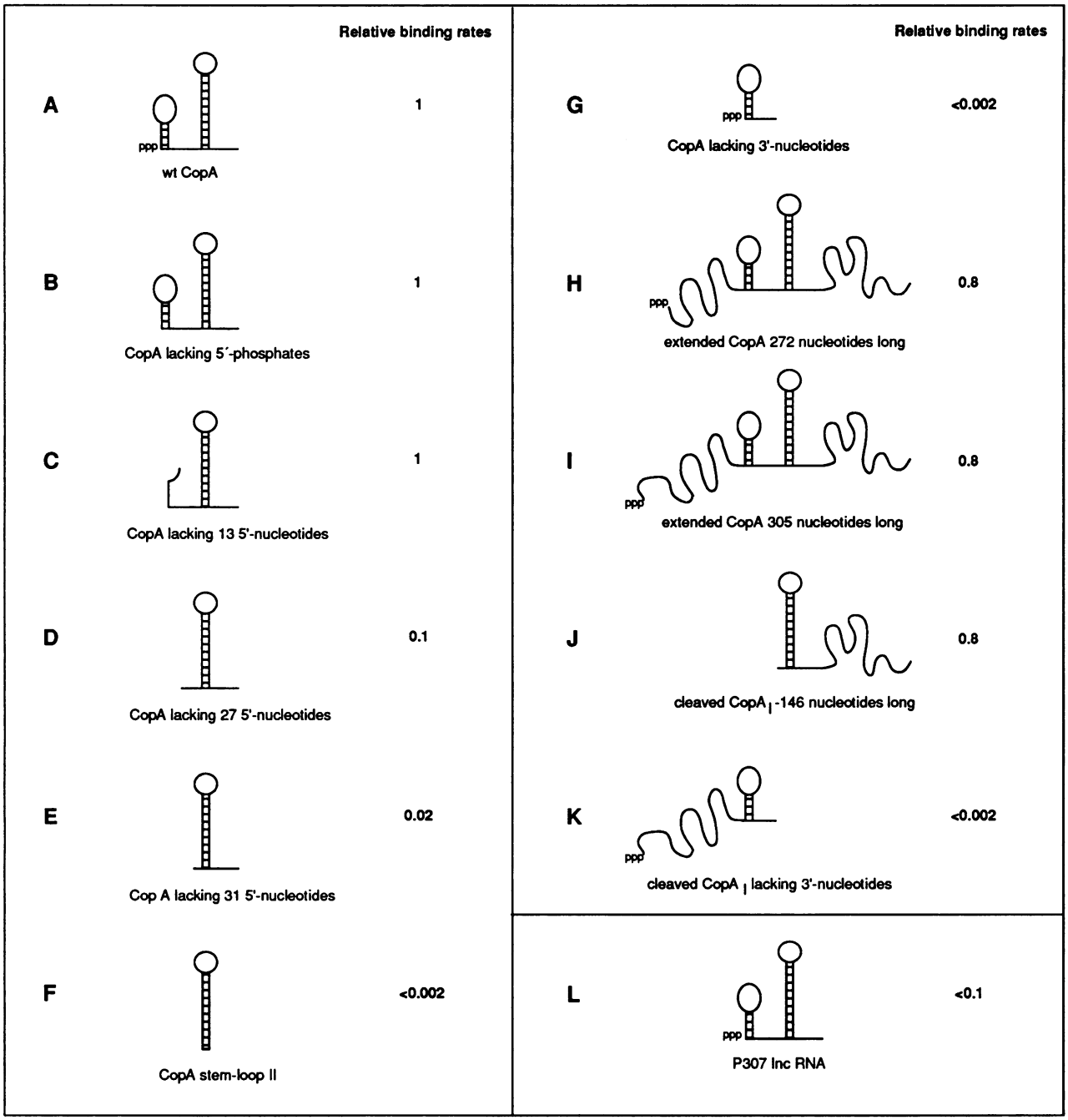

Fig. 2. Comparison between the binding rates of various CopA RNA species and $CopT_{322}$ RNA. CopA derivatives were obtained either by cleavage of wild-type RNA (or extended CopA RNA) or by transcription from other templates (Materials and methods). The experiments were performed using unlabelled $CopT_{322}$ RNA at a concentration of 5.4×10^{-9} M for binding with $CopA_H$, and at 6.2×10^{-9} M for binding with the rest of the molecules. The concentrations of the radiolabelled CopA species were about 10- to 20-fold lower than that of CopT. The values given are averages from several experiments performed with each CopA derivative. In the experiment with P307 Inc RNA (species L), labelled $CopT_{322}$ RNA was mixed with unlabelled Inc RNA (4.2×10^{-9} M).

was believed to be a result of convergent transcription initiating at the *repA* promoter, had earlier been suggested to lack CopA function (Stougaard *et al.*, 1982). We tested a CopA RNA that was 85 nucleotides longer on the 3′ side and 94 nucleotides longer on the 5′ side (272 nucleotides long, species H in Figure 2). All extra nucleotides, except for six at the very 3′ end, were complementary to CopT. $CopA_H$ was essentially as effective as intact CopA in binding CopT (Figures 2 and 3H). In addition, we tested another extended molecule, $CopA_I$, which is 305 nucleotides long and contains 34 5′-nucleotides that are derived from vector sequences and are, thus, not complementary to CopT. This molecule also exhibited almost wild-type binding properties. It is therefore likely that stem–loop II constitutes a structurally stable domain of the extended CopA molecules.

In the experiments reported above, a single-stranded stretch of >7 nucleotides positioned 5′ to stem–loop II was required for normal binding kinetics. It is, however, conceivable that single-stranded regions located elsewhere could substitute for this requirement in artificial CopA RNA

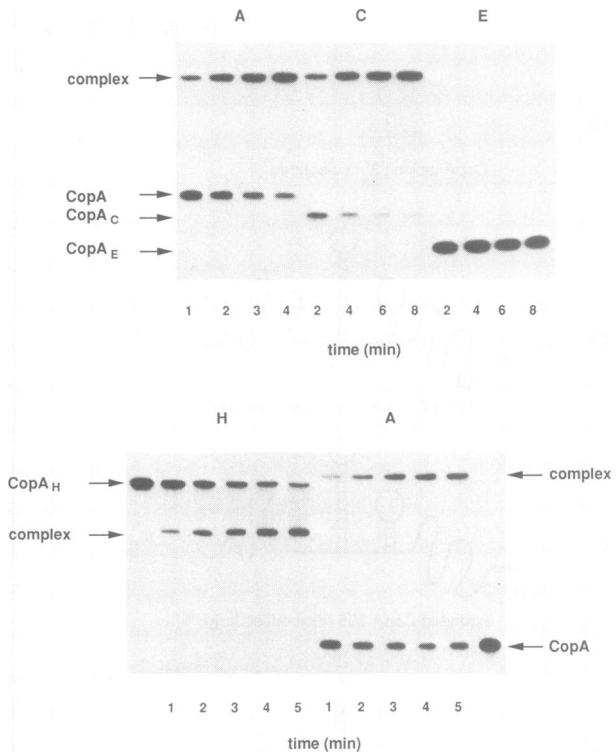

Fig. 3. Time course of duplex formation between various CopA RNA species and CopT RNA. Each radiolabelled CopA species was incubated with unlabelled CopT RNA in binding buffer at 37°C. At the times indicated aliquots were diluted with gel application buffer at 0°C to stop the reaction. The samples were subsequently electrophoresed on polyacrylamide gels containing 7 M urea. The CopT RNA used in these experiments was 322 nucleotides long ($CopT_{322}$ RNA). **Upper panel**: the concentration of the unlabelled CopT was 6.2×10^{-9} M and the concentration of each labelled CopA species was approximately 10-fold lower. **Lower panel**: the concentration of unlabelled CopT was 5×10^{-9} M. The concentration of each labelled CopA species was 25- or 15-fold lower, respectively. The leftmost and rightmost lanes contain radiolabelled $CopA_H$ and wild-type CopA, respectively.

species. We therefore asked whether an extended 3′ tail improved the binding properties of a previously poorly active CopA species. Such a molecule was made by cleaving $CopA_I$ in the same way as we had previously cleaved CopA to yield $CopA_E$. The new molecule, $CopA_J$, bound CopT at about the same rate as wild-type CopA, even though it lacks the 5′ single-stranded region (Figure 2). These results indicate that (i) the single-stranded region can be positioned on either side of stem−loop II, and (ii) there is no specific sequence requirement for this part of the binding reaction. The 5′ half resulting from cleavage of $CopA_I$, i.e. $CopA_K$, which contains as many as 125 nucleotides complementary to CopT, bound CopT extremely poorly (Figure 2).

To summarize, we conclude that stem−loop II of CopA is essential for binding, since it is present in all molecules that bind CopT at wild-type rates. On the other hand, stem−loop II is not sufficient for efficient duplex formation. A single-stranded region of more than about 10 bases is necessary, but its position can be either 3′ or 5′ of stem−loop II.

Binding between a truncated CopT molecule and a CopA derivative with an extended 3′ tail

In a previous communication (Persson *et al.*, 1988) we presented data on the binding of CopA RNA to CopT RNAs

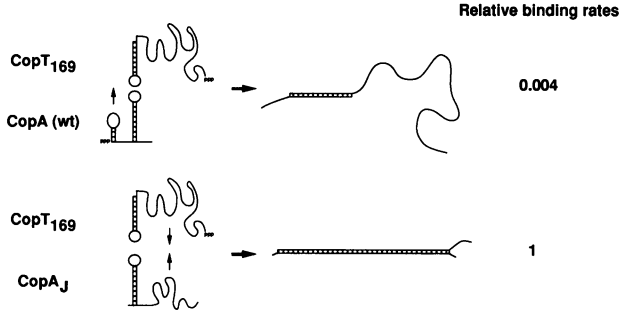

Fig. 4. Schematic representation of the binding between wild-type CopA and $CopT_{169}$ RNA and $CopA_J$ and $CopT_{169}$ RNA. Note that the molecules are not drawn to scale. The rate of binding between $CopA_J$ and $CopT_{169}$ RNA is about two-fold higher than that between CopA and $CopT_{322}$ RNA (data not shown).

of different lengths. We showed that there were some variations in the relative binding rates of the different CopT RNAs tested, but that all CopT RNAs containing complete target regions were good substrates for the binding reaction. One molecule was tested which lacked part of the target region. This RNA, $CopT_{169}$, which is 169 nucleotides long and lacks the nucleotides complementary to the 5′-most 37 nucleotides of CopA RNA, bound CopA very poorly. Previously we could not discriminate between two possible explanations for the observed effect: either the lack of part of the target region as such, or an altered secondary structure. The identification of the 5′ side of CopA as essential for complete duplex formation suggested the former explanation to be more likely. To test this, we performed an experiment where $CopT_{169}$ RNA was allowed to react with $CopA_J$. The data in Figure 4 indicate that the presence of an extended (and presumably at least partly single-stranded) 3′ tail in CopA, complementary to CopT, can allow $CopT_{169}$ RNA to function efficiently in binding.

Duplex formation between CopA and CopT is inhibited by a shortened CopA molecule

Tomizawa (1984) suggested that pairing leading to complete duplex formation between RNAI and RNAII of ColE1 occurs via a series of steps. The first step is a transient interaction between loops of folded structures of RNAI and RNAII. This initial interaction was termed 'kissing', to distinguish it from stable pairing. Later, kinetic analysis of the inhibition of binding between RNAI and RNAII by a heterologous RNAI demonstrated that a kissing intermediate was formed (Tomizawa, 1985). The results presented above show that stem−loop II is needed for efficient pairing between CopA and CopT. However, stem−loop II is not sufficient, since CopA species lacking regions on the 5′ side bound CopT RNA at lower rates. This suggested that the formation of a complete duplex between CopA and CopT also proceeds in at least two steps and that stem−loop II could be involved in the 'kissing' interaction. The reaction between CopA and CopT can be represented by:

$$A + T \rightleftharpoons A{:}T \rightarrow AT$$

where A, T, A:T and AT are CopA RNA, CopT RNA, the kissing intermediate and the complete duplex, respectively. If stem−loop II is involved in the kissing interaction, a molecule such as $CopA_E$ should be able to competitively inhibit binding between wild-type CopA and CopT by interfering with the formation of the kissing intermediate.

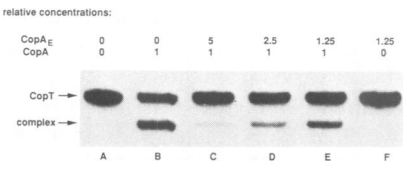

Fig. 5. CopA$_E$ RNA competitively inhibits binding between CopA RNA and CopT RNA. Labelled CopT$_{322}$ RNA ($\sim 1 \times 10^{-10}$ M) was incubated with unlabelled CopA (4×10^{-9} M) either in the absence of CopA$_E$ (lane B) or in the presence of various concentrations of CopA$_E$ (lanes C, D and E). In lane F, CopT$_{322}$ RNA was incubated together with CopA$_E$ in the absence of CopA. The reactions were stopped after 3 min, and the reaction mixtures analysed on a 4% polyacrylamide gel containing 7 M urea.

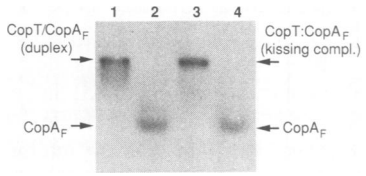

Fig. 6. Stable kissing complex formation analysed on a native polyacrylamide gel. Kissing complex was formed between CopA$_F$ (uniformly ^{32}P-labelled) and unlabelled CopT as described in Materials and methods. Electrophoresis was performed on a native 7% polyacrylamide gel in TMN buffer with constant recycling of running buffer to avoid changes in pH. An autoradiogram of the gel is shown. Lane 1: labelled CopA$_F$ annealed to CopT by boiling–slow cooling. Lane 2: labelled CopA$_F$. Lane 3: labelled CopA$_F$ in kissing complex with CopT. Lane 4: unlabelled CopA$_F$ was allowed to form a complex with CopT, then labelled CopA$_F$ was added prior to electrophoresis. The details of the protocol are given in Materials and methods.

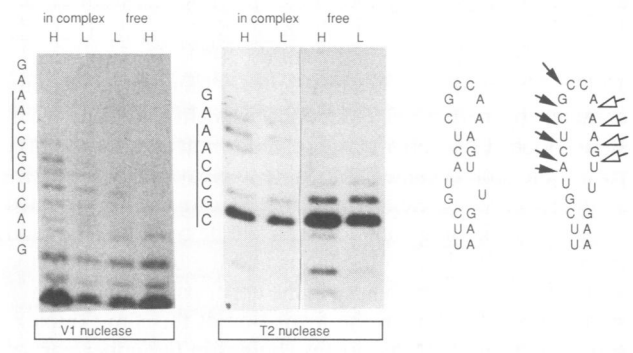

Fig. 7. Secondary structure probing of the upper stem–loop II region of CopA$_F$, free or in a kissing complex. The **left-hand panels** show autoradiograms of partial ribonuclease digests. The protocol for the formation of the kissing intermediate between CopA$_F$ and CopT as well as for the cleavage reactions is described in Materials and methods. H and L refers to high and low nuclease concentration, repectively [V_2: (H) 0.07 U; (L) 0.01 U; T_2: (H) 3 U; (L) 0.8 U]. Electrophoresis was done on 12% sequencing gels. Partial digests with ribonucleases T_1 and A, and phosphorylated deoxyoligonucleotides of known length, were used as markers (data not shown). The **right-hand panels** show the upper stem–loop II region of CopA RNA (or CopA$_F$). The secondary structure determined for free CopA RNA in solution (Wagner and Nordström, 1986) is indicated in the left-hand drawing. The right-hand drawing indicates changes occurring due to complex formation with CopT. The filled arrows indicate positions of duplex-dependent enhanced cleavages by V_1, the open arrows indicate enhanced cleavages by T_2.

We have performed an experiment in which both CopA$_E$ and CopA RNA were allowed to react with CopT simultaneously. Our results show that the reaction between CopA and CopT was inhibited by the CopA$_E$ molecule (Figure 5). We therefore conclude that CopA$_E$ can readily form a kissing complex with CopT although it is a poor substrate for complete duplex formation.

The kissing complex can be identified on native polyacrylamide gels

The experiment in Figure 5 showed that CopA$_E$ was capable of inhibiting CopA RNA binding to CopT, implying that such a species can form the postulated kissing complex. We had observed previously that this complex is unusually stable, with a half-life of >60 min (under standard conditions at 37°C; unpublished results). The least active species in forming the stable duplex with CopT was CopA$_F$ (Figure 2). This RNA does not form a stable duplex (detectable on denaturing gels) with CopT even after long incubations. We therefore tested whether the stability of the inferred kissing complex is sufficient for gel electrophoretic separation in native polyacrylamide gels in binding buffer. Figure 6 shows that a kissing complex between CopA$_F$ and CopT can be resolved from free CopA$_F$, and that this complex does not dissociate appreciably during the 7 h gel run, indicating a very small dissociation rate constant (see also accompanying paper). As a control, preformed kissing complex between unlabelled CopA$_F$ and CopT was challenged with labelled CopA$_F$. Lane 4 in Figure 6 shows that the labelled CopA$_F$ does not exchange with the unlabelled species.

The kissing complex involves base-pairing in the upper stem region of stem–loop II of CopA

We now studied the loop region of free CopA$_F$ and compared it to that of the same molecule in a kissing complex. The single-strand-specific ribonuclease T_2 and the double-strand-specific ribonuclease V_1 were used to probe 5′ end-labelled CopA$_F$, free or complexed with CopT RNA (see Materials and methods). The results in Figure 7 suggest that the short, double-stranded helix between the six-base loop and the upper bulge is opened upon the formation of the kissing complex. This is indicated by the accessibility to T_2 cleavage of four bases on the 3′ side of the loop which are not cleaved by T_2 in the free RNA. Likewise, CopA$_F$ becomes susceptible to the action of nuclease V_1 at 4–6 consecutive nucleotides, suggesting that these nucleotides on the 5′ side of the loop are hydrogen-bonded to the complementary bases in the CopT loop II. From the experiment in Figure 7 and additional cleavages (data not shown) we estimate that 6–8 nucleotides of CopA are involved in base-pairing to form the kissing complex, and that formation of this short inter-strand-helix proceeds by a disruption of the upper part of the helical stem in CopA (in the 5′ direction with respect to the CopA molecule).

Binding properties of an Inc RNA derived from plasmid P307

Plasmid P307 belongs to the IncFI group and is completely compatible with plasmid R1. However, P307 encodes an Inc RNA homologous to CopA RNA in all but seven bases located in the upper region of stem II and 2 bases located at the foot of the stem (Saadi *et al.*, 1987). This Inc RNA was used in binding experiments with CopT RNA of wild-type R1 sequence. The P307 Inc RNA bound CopT RNA very poorly (species L in Figure 2), in accordance with the fact that plasmids P307 and R1 are compatible *in vivo*. Furthermore, the P307 Inc RNA did not inhibit binding between CopA and CopT RNAs (data not shown). This is

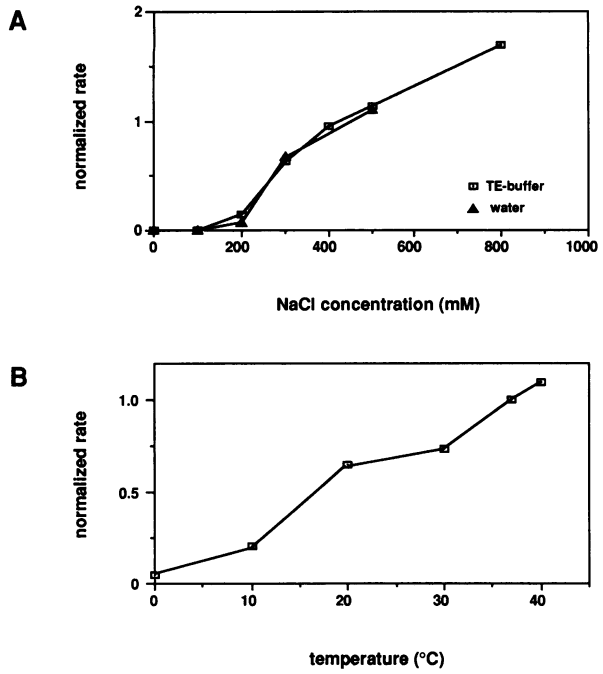

Fig. 8. Salt and temperature dependence of the binding rate between CopA RNA and CopT RNA. (**A**) Binding experiments, using radiolabelled $CopT_{322}$ and unlabelled CopA, were performed in H_2O or TE buffer at 37°C, in the absence or presence of various concentrations of NaCl. For each experiment the binding rates were calculated and expressed relative to the binding rate of the reactants in standard binding buffer. The concentration of CopA was 3 or 4 × 10^{-9} M and that of CopT was ~10- to 20-fold lower. (**B**) A binding experiment performed using labelled $CopT_{322}$ RNA (1.2 × 10^{-10} M) and unlabelled CopA (1.2 × 10^{-9} M) in TMN buffer. The reaction temperatures were varied as indicated and the calculated binding rates are expressed relative to the binding rate at 37°C.

somewhat surprising, since all bases in the Inc RNA loop II are complementary to the corresponding bases in CopT. However, the results presented above indicate that formation of a kissing complex between R1 CopA and R1 CopT involves not only the nucleotides of loop II but also several nucleotides in the upper part of stem II. Thus, it seems likely that the interaction between P307 Inc RNA and R1 CopT RNA does not result in the formation of a kissing complex, due to the mismatches in the upper stem regions.

Effects of salt and temperature upon the rate of complex formation

The breaking and reforming of hydrogen bonds, in a situation where intra-strand bonds are replaced by inter-strand bonds in an RNA duplex, bears similarities to cruciform formation in DNA. Even though the phenomena are not identical we expect the effects of salt concentration and temperature to be similar in both cases (for effects of salt and temperature in cruciform extrusion, see Lilley, 1985). We have therefore analysed the effects of salt and temperature on the binding reaction between CopA and CopT. In the absence of Mg^{2+} ions, binding was completely dependent on NaCl concentration. In water or TE buffer, without added NaCl or in the presence of 100 mM NaCl, the rate of binding was decreased at least 100-fold as compared to the rate in TMN buffer. At 200 mM NaCl the rate was still 10-fold lower than the rate of TMN, and continued to increase with increasing concentrations of NaCl (Figure 8). The rate of complex formation at different temperatures (in TMN buffer) was also determined. At 0°C and at 10°C it was about 10-fold and 5-fold lower than at 37°C. In the range between 20°C and 42°C it varied from 0.6- to 1.2-fold as compared with the rate at 37°C (Figure 8). These results suggest a mechanism which involves gradual intra-strand breakage and inter-strand reunion rather than a mechanism where a considerable unfolding of the molecules precedes duplex formation.

Discussion

Tomizawa (1984) proposed a stepwise model for the formation of a duplex between RNAI and RNAII from plasmid ColE1. The first step is the formation of a kissing complex between the complementary loops in the two highly structured RNA molecules. The subsequent steps start with base-pairing between the 5′ single-stranded end of RNAI and the complementary region of RNAII and eventually lead to the formation of a complete RNA duplex. In the present communication, we conclude that the interaction between CopA and CopT RNAs from plasmid R1 also takes place in (at least) two experimentally distinguishable steps. The data in Figure 5 show that an RNA molecule derived from CopA lacking the 5′-most 31 nucleotides, is ineffective in forming an RNA–RNA duplex with CopT RNA, but very effective in inhibiting binding between full length CopA and CopT. Thus, stem–loop II was shown to be the structure in CopA necessary and sufficient for the first step. This is in agreement with earlier data, from both *in vivo* and *in vitro* experiments, which demonstrate that the specificity of the initial recognition reaction resides in loop II in CopA RNA and the corresponding loop in CopT RNA (Rosen *et al.*, 1980; Danbara *et al.*, 1981; Stougaard *et al.*, 1981; Brady *et al.*, 1983; Givskov and Molin, 1984; Persson *et al.*, 1988). The results shown in Figure 6 identify the kissing complex by native gel electrophoresis. In addition, when stem–loop II of CopA was allowed to interact with CopT, part of the loop became susceptible to cleavage by a double-strand specific nuclease. The CopA–CopT double-stranded region in the kissing complex seems to be 6–8 bp long, and involves nucleotides in the loops and the upper part of the stems (Figure 7). The bulges in the upper part of the CopA/T stem II probably favour the formation of this short duplex; several antisense RNAs possess bulges in the upper part of their major stems (see review by Simons and Kleckner, 1988). Experiments to test the involvement of the bulges in the reaction pathway are in progress in our laboratory.

It was shown that a single-stranded region in CopA, and its counterpart in CopT, are essential for the second step(s) to occur efficiently. In natural CopA this single-stranded region is located 5′ to stem–loop II. However, we showed that the location of the single-stranded region with respect to stem–loop II is not important. An artificial CopA–CopT pair, in which the CopA 5′ region and its CopT 3′ counterpart were deleted and replaced by an extended CopA 3′ region and its CopT 5′ counterpart, respectively, formed a complete duplex very efficiently. The length of the single-stranded region is important, though, since the wild-type 3′ tail could not substitute for the 5′ region in those CopA molecules that were shortened on the 5′ side. The data in Figure 2 indicate some correlation between the length of the shortened CopA 5′ region and the rate of duplex formation.

The 5′ single-stranded region of CopA is bordered by stem–loop I on its 5′ side. Our results indicate that the latter

structure is not needed for binding between CopA and CopT, since it can be disrupted with no effect on binding rates (see molecule C in Figure 2). Furthermore, there exists a natural antisense RNA (derived from an IncB miniplasmid; Praszkier *et al.*, 1989) which resembles CopA RNA in sequence and predicted structure except that it lacks a stem–loop at its 5′-most end, also indicating that such a structure is not required for antisense RNA function. This situation is distinctly different from the ColE1 system, where an intact 5′ tail is required for efficient duplex formation (Tomizawa, 1984). The results presented here do not exclude the possibility that the interaction leading to complete duplex formation might start at the single-stranded nucleotides of loop I in wild-type CopA. However, we do not favour this hypothesis because of topological constraints imposed on the RNA molecules if pairing is to start in a loop structure.

The pairing reactions between P22 *sar* RNA, Tn10 RNA-OUT, ColE1 RNAI, and R1 CopA RNA, and their respective target RNAs have been studied *in vitro* (Liao *et al.*, 1987; Kittle *et al.*, 1989; Tomizawa and Itoh, 1981; Tomizawa 1984, 1985; Persson *et al.*, 1988). The apparent second-order rate constant for the binding between antisense RNA and target RNA has been reported in three of the cases: it was 3×10^5 M^{-1} s^{-1} for the RNA-OUT–RNA-IN pair, 7.1×10^5 M^{-1} s^{-1} for the RNAI–RNAII pair and 1×10^6 M^{-1} s^{-1} for the CopA–CopT pair. It is evident that the overall rate of duplex formation is of the same order of magnitude, even though the three antisense RNAs/target RNAs differ in structure, and to some extent in the mechanism of pairing. RNA-OUT forms a structure composed of a stem domain topped by a loop domain, which contains two small duplex regions and a six-nucleotide loop. The 5′ end of RNA-IN, which lacks ordered secondary structure, is complementary to the top of the loop of RNA-OUT and complementarity extends down the 5′ side of RNA-OUT, over a total stretch of 35 nucleotides. Initiation of base-pairing at the 5′ end of RNA-IN is required for complete duplex formation, and may also be important for the formation of an initial complex. RNAI and RNAII, on the other hand, contain three complementary stem–loop structures, and the interaction between the molecules initiates at these loops. Persistent pairing, however, initiates in a separate region involving the 5′ end of RNAI, and base-pairing then propagates towards the 3′ end. CopA RNA contains two stem–loop structures, but it is only the major stem and its six-nucleotide loop (stem–loop II) that are involved in the initial interaction. Complete duplex formation probably initiates in the single-stranded region located between stem–loops I and II, propagating towards both ends of the CopA RNA. Thus, the CopA–CopT pairing mechanism resembles the RNAI–RNAII mechanism in that there is a structural separation between the region at which the initial contact takes place and the region at which complete pairing initiates. This structural separation is probably due to topological constraints that make complete pairing starting from the paired loops impossible. An alternative mechanism for duplex formation starting from the loops would require melting of the stems prior to complete duplex formation. This is, however, energetically unfavourable, and as discussed below, not compatible with our experimental data. We suggest that the reason for the similarity of the apparent second-order binding rate constants in the various antisense RNA–target RNA reactions lies in the primary interaction. This first interaction involves collision between the RNA molecules, as was suggested by Tomizawa (1985), and as is shown in our companion paper. The formation of the kissing complex is the rate-limiting step in the formation of a complete duplex (Persson *et al.*, accompanying paper).

Both CopA and CopT RNA molecules have a high degree of secondary structure, with the very stable stem–loop II as the most prominent feature. At some step during complex formation between the two RNAs, intra-strand unfolding takes place, and an inter-strand RNA helix is formed. Such a process, where structural rearrangements entail the breaking of many hydrogen bonds and the formation of a new stretch of hydrogen-bonded nucleotides, is reminiscent of DNA cruciform extrusion. Lilley (1985) has proposed two mechanisms for the transformation of a double helical inverted repeat into a cruciform. One mechanism invokes the formation of a large unpaired region prior to the formation of a complete cruciform. In the other mechanism a relatively small unpaired region forms a protocruciform which is gradually converted to the fully extended cruciform by a number of branch migration steps. The reactions differ dramatically in their temperature and salt concentration dependences. The results of binding experiments performed at different salt concentrations and temperatures (Figure 8) suggest that the binding reaction between CopA and CopT occurs through continuous breaking and formation of hydrogen bonds rather than through a general unfolding of the molecules prior to complete duplex formation.

Plasmids with different copy number control systems are able to coexist stably in the same cell, i.e. they are compatible. If copy number control is exerted by an antisense RNA, the antisense RNAs of compatible plasmids must have different sequence specificities which preclude binding of the heterologous target RNAs. Plasmid R1 and plasmid P307 are compatible (Saadi *et al.*, 1987). However, close examination of their regulatory regions reveals a high degree of homology. The antisense RNAs (*inc* RNAs) encoded by these plasmids differ by only nine nucleotides in the major stem–loop. Seven of the differing nucleotides are located close to the loop, whereas the six nucleotides of the loop are conserved. We propose that the Inc RNA of P307 (Figure 2), even though it might recognize the loop II sequence of CopT RNA, is not able to form a kissing intermediate with CopT RNA due to the differences in the upper part of the major stem. This would indicate that stable duplex formation is only possible if the RNAs are initially held together by a relatively long-lived, but likewise transient, interaction. An alternative explanation for the lack of binding between P307 Inc RNA and CopT RNA is that P307 Inc RNA might adopt an altogether different structure unfavourable for proper binding to the wild-type R1 CopT RNA.

All natural antisense RNAs discovered so far contain stem–loop structures and their targets often have complementary stem–loops. What is the function of these structures? Previously, we proposed three possible alternatives (Persson *et al.*, 1988): (i) relatively large regions of double-strandedness restrict the regions that are free to interact to a few single-stranded loops, thereby increasing the fidelity of the interaction, (ii) stem–loops affect the rate by which duplex formation can occur, either positively or negatively, and (iii) stem–loops occur simply because of the need to terminate transcription, and are unimportant for

the function of the RNAs. It is clear from the results presented in this communication that a highly ordered structure greatly increases the rate of duplex formation at physiological temperatures; compare e.g. molecules J and K in Figure 2. Conceivably, one of the reasons why artificial antisense RNAs have been used with variable success is that the structures of these inhibitors are sometimes far from optimal for high reaction rates. A detailed knowledge of the structures that favour rapid and specific binding will undoubtedly be important to guide the search for effective artificial antisense RNA control systems.

Materials and methods

Plasmids

Plasmid pGW543 is a derivative of pHST0 (Jobling *et al.*, 1988), with the 262 bp *Sau*3AI fragment encoding CopA−CopT (see Figure 1) inserted into the *Bgl*II site. The orientation of the *Sau*3AI fragment is such that transcripts directed from the SP6 promoter are in the sense orientation with respect to CopA.

Plasmid pGW640 is a pSP64 (Melton *et al.*, 1984) derivative containing R1 sequences oriented such that transcription from the SP6 promoter yields RepA mRNA transcripts in the sense orientation (Persson *et al.*, 1988).

Plasmid pGW643 is a pSP64 derivative containing the 262 bp *Sau*3AI fragment encoding CopA−CopT inserted into the *Bam*HI site. The orientation of the *Sau*3AI fragment is such that transcripts directed from the SP6 promoter are in the sense orientation with respect to CopA (Wagner *et al.*, 1987).

Plasmid pGW401 is a pUC19-replicon (Yanisch-Perron *et al.*, 1985), which contains a T7 promoter in front of sequences coding for stem−loop II of CopA RNA. The pUC19 vector was cleaved with *Eco*RI and *Hin*dIII, then two double-stranded synthetic deoxyoligonucleotides were ligated to the vector DNA. The first two complementary single-stranded oligonucleotides, 5′-dAATTCGAAAT TAATACGACT CACTA TAG-3′ and 5′-dCTATA GTGAG TCGTA TTAAT TTCG-3′, contain a bacteriophage T7 promoter sequence and have the 5′-overhanging nucleotides of an *Eco*RI site at one end, while the other end is blunt. The second two complementary single-stranded oligos, 5′-dGGCCCCGTA ATCTTTTCGT ACTCGC-CAAA GTTGAAGAAG ATTATCGGGG TTTAAATGCA GGTA-3′ and 5′-dAGCTTACCTG CATTTAAACC CCGATAATCT TCTTCAACTT TGGCGAGTAC GAAAAGATTA CCGGGGCC-3′, yield blunt ends and four 5′-protruding nucleotides compatible with *Hin*dIII sites. Ligation recreated both the *Eco*RI and *Hin*dIII sites and brought the T7 promoter in front of the sequence coding for the stem−loop II of CopA RNA. The construct was verified by DNA sequencing. Cleavage with either *Dra*I or *Bsp*MI yields linear DNA suitable for T7 transcription of the CopA species F (see Figure 2). The oligonucleotides were synthesized in our laboratory, using a Pharmacia LKB Gene Assembler Plus DNA synthesizer.

Plasmid pSS288 is a pBR322 derivative containing the Inc RNA-coding region of P307 (Saadi *et al.*, 1987).

Preparation of CopA RNA and CopT RNA

CopA RNA was synthesized *in vitro* using plasmid pGW640 as a template. Plasmid pSS288 was utilized to transcribe P307 Inc RNA. Transcription and gel purification were performed as previously described for CopA RNA (Persson *et al.*, 1988). To synthesize the extended CopAs we used either plasmid pGW543, linearized with *Hin*dIII, or plasmid pGW643, linearized with *Ava*I, as templates. The transcripts were 272 or 305 nucleotides long, respectively. These CopA derivatives were transcribed from an SP6 promoter using a protocol previously described (Persson *et al.*, 1988). CopT RNA was transcribed using *Bst*NI linearized pGW640 as template, yielding a 322 nucleotide long transcript. Linearization of pGW640 with *Nci*I yielded $CopT_{169}$. Transcription and purification protocols were as described in Persson *et al.* (1988).

Binding between CopA RNA and CopT RNA

Binding experiments were performed as described earlier (Persson *et al.*, 1988). The standard binding buffer (TMN buffer) contains 20 mM Tris−OAc, pH 7.5, 10 mM $Mg(OAc)_2$, and 100 mM NaCl. The experiments were performed at 37°C unless otherwise stated, concentrations of reactants are given in the figure legends. The salt dependence experiments were performed such that CopA and CopT RNAs were dissolved in either water or TE buffer (10 mM Tris−HCl, pH 7.9, 1 mM Na_2EDTA) after purification. The final salt concentration was adjusted by adding an appropriate volume of 1 M or 2 M NaCl directly into the reaction tube.

Preparation of CopA RNA segments

Cleaved CopA RNA molecules were obtained by annealing octameric or dodecameric oligodeoxyribonucleotides complementary to different regions of CopA RNA and subsequently cleaving with RNase H, as described in Öhman and Wagner (1989). The sequences of the oligodeoxyribonucleotides were: 5′-dCCCACTTA-3′ (KabiGen, Stockholm) which gave the derivatives E, G, J and K, and 5′-dCCAACAATTCAG-3′ (synthesized in our laboratory) which gave the derivative C (see Figure 2). Derivative D appears spontaneously in purifiable amounts upon prolonged incubation in the absence of oligonucleotides and RNase H. The cleaved CopA products were chromatographed on 6% or 8% polyacrylamide gels containing 7 M urea and each fragment was purified as described previously for CopA RNA (Persson *et al.*, 1988). The new 5′-endpoints of the obtained CopA derivatives were mapped by reverse transcription using an oligodeoxyribonucleotide complementary to loop II of CopA as primer. The procedure has been described in Öhman and Wagner (1989). The sequence markers were synthesized from suitable plasmids carrying the subcloned replication control region of R1, sequenced with the Klenow fragment of DNA polymerase I (BioLabs) in the Gem Seq K/RT system (Promega). The primer used was identical to the one used for reverse transcription of the CopA derivatives.

Preparation of $CopA_F$ RNA

$CopA_F$ RNA was transcribed using T7 RNA polymerase and plasmid pGW401, linearized by cleavage with *Dra*I, as template. An incubation mixture for the synthesis of uniformly ^{32}P-labelled $CopA_F$ RNA contained in 50 μl: 40 mM Tris−HCl, pH 7.9, 6 mM $MgCl_2$, 20 mM DTT, 2 mM spermidine, 0.01% Triton X-100, 0.05 mM UTP, 0.05 mM GTP, 0.2 mM ATP, 0.2 mM CTP, 50 μCi [α-^{32}P]UTP (New England Nuclear, 800 Ci/mmol), 20 U of RNasin (Promega Biotech), 40 U of T7 RNA polymerase (New England Biolabs) and ~1 μg of linearized plasmid DNA. An aliquot of the reaction mixture was withdrawn for determination of specific radioactivity. The remaining mixture was incubated at 37°C for 2 h, then 3 U of RQ1 DNase (Promega Biotech) were added and incubation was continued for another 15 min. Subsequent purification, including gel electrophoresis and elution, were performed as above.

Analysis of the kissing complex on native polyacrylamide gel

The protocol for formation and analysis of the kissing complex on a native polyacrylamide gel was as follows. The binding reaction was initiated by mixing $CopA_F$ labelled (0.3 pmol, 50 000 c.p.m./pmol) or unlabelled (1.4 pmol) with 0.7 pmol CopT in 10 μl of TMN. The CopT species is analogous to the one used in the binding experiments in Figure 2. The incubation was at 37°C for 5 min. For the incubation corresponding to lane 4 in Figure 6, the labelled $CopA_F$ was added after formation of the unlabelled $CopA_F$−CopT complex. When indicated (see Figure legend), CopT was omitted. For the incubation corresponding to lane 1 in Figure 6, labelled $CopA_F$ and unlabelled CopT were heated to 95°C and slowly cooled to anneal the RNAs to a complete duplex. 5 μl of TMN buffer containing 30% glycerol, 0.05% bromophenol blue and 2 μg of tRNA was added to all samples, and electrophoresis was performed on a 7% polyacrylamide gel in TMN buffer at 4°C. A stable pH was ensured throughout the 7 h run (50 V/20 mA) by constant recycling of the buffer between the reservoirs. The gel was additionally cooled by placing a fan in front of the gel plates. The gel was autoradiographed on Kodak X-omat S film overnight.

Secondary structure probing of $CopA_F$ RNA, free and in complex with CopT RNA

For the experiment in Figure 7, $CopA_F$ RNA was synthesized as above, except that [γ-^{32}P]GTP (New England Nuclear) was used in order to 5′ end-label the transcripts. Unlabelled $CopA_F$ RNA was prepared in parallel by the omission of the radioactive nucleoside triphosphate. Samples of free and complexed labelled $CopA_F$ RNA were obtained in the following way: CopT RNA (1.5×10^{-8} M) was preincubated in TMN buffer either with unlabelled $CopA_F$ (~5×10^{-8} M) (A) or labelled $CopA_F$ (1×10^{-9} M) (B) at 37°C for 3 min. Labelled $CopA_F$ and unlabelled $CopA_F$ (at the same concentrations as above) were then added to mixtures A and B, respectively. Due to the slow dissociation rate of the kissing complex, this protocol ensures that both incubations contain the same amount of labelled and unlabelled RNAs, but only in tube B will the kissing complex be formed with the labelled $CopA_F$ species. Aliquots were diluted 4-fold into buffer containing either ribonuclease T_2 (Bethesda Research Laboratories) or V_1 (Pharmacia), and the incubation was continued for 5 min. After addition of an equal

volume of sample buffer (90% formamide, 10 mM Na_2EDTA, 0.025% xylene cyanol, 0.025% bromophenol blue), the samples were boiled and electrophoresed on sequencing gels.

Acknowledgements

We thank Lena Möller for expert technical assistance. We are grateful to Dr Måns Ehrenberg for critical reading of the manuscript. The technical assistance of Solveig Andersson is also gratefully acknowledged. This work was supported by the Swedish Natural Science Research Council and the Swedish Cancer Society.

References

Blomberg,P., Wagner,E.G.H. and Nordström,K. (1990) *EMBO J.*, **9**, 2331–2340.
Brady,G., Frey,J., Danbara,H. and Timmis,K.N. (1983) *J. Bacteriol.*, **154**, 429–436.
Cole,S.T. and Honoré,N. (1989) *Mol. Microbiol.*, **3**, 715–722.
Danbara,H., Brady,G., Timmis,J.K. and Timmis,K.N. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 4699–4703.
Dempsey,W.B. (1989) *J. Bacteriol.*, **171**, 2886–2888.
Gerdes,K., Helin,K., Wiborg Christensen,O. and Løbner-Olesen,A. (1988) *J. Mol. Biol.*, **203**, 119–129.
Givskov,P.J. and Molins,S. (1984) *Mol. Gen. Genet.*, **194**, 286–292.
Green,P.J., Pines,O. and Inouye,M. (1986) *Annu. Rev. Biochem.*, **55**, 569–597.
Ho,Y.S. and Rosenberg,M. (1985) *J. Biol. Chem.*, **260**, 11838–11844.
Hoopes,B.C. and McClure,W.R. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3134–3138.
Jobling,S.A., Cuthbert,C.M., Rogers,S.G., Fraley,R.T. and Gehrke,L. (1988) *Nucleic Acids Res.*, **16**, 4483–4498.
Kim,K. and Meyer,R.J. (1986) *Nucleic Acids Res.*, **14**, 8027–8046.
Kittle,J.D., Simons,R.W., Lee,J. and Kleckner,N. (1989) *J. Mol. Biol.*, **210**, 561–572.
Krinke,L. and Wulff,D.L. (1987) *Genes Dev.*, **1**, 1005–1013.
Kumar,C.C. and Novick,R.P. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 638–642.
Liao,S.-M., Wu,T., Chiang,C.H., Susskind,M.M. and McClure,W.R. (1987) *Genes Dev.*, **1**, 197–203.
Light,J. and Molin,S. (1983) *EMBO J.*, **2**, 93–98.
Lilley,D.M.J. (1985) *Nucleic Acids Res.*, **13**, 1443–1465.
Melton,D.A., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zinn,K. and Green,M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
Mizuno,T., Chou,M.-Y. and Inouye,M. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1966–1970.
Nordström,K., Molin,S. and Light,J. (1984) *Plasmid*, **12**, 17–90.
Öhman,M. and Wagner,E.G.H. (1989) *Nucleic Acids Res.*, **17**, 2557–2579.
Okamoto,K. and Freundlich,M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5000–5004.
Patel,I. and Bastia,D. (1987) *Cell*, **51**, 455–462.
Persson,C., Wagner,E.G.H. and Nordström,K. (1988) *EMBO J.*, **7**, 3279–3288.
Persson,C., Wagner,E.G.H. and Nordström,K. (1990) *EMBO J.*, **9**, 3777–3785.
Praszkier,J., Bird,P. Nikoletti,S. and Pittard,J. (1989) *J. Bacteriol.*, **171**, 5056–5064.
Rosen,J., Ryder,T., Inokuchi,H., Ohtsubo,H. and Ohtsubo,E. (1980) *Mol. Gen. Genet.*, **179**, 527–537.
Rownd,R.H., Womble,D.D., Dong,X., Luckow,V.A. and Wu,R.P. (1985) In Helinski,D.R., Cohen,S.N., Clewell,D.B., Jackson,D.A. and Hollander,A. (eds), *Plasmids in Bacteria*. Plenum Publishing Co., New York, pp. 335–354.
Ryder,T.B., Davison,D.B., Rosen,J.J., Ohtsubo,E. and Ohtsubo,H. (1982) *Gene*, **17**, 299–310.
Saadi,S., Maas,W.K., Hill,D.F. and Bergqquist,P.L. (1987) *J. Bacteriol.*, **169**, 1836–1846.
Simons,R.W. and Kleckner,N. (1983) *Cell*, **34**, 683–691.
Simons,R.W. and Kleckner,N. (1988) *Annu. Rev. Genet.*, **22**, 567–600.
Stougaard,P., Molin,S., Nordström,K. and Hansen,F.G. (1981) *Mol. Gen. Genet.*, **181**, 116–122.
Stougaard,P., Light,J. and Molin,S. (1982) *EMBO J.*, **1**, 323–328.
Tomizawa,J. (1984) *Cell*, **38**, 861–870.
Tomizawa,J. (1985) *Cell*, **40**, 527–535.
Tomizawa,J. and Itoh,T. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 6096–6100.
Tomizawa,J., Itoh,T., Selzer,G. and Som,T. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 1421–1425.
Wagner,E.G.H. and Nordström,K. (1986) *Nucleic Acids Res.*, **14**, 2523–2538.
Wagner,E.G.H., von Heijne,J. and Nordström,K. (1987) *EMBO J.*, **6**, 515–522.
Yanisch-Perron,C., Viera,J. and Messing,J. (1985) *Gene*, **33**, 103–119.

Received on July 3, 1990

Note added in proof

Since the submission of our manuscript, Tomizawa has published a paper (Tomizawa,J. (1990) *J. Mol. Biol.*, **212**, 683–694) in which he studies the system that controls replication of plasmid ColE1. He identified intermediates in the formation of an RNA duplex between the antisense RNA I and its target RNA II. He was also able to demonstrate that the formation of the RNA duplex takes place in several steps and to determine the kinetic constants for the formation of the various intermediates.