CopA and CopT: The Perfect RNA Couple

BY

JACOBA G. SLAGTER-JÄGER

ACTA UNIVERSITATIS UPSALIENSIS
UPPSALA 2003
Dissertation presented at Uppsala University to be publicly examined in C10:305, BMC, Uppsala, Thursday, June 5, 2003 at 10:00 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

Antisense RNAs regulate gene expression in many bacterial systems. The best characterized examples are from prokaryotic accessory elements such as phages, plasmids and transposons. Many of these antisense RNAs have been identified as plasmid copy number regulators where they regulate the replication frequency of the plasmid by negative feedback. Instability and fast binding kinetics is crucial for the regulatory efficiency of these antisense RNAs.

In this thesis, the interaction of the cis-encoded antisense RNA CopA with its target CopT was studied in detail using in vivo reporter gene fusion expression and different in vitro methods, such as surface plasmon resonance, fluorescence resonance energy transfer, and gel-shift assays.

Formation of inhibitory complexes differs from simple hybridization reactions between complementary strands. E.g., the binding pathway of CopA and CopT proceeds through a hierarchical order of steps. It initiates by reversible loop-loop contacts, resulting in a helix nucleus of two or three base pairs. This is followed by rapid unidirectional helix progression into the upper stems, resulting in a four-way helical junction structure. It had been suggested that the loop of CopT carries a putative U-turn, a structure first found in tRNA anticodon loops. We showed that this putative U-turn is one of the structural elements of CopA/CopT required to achieve fast binding kinetics. Furthermore, the hypothetical U-turn structure determines the direction of helix progression when the kissing complex progresses to a four-way helical junction structure. Another structural element in CopT is the helical stem adjacent to the recognition loop. This stem is important to present the recognition loop appropriately to provide a scaffold for the U-turn.

Furthermore, the role of protein Hfq in the interaction of antisense/target RNA was investigated, since several trans-encoded antisense RNAs had been shown to need this protein to exert their function. In contrast, studies of two cis-encoded antisense RNA systems showed that these antisense RNAs do not rely on Hfq for activity. In this study it was also shown that MicF, a trans-encoded antisense RNA which is dependent on Hfq, is greatly stabilized by this protein.

Keywords: antisense RNA, plasmid, U-turn, four-way junction, RNA-RNA interaction, Hfq

Jacoba G. Slagter-Jäger, Department of Cell and Molecular Biology, Box 596, BMC, Uppsala University, SE-75124 Uppsala, Sweden

© Jacoba G. Slagter-Jäger 2003

ISSN 1104-232X
ISBN 91-554-5657-X
urn:nbn:se:uu:diva-3465 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-3465)
This thesis is based on the following publications, which are referred to by their Roman numerals:


Reprints have been made with permission from the publishers.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ABBREVIATIONS</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREFACE</td>
<td>5</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>6</td>
</tr>
<tr>
<td>What is antisense RNA?</td>
<td>6</td>
</tr>
<tr>
<td>Antisense RNA involved in regulation of plasmid copy number</td>
<td>7</td>
</tr>
<tr>
<td>The case of Plasmid R1: regulation through inhibition of translation</td>
<td>9</td>
</tr>
<tr>
<td>Other antisense RNA regulated plasmid copy number systems</td>
<td>12</td>
</tr>
<tr>
<td>ColIb-P9: inhibition of pseudoknot formation</td>
<td>12</td>
</tr>
<tr>
<td>ColE1-relatives: inhibition of primer formation</td>
<td>14</td>
</tr>
<tr>
<td>Plasmid pIP501 and pT181: transcription attenuation</td>
<td>14</td>
</tr>
<tr>
<td>Antisense RNA systems encoded by the bacterial chromosome</td>
<td>16</td>
</tr>
<tr>
<td>MicF</td>
<td>16</td>
</tr>
<tr>
<td>OxyS</td>
<td>16</td>
</tr>
<tr>
<td>DsrA</td>
<td>17</td>
</tr>
<tr>
<td>RprA</td>
<td>17</td>
</tr>
<tr>
<td>The role of proteins in antisense/target RNA binding</td>
<td>17</td>
</tr>
<tr>
<td>hfq</td>
<td>18</td>
</tr>
<tr>
<td>Kinetics of binding of two hybridizing nucleic acids</td>
<td>19</td>
</tr>
<tr>
<td>Non-structured RNAs</td>
<td>19</td>
</tr>
<tr>
<td>Folded RNAs</td>
<td>20</td>
</tr>
<tr>
<td>Important structures for fast kinetics</td>
<td>21</td>
</tr>
<tr>
<td>Hairpin loops</td>
<td>21</td>
</tr>
<tr>
<td>U-turn</td>
<td>22</td>
</tr>
<tr>
<td>Bulges</td>
<td>23</td>
</tr>
<tr>
<td>Single stranded regions</td>
<td>24</td>
</tr>
<tr>
<td>Helical junctions</td>
<td>24</td>
</tr>
<tr>
<td>Methods for studying antisense-target RNA in real time</td>
<td>24</td>
</tr>
<tr>
<td>Surface plasmon resonance</td>
<td>25</td>
</tr>
<tr>
<td>Fluorescence resonance energy transfer</td>
<td>26</td>
</tr>
<tr>
<td>THE PRESENT INVESTIGATION</td>
<td>27</td>
</tr>
<tr>
<td>Scope of this investigation</td>
<td>27</td>
</tr>
<tr>
<td>Extended complex needed for rapid pairing (paper I)</td>
<td>28</td>
</tr>
<tr>
<td>Preference of directionality (paper I)</td>
<td>29</td>
</tr>
</tbody>
</table>
The U-turn determines directionality of helix progression (paper IV) ..........................................................................................................................29
The upper stem segment functions as a scaffold (paper IV) ........31
The role of Hfq (paper V) ............................................................................31
SPR studies (paper II) ................................................................................32
FRET studies (paper III) ................................................................................33

DISCUSSION ........................................................................................................35
The binding of CopA and CopT .................................................................35
Real time kinetic studies ........................................................................36
The role of Hfq in antisense regulation .....................................................37

CONCLUDING REMARKS AND FUTURE PROSPECTS ...............40

ACKNOWLEDGMENTS .................................................................................41

REFERENCES ..............................................................................................43
### ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>adenosine</td>
</tr>
<tr>
<td>2-AP</td>
<td>2-aminopurine</td>
</tr>
<tr>
<td>C</td>
<td>cytosine</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>et al.</td>
<td>and others</td>
</tr>
<tr>
<td>FRET</td>
<td>Fluorescence resonance energy transfer</td>
</tr>
<tr>
<td>G</td>
<td>guanosine</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>N</td>
<td>any nucleotide</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotide(s)</td>
</tr>
<tr>
<td>P</td>
<td>purine</td>
</tr>
<tr>
<td>R</td>
<td>pyrimidine</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>SD</td>
<td>Shine-Dalgarno</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface plasmon resonance</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>U</td>
<td>uracil</td>
</tr>
</tbody>
</table>


There has been an increased amount of interest of the scientific community in small regulatory RNAs. During the last few years several research groups have performed and succeeded with the identification of small regulatory RNAs. Chromosomally encoded small RNAs have been isolated from both prokaryotes and eukaryotes. This, and the discovery of the RNA Interference (RNAi) phenomenon and its applications, has put small regulatory RNAs in the spotlight of research in molecular biology. However, the first naturally occurring antisense RNA was discovered about 20 years ago. This antisense RNA belongs to the large group of antisense RNAs encoded by bacterial accessory elements, like plasmids, phages or transposons. This thesis focuses on one of these antisense RNAs; CopA and its target CopT, involved in regulation of copy number of plasmid R1.

Natural antisense RNAs need to be metabolically unstable and fast binders to be able to exert their regulatory functions. They are characterized by distinct secondary and tertiary structures and particular motifs appear to correlate with regulatory performance.

This thesis summarizes the studies on the early steps of the binding pathway of the antisense RNA CopA and its target CopT, to elucidate how the initial stages of the binding takes place, and which structures are important and involved in fast binding. As a strategy, we designed different mutants of which we investigated the binding kinetics in vivo and in vitro. Furthermore, different techniques for studying the interaction kinetics were investigated.
INTRODUCTION

What is antisense RNA?

Antisense RNAs are RNA molecules that are complementary or partial complementary to target RNAs. Complementarity ensures that antisense and target RNA can bind to each other, forming a partial or complete RNA duplex. In most cases this results in inhibition of target RNA function, e.g. if the target RNA is an mRNA, then RNA duplex formation in a functionally important region may inhibit translation. This implies that antisense RNAs are inhibitors of gene expression at the post-transcriptional level.

Apart from blocking translation, antisense RNAs can use other mechanisms for regulation of gene expression. They can induce premature target RNA termination (Brantl and Wagner, 1994), facilitate mRNA decay (Krinke and Wulff, 1990), or inhibit formation of a mature replication primer (Masukata and Tomizawa, 1986).

Naturally occurring antisense RNAs are small, highly structured, unstable, diffusible RNA molecules that are generally not translated. Antisense RNAs as regulators of gene expression are rather widespread, from a few cases in eukaryotes and archea to many antisense RNA regulated systems in bacteria. The majority of known cases are found in accessory genetic elements such as plasmids, transposons and phages (Wagner and Simons, 1994; Wagner et al., 2002). However, lately many small chromosomally encoded RNAs from prokaryotes have been discovered in searches for non-coding RNAs (Argaman et al., 2001; Wassarman et al., 2001; Rivas et al., 2001; Chen et al., 2002). Although the function of most of these small RNAs is still unknown, at least one of them has been implicated to function as an antisense RNA (Masse and Gottesman, 2002).

Antisense RNAs are extremely efficient inhibitors of gene expression. Sequence and structure of antisense and target RNA are important to promote high pairing rates and have evolved to optimize the regulatory functions of the antisense RNA. Generally, antisense RNAs contain one or more stem-loop structures that are complementary to sequences in target RNAs. Interaction with a target most often initiates at specific loop sequences of both antisense RNA and target, or between a loop and a single stranded segment of the
RNAs. Bulged-out nucleotides are functionally important, and single-stranded regions are important for the pairing process.

In all antisense RNA controlled mechanisms, control is kinetic. The rate of antisense RNA binding determines the fraction of the target RNA that becomes functionally inactivated. In *Escherichia coli*, there are essentially two types of antisense RNAs that show the highest *in vitro* binding rate constant of $10^6 \text{ M}^{-1}\text{s}^{-1}$. The high copy number plasmids including pMB1 and ColE1 encode one type of such antisense RNAs. These antisense RNAs have sizes of about 100 nt, form three stem-loops with a single-stranded 5’ leader, and bind to the preprimer RNAs for the leading strand DNA synthesis, thereby inhibiting formation of mature primer RNAs. The other type of of rapid binders are encoded by low-copy-number plasmids including the members of IncFII, IncIalpha and IncB groups. These antisense RNAs are 70-100 nt long, fold essentially into a single large stemloop of 50 nt and control the expression of the plasmid replication initiator genes at the translational level.

**Antisense RNA involved in regulation of plasmid copy number**

Plasmids are extra-chromosomal elements consisting of circular or linear double stranded DNA. They are found in bacteria and in some eukaryotic cells like plant and yeast cells. Some plasmids consist of only 1000 base pairs, while others can have the size of several hundred thousand base pairs.

They are self-replicating entities and those carrying transfer genes are often capable of cell to cell transfer across species and generic boundaries. Plasmids carry several genes needed for their survival, but they also rely on proteins encoded by the chromosome of their host. Furthermore, they may carry genes conferring resistance to several kinds of antibiotics to the bacterial cell carrying the plasmid. Other kinds of plasmids may carry genes encoding metabolizing enzymes or for example plant infecting compounds enabling the bacterial host to survive and grow on specific metabolites or in a specific environment.

Some plasmids display a very narrow host range whereas others are promiscuous. Plasmids that belong to the same incompatibility group share the same copy number control elements. Due to this, plasmids belonging to the same incompatibility group cannot be maintained in
the same host cell, which results in loss of one of the plasmids from the cell.

Plasmid copy number control is homeostatic. Fluctuations in copy number are adjusted by changing the replication frequency of the plasmid. A plasmid can "measure" its copy number in the cell, and use this information to adjust the replication frequency of the plasmid. Plasmids can use different devices to accomplish this: some plasmids use iterons (short repetitive DNA sequences) for control (see review by Chattoraj, 2000), others use antisense RNAs, and a third group uses an antisense RNA in combination with a protein.

Many of the prokaryotic antisense RNAs have been identified as plasmid copy number regulators where they regulate replication frequency of the plasmid by negative feedback. In the case of ColE1, the antisense RNA I inhibits formation of primer RNA for leading strand synthesis (see below), and in the case of pT181-, IncFII-, IncIalpha/IncB-type and other types of plasmids, antisense RNAs control the synthesis of plasmid-encoded replication initiator proteins.

The antisense RNAs act here as key regulators. They are metabolically unstable and constitutively transcribed. As a result, their intracellular concentration is correlated with gene dosage (= plasmid copy number). This results in a balanced, defined copy number in individual cells by means of a negative feedback loop: if copy numbers rise, antisense RNA concentration is increased, inhibition of the replication function is increased and therefore replication of the plasmid decreases. In the same way, at too low copy number, replication is allowed to occur more often.

Apart from the basic replicon which is defined as the region containing all genes and sites required for wild-type copy number control (Rosen et al., 1981), additional systems are encoded by many plasmids, in particular those that have a low copy number, to assure a stable inheritance. The partitioning system ensures equal distribution of plasmids to daughter cells (see a recent review by Pogliano (2002)), a post-segregational killing system kills daughter cells devoid of plasmid after cell division (Gerdes et al., 1997), and a site-specific recombination system converts plasmid dimers arisen by homologous recombination back to monomers (Summers, 1998).
The case of plasmid R1: regulation through inhibition of translation

Plasmid R1 is a 90 kb low-copy number plasmid, belonging to incompatibility group IncFII. Other members of this group are R100 and R6-5. R1 resides in *Escherichia coli* and other closely related bacteria and carries genes that confer resistance to the antibiotics kanamycin, ampicillin, sulfonamides, streptomycin, spectinomycin, and chloramphenicol. It also carries genes that assure its stable maintenance, and genes involved in conjugal transfer.

All genes and sites required to obtain the characteristic copy number of R1 of about 4 copies per exponentially growing cell or about 1.5 copies per bacterial chromosome are located in a contiguous 2 kb segment of R1 DNA, called the basic replicon (see Figure 1) (Kollek et al., 1978).

Initiation of replication requires the binding of multiple copies of RepA to oriR1 (Masai and Arai, 1987). RepA is a 33kDa plasmid-encoded protein that acts in cis and has to be synthesized de novo for replication to occur. The frequency of replication is determined by the amount of RepA synthesized and this is regulated at two levels. The repA mRNA can be synthesized from two different promoters, the repA promoter and the constitutive copB promoter. Transcripts started at the copB promoter encode the CopB protein and the RepA protein. CopB is a transcriptional repressor that shuts off the repA promoter at normal copy number (Light and Molin, 1982a, 1982b). Translation of the repA mRNA is regulated by an antisense RNA, CopA. CopA is transcribed from a promoter on the opposite strand of repA. The
product is a 90 nt long, unstable, untranslated and highly structured RNA, which inhibits \textit{repA} mRNA translation by binding to its target CopT in the leader region of the \textit{repA} mRNA (Light and Molin, 1982a).

An open reading frame encoding a 24 amino acid long peptide called translational activator peptide (Tap) overlaps with the \textit{repA} gene (Blomberg et al., 1992). The \textit{tap} stop codon is located two nucleotides downstream of the \textit{repA} start codon. Translation of \textit{tap} is required for efficient \textit{repA} expression, i.e. translation of \textit{tap} disrupts the stable secondary structure which prevents \textit{repA} translation. As a result, the \textit{repA} ribosome binding site (RBS) becomes accessible to the ribosome (Blomberg et al., 1994). In addition, \textit{tap} translation enables translation of \textit{repA} by translational coupling between \textit{tap} and \textit{repA}. Binding of CopA to CopT inhibits \textit{tap} translation directly by interfering with ribosome binding to the \textit{tap} RBS (Malmgren et al., 1996). Thus, CopA inhibits \textit{repA} synthesis indirectly via \textit{tap}.

The binding process of CopA and CopT occurs in at least two experimentally distinguishable steps (see figure 2). The first step leads to a loop-loop contact between CopA and CopT and is called "kissing" (Persson et al., 1988;1990a; 1990b). Formation of the first binding intermediate permits the RNAs to proceed to a four-helical junction structure by formation of intermolecular helices B and B’, intramolecular helices A and A’ and an additional helix C between the 5’ tail of CopA and the complementary part of CopT (Kolb et al., 2000). This structure (shown in Figure 3), which could be called the "hugging complex" is able to inhibit \textit{repA} expression. Full duplex formation is slow and is not required for inhibition of \textit{repA} translation (Malmgren et al., 1997).
Figure 2: Binding pathway of CopA and CopT.

Figure 3: The four-way junction complex of CopA and CopT.
Other antisense RNA regulated plasmid copy number systems

Antisense RNAs are involved in the regulation of a great number of other plasmids than R1. To give an overview of the different ways they can be involved in copy number regulation, some of the best defined examples are described below.

**ColIb-P9; inhibition of pseudoknot formation**

The group I plasmids, such as the IncB and IncIα have a similar organization of their replication control regions. The IncB plasmid pMU720 (Praskier et al., 1992; Wilson et al., 1994), e.g., and the IncL/M plasmid pMU604 (Athanasopoulos et al 1995; 1999) share a similar kind of replication control mechanism as ColIb-P9 of the IncIα-group (Asano et al., 1991), which is described below. ColIb-P9 is maintained at one or two copies in enterobacteria. Its way of copy number control shows great similarities to the systems of the IncFII-plasmids R1 and R100, which are distant relatives of ColIb-P9. The replication region contains the repZ gene encoding a 39 kDa replication initiation protein. The inc gene resides 5' to repZ and encodes the antisense Inc RNA of 70 nt. The target site of the Inc RNA is folded into a stem-loop with a size of 51 nt. Translation of repZ depends on translation of repY, a short ORF encoding 29 amino acids. The main function of repY translation is to disrupt a stem-loop at the repZ ribosome loading site. This in turn allows a pseudoknot to form which facilitates the recognition of the repZ RBS by the ribosome. The function of the Inc antisense RNA is two-fold: the binding of Inc RNA to its target blocks translation of repY, and it inhibits formation of the pseudoknot, thereby inhibiting repZ translation tightly (Asano et al., 1991; Asano and Mizobushi, 1998). The complex formed between Inc RNA RepZ shows a similar four-way junction structure to the one observed for CopA-CopT complex, despite of sequence differences (Kolb et al., 2001a). Comparisons of several target and antisense RNAs encoded by other plasmids indicate that similar binding pathways are used to form inhibitory antisense-target RNA complexes (Kolb et al., 2001a).
Figure 4. Proposed regulatory mechanism of \textit{repZ} gene expression of plasmid ColIb-P9. Genes \textit{repY} and \textit{repZ} are translationally coupled. On the mRNA, the Shine-Dalgarno sequence (SD) is exposed to ribosomes. The SD and the initiation codon of \textit{repZ} are sequestered in stem-loop III. Without binding of Inc antisense RNA, \textit{repY} translation takes place and the ribosomes at the end of \textit{repY} unfold structure III, allowing the formation of a pseudoknot, facilitating binding of ribosomes to the \textit{repZ} SD, followed by translation of \textit{repZ}. Inc RNA hinders pseudoknot formation and translation of \textit{repY}, leading to inhibition of \textit{repZ} translation. The figure is based on Del Solar and Espinosa (2000).
**ColE1-relatives: inhibition of primer formation**

ColE1 is the prototype of many closely related high-copy number plasmids and was the first plasmid shown to be regulated by antisense RNA (Tomizawa et al., 1981). It does not require any plasmid-encoded proteins for regulation of its copy number, instead it relies entirely on host factors.

Replication of ColE1 requires the synthesis of a pre-primer, RNAII, which binds to the plasmid DNA within the origin region. The RNA strand of an unusually long RNA/DNA hybrid is then cleaved by RNase H, and the resulting mature primer is extended by DNA polymerase I. Upon binding, the 108 nt tightly folded antisense RNAI induces a change in the folding of the nascent primer, preventing the formation of a stable RNA/DNA structure and thereby the formation of the mature primer. For regulation to work, timing is important, since there is only a small time window of four or five seconds during which RNAI binding is inhibitory (Tomizawa, 1986). I.e. RNAI must bind while RNAII is between 100 and 360 bases long. When RNAII is shorter than 100 bases the RNA polymerase is restricting access to the primer RNA for RNAI. When RNAII is longer than 360 bases, RNAI can bind, but RNAII is already committed to fold into the active structure.

The antisense systems of plasmids ColE1 and R1 have some similarities. In both systems, antisense/target RNAs start binding through a kissing intermediate and bind rapidly with an apparent binding rate constant of around $10^6 \text{M}^{-1} \text{s}^{-1}$. Furthermore, their antisense RNAs are highly structured and present recognition sequences in their loop structures, whereby a transient kissing interaction is initiated. As seen in CopA/CopT binding, complete basepairing between RNAI and RNAII is a very slow process and seems to be irrelevant for inhibition in vitro (Tomizawa et al., 1990).

**Plasmid pIP501 and pT181: transcription attenuation**

In plasmids pIP501 and pT181, similar to the ColE1 case, binding of the antisense RNA must occur within a short time window (estimated to be 10-20 seconds) to be productive. Here, the repR mRNA encoding the replication initiator protein is the target. The repR RNA can adopt two mutually exclusive structures, one of which is induced by the antisense RNA. Antisense binding results in the formation of a rho-independent transcriptional terminator, and therefore in premature termination of repR mRNA transcription (Brantl et al., 1993). Escape
from antisense RNA binding silences the transcriptional terminator by refolding of the mRNA.

In contrast to the previously described cases, pIP501 encodes an unusually stable antisense RNA (RNAIII). This is expected to hamper proper maintenance of copy number. However, this problem is solved by a second control mechanism provided by a small transcriptional repressor CopR that binds to the repR promoter pII leading to a 10-20 fold decrease of transcription of the repR mRNA and consequently of the pIP501 copy number. A second role of CopR is to prevent convergent transcription from promoters pII and pIII, thus increasing the initiation frequency from pIII, resulting in higher levels of RNAIII (Brantl and Wagner, 1997). See Figure 5 for a schematic representation of the dual regulation of plasmid pIP501 copy number.

I.e., if copy numbers increase, RNAIII is sufficient to downregulate replication. Higher RNAIII levels lead to lower copy number. If copy numbers decrease, the intracellular CopR concentration decreases, derepressing repR transcription. At the same time, convergent transcription from pII and pIII results in a decrease of initiation at pIII and thus lower RNAIII levels. In addition, the now higher amounts of repR mRNA titrate the remaining long-lived RNAIII.

Figure 5. Dual regulation of plasmid pIP501 copy number. The initiator RepR protein acts on the origin of replication. The transcriptional repressor CopR regulates transcription from pII, which is the only promoter that directs the expression of repR. CopR is synthesized from promoter pI, which is constitutive. From promoter pIII, RNA III is synthesized. Interaction of RNA III with its target in the leader region of the repR mRNA induces generation of a mRNA secondary structure, which acts as a transcriptional attenuator, avoiding repR expression. CopR-dependent inhibition of pII increases transcription from pIII by alleviating the effect of convergent transcription. The figure is based on Del Solar and Espinosa, 2000.
Antisense RNA systems encoded by the bacterial chromosome

Independent approaches of genome wide searches in *E. coli* for small non-coding RNAs have resulted in a list of chromosomally encoded antisense RNAs (Argaman et al., 2001; Rivas et al., 2001; Wassarman et al., 2001; Chen et al., 2002). These chromosomally encoded antisense RNAs are usually *trans*-encoded and have only partial complementarity to their targets, making it possible for an antisense RNA to have multiple targets and different regulatory roles. Several *trans*-encoded antisense RNAs have been shown to need protein Hfq to exert their function (described below) (Zhang et al., 2002; Möller et al., 2002). Some of the chromosomally *trans*-encoded small regulatory RNAs are described here.

**MicF**

About 15 years ago, the first chromosomally encoded antisense RNA was isolated: MicF (Andersen et al., 1987). It is induced under stress conditions like elevated temperature, high osmolarity and redox stress. The target of MicF is the *ompF* mRNA coding for a porin of the *E. coli* outer membrane whose expression is maximal at low temperature and low osmolarity. Formation of a partial duplex between MicF and *ompF* mRNA inhibits *ompF* translation and promotes *ompF* mRNA degradation (Andersen and Delihas, 1990). So far, *ompF* mRNA is the only target of MicF known.

**OxyS**

OxyS RNA is a 109 nt untranslated RNA induced in response to oxidative stress in *E. coli*. OxyS represses translation of two of its target genes: *fhlA*, encoding a transcriptional activator and *rpoS*, encoding the alternate sigma factor sigma S. OxyS repression of *fhlA* is achieved through two basepairing interactions (Altuvia et al., 1998; Argaman and Altuvia, 2000). One site overlaps the ribosome binding site and a second site resides within the coding sequence of the *fhlA* mRNA. OxyS RNA-*fhlA* mRNA base pairing prevents ribosome binding and thereby represses translation.

The mechanism of OxyS RNA repression of *rpoS* translation is less clear. Here, a rather long A-rich single stranded region between stem loops 2 and 3 of OxyS is involved, although this part of OxyS does not show significant sequence complementarity to rpoS mRNA. The involvement of protein Hfq (described below) might be an important
factor in the OxyS-rpoS interaction. OxyS might sequester Hfq or form a translationally incompetent ternary complex with Hfq and rpoS mRNA (Zhang et al., 1998).

**DsrA**
At low temperature, activation of translation of rpoS is regulated by DsrA, another small untranslated RNA (Sledjeski et al., 1996). Besides rpoS, DsrA has another target, hns mRNA (Sledjeski and Gottesman, 1995). DsrA is predicted to form a three-stem-loop structure. The first stem-loop is involved in the regulation of translation of the rpoS mRNA. rpoS mRNA translation is activated by disrupting an inhibitory structure that normally sequesters the RBS. The second stem-loop of DsrA is necessary for the repression of hns translation by destabilization of the mRNA and blocking of the RBS (Lease and Belfort, 2000). The third stem-loop functions as a transcription terminator. So, DsrA is an example of small RNAs that can both activate and repress target genes.

**RprA**
RprA is the third known antisense RNA involved in rpoS translational control. It interacts with the same region of the rpoS leader as DsrA, and, similar to DsrA, activates translation of rpoS. RprA synthesis is regulated by the RcsC/RcsB phosphorelay system, which regulates capsule synthesis (Majdalani et al., 2002). RcsB possibly acts directly on the RprA promoter, stimulating the binding of RNA polymerase. RcsC in turn regulates RcsB, by phosphorylation. However, the exact signal for RprA regulation of rpoS remains to be discovered.

These examples of small regulatory RNAs illustrate the flexibility in ways by which they can be used to regulate gene expression: the same small RNA can bind multiple targets, and in addition, more than one small RNA can be acting on the same target.

**The role of proteins in antisense/target RNA binding**

In the plasmid copy control mechanisms described above the antisense RNA is usually the key regulator and proteins play auxiliary roles.
In the copy number regulation of plasmid R1, the CopB repressor protein represses transcription from the repA promoter, but is non-essential for the replication control mechanism (Riise et al., 1982). Also proteins that influence antisense/target RNA binding can be dispensable in regulation of copy number. The Rom protein of plasmid CoLE1 is an example of this. Rom binds specifically to the transiently formed intermediate complex of RNAI and RNAII, suppressing dissociation of the two RNAs and facilitating formation of a persistent hybrid duplex of the two RNAs (Tomizawa, 1986).

hfq
Another protein proposed to be involved in antisense RNA/target RNA binding is Hfq, although at this point such an activity has only been shown in trans-encoded antisense RNA systems. Hfq was identified as an E. coli host factor required for the replication of RNA bacteriophage Qβ (Franze de Fernandez et al., 1968). Hfq is a small, 11.2 kDa, abundant protein (60,000 molecules per cell), mainly located in the cytoplasm in association with ribosomes. Inactivation of the hfq gene in E. coli causes pleiotropic effects like decreased growth rate, sensitivity to UV-light and mutagens and, increased cell length (Tsui et al., 1994; Muffler et al., 1997). It has been shown to participate in many aspects of RNA metabolism in the cell, including mRNA stability, mRNA polyadenylation and translation. Furthermore, Hfq is involved in targeting mRNA for degradation either by increasing polyadenylation (Hajnsdorf and Regnier, 2000) or by interfering with ribosome binding (Vytvytska et al., 2000). Hfq binds single stranded RNA and has a preference for A/U rich regions. These features of Hfq show great similarities to the RNA binding proteins Sm and Sm-like (LSm) proteins, which also play wide-ranging roles in RNA metabolism. The Sm proteins contain two conserved regions: the Sm1 and Sm2 motifs. Crystal structures of Sm proteins show a N-terminal α-helix followed by a twisted five-stranded β-sheet. This Sm-fold oligomerizes and forms homo-heptameric rings in archea and hetero-heptameric rings in humans (Walke et al., 2001). In contrast, crystal structures of Staphylococcus Hfq revealed that the protein only contains one Sm1 motif and forms a homo-hexameric ring structure (Schumacher et al., 2002).

Part of the pleiotropic effects of hfq mutation can be explained by the role Hfq plays in expression of the Sigma factor encoded by rpoS. The pleiotropic phenotype of the hfq mutant resembles the phenotype
of a rpoS mutant. Muffler et al. (1996) reported that Hfq is required for efficient translation of the rpoS mRNA. Hfq recognizes and binds A-rich sequences in the vicinity of the ribosome binding site in the rpoS mRNA, and thereby, releases a secondary structure sequestering the Shine-Dalgarno sequence and allows translation (Zhang et al., 1998).

Hfq has been shown to bind to several small RNAs involved in the expression of rpoS: OxyS, RprA and DsrA (Zhang et al., 1998; Wassarman et al., 2001; Sledjeski et al., 2001). The way how Hfq is involved in the regulatory activities of these antisense RNAs is not clear. Studies by Sledjeski et. al. (2001) on DsrA suggested that Hfq acts as a protein cofactor for the regulatory activities of DsrA by either altering the structure of DsrA or forming an active DsrA-Hfq complex. A more recent report from the same group (Brescia et al., 2003) suggests that Hfq does not alter DsrA secondary structures, but might affect its tertiary conformation.

Hfq has been shown to promote contact between the OxyS and fhlA RNA (Zhang et al., 2002), and between Spot42 and galK mRNA (Møller et al., 2002) suggesting that Hfq assists in bimolecular RNA-RNA interactions, probably by unfolding structures to allow RNA-RNA interaction. Moll et al. (2003) proposed that Hfq acts as an RNA-chaperone as it was able to rescue an RNA folding trap in a splicing defective T4 bacteriophage td gene in vivo.

**Kinetics of binding of two hybridizing nucleic acids**

For efficient regulation of plasmid copy number, it is important that the antisense RNA is unstable and a fast binder. Different subjects concerning binding kinetics, such as RNA structure and methods for studying antisense-target RNA interaction are discussed below.

**Non-structured RNAs**

Complementary nucleic acids in solution hybridize when they encounter each other. Upon encounter, two complementary strands may form a precomplex, which does not involve sequence specific interactions. During the lifetime of this precomplex, short-distance dissociation followed by reassociation may occur at high frequency. This process is called microcollisions. Subsequently, two scenarios are possible: the precomplex dissociates or a first base-specific complex may be formed. The formation of the first few base pairs is called
nucleation. Pörschke and Eigen (1971) reported that the rate constant of hybridization is almost independent of the chain length of the hybridizing molecules, suggesting that the nucleation process is the rate limiting step in the hybridization reaction.

There are several reasons why the first few base pairs might form more slowly than a base pair at the end of a helix. Formation of the first pair is not stabilized by stacking free energy. A second base pair can form next to the first only if the coils are aligned in a proper orientation. With such difficulties involved in initiation of a helix, the first base pairs might often break before the helix could begin to grow, hence two or three bases are pairing and unpairing in rapid equilibrium, until a stable nucleus is formed. Once there is a nucleus of about 3 base pairs for helix growth, the helix will "zip up", and a limiting rate of base pair addition of around $10^6 \text{ s}^{-1}$ will be attained (Craig et al., 1971).

The rate of helix dissociation depends on the number of base pairs that must break to reach the pre-helix intermediate, which then dissociates very rapidly. Hence the rate and activation energy of helix dissociation are strongly dependent on the length of the helix (Craig et al., 1971).

**Folded RNAs**

The current understanding is that simple sequence complementarity-dependent hybridization cannot account for the high specificity and efficiency shown by antisense control systems. Rather, evolved three-dimensional structures of both antisense and target RNA are important determinants for high specificity and efficiency. Furthermore, recent data indicate that antisense RNAs mediate inhibition by forming complexes that involve limited numbers of base pairs with their targets (Franch et al., 1999; Kolb et al., 2000a). Formation of full duplexes is too slow to account for the observed high binding rates (Malmgren et al., 1997).

However, as described for unstructured RNAs, the initial step of the binding pathway has to be the formation of the binding nucleus. Formation of this binding nucleus is highly dependent on the structures of the RNA molecules involved. Nucleotides involved in intramolecular base-pairing do not serve as nucleation sites. Secondary and tertiary structures in antisense RNAs might display possible nucleation sites to other molecules and enable in this way nucleation sites to find each other very fast, increasing the apparent annealing
rates. This initial base-pairing interaction between loops or a loop and an unstructured RNA region is rate-limiting for the binding reaction.

The fastest antisense RNAs form stable antisense/target complexes at very similar second order association rates of $10^6$/M/s. In spite of great differences in sequences and structures of antisense and target RNAs as well as the binding pathways used, this suggests that there is a similar or identical rate limiting step in the binding process, which might be the formation of the binding nucleus.

In the case of CopA/CopT, the rate-limiting initiation complex is reversible but is followed by a rapid first order reaction to form an irreversible stable complex. The formation of this first short intermolecular helix comprises of 6-8 base pairs involving nucleotides in the loop and the upper part of the stem of each RNA. The association rate constant for the stable complex formation is nearly the same as for initiation complex formation, which suggests that binding nucleus formation is rate limiting, also for stable complex formation.

In general, the efficiency of an antisense RNA acting on its target is determined by two parameters, the binding rate constant ($k_{app}$) and its intracellular concentration. The binding kinetics resembles Briggs-Haldane kinetics, in that the formation of the primary intermediate (the kissing complex in the case of CopA/CopT) determines the overall rate of the reaction, and the probability of dissociation of the complex is small compared to that of conversion to a stable complex. This means that binding appears to be diffusion controlled; the rate of stable complex formation increases as a function of antisense RNA concentration far beyond the $K_D$ of the kissing complex.

**Important structural features for fast kinetics**

The structural features that favor the rapid formation of inhibitory complexes are not well understood and the complexity of binding pathways makes general conclusions difficult. However, several structures have been shown to be important for fast binding kinetics and effective regulation of target gene expression:

*Hairpin-loops*

The initial contact between the two RNAs always involves single stranded regions in both RNAs, such as loops or short unpaired
structures. In these regions, binding can start similar to as described for unfolded RNAs, with formation of a binding nucleus. Loop sizes for antisense/target RNA are crucial: 5-7 nt unpaired loop nucleotides result in maximal binding rates in the case of CopA/CopT (Hjalt and Wagner, 1992). Stem segments closing the loop are important for providing a scaffold for structures like the U-turn to be able to form in the loop (paper IV).

In many cases, if not all, the sequence of the loop bases is important. G/C rich loop sequences maximize the binding energy for nucleus formation. Several different motifs have been reported to form specific loop-structures. Examples of these are the tetra-loops like GNRA-loops (N=any nucleotide, R= G or A, Woese et al., 1990) and the YUNR-motif encoding the U-turn structure described below.

Furthermore, the bases closing the loop can be of importance for stability of the formed complex. In aptamers forming a loop-loop complex with TAR RNA it was shown that the GA pair is crucial for the formation of the TAR-RNA aptamer complex and that any other "pairs" lead to TAR-RNA aptamer complexes whose stability decreases in the order AG > GG > GU > AA > GC > UA >> CA, CU (Duconge et al., 2000). Similar, replacement of the wild type GC pair next to the loop of RNAI by a GA pair stabilizes the RNAI-RNAII loop-loop complex from the ColE1 plasmid replication control system, suggesting that a closing GA pair could be preferred for kissing complexes (Duconge et al., 2000).

In addition to antisense RNAs, other RNAs have been reported to interact through stem-loop structures. E.g. dimerization of HIV-1 genomic RNA is mediated by formation of a loop-loop kissing complex between both monomers (Paillart et al., 1996). Furthermore, Wagner et al. (2001) reported that localization of the bicoid mRNA to the anterior pole of the Drosophila melanogaster egg requires dimerization through loop structures in the 3’ untranslated regions of the RNAs.

U-turn

U-turns are structural elements first found in the anticodon loops of tRNAs (Quigley and Rich, 1976), where they facilitate rapid codon-anticodon interaction (Ashraf et al., 1999). Asano et al. (1998) reported on the presence of the sequence 5’-rUUGGCG-3’, encoding a putative U-turn in the loop of the target RNA involved in antisense mediated copy number regulation of plasmid ColIb-P9. Other antisense RNAs
having a similar loop-sequence were mentioned in this report. Comparative analysis of recognition loops of antisense RNAs and their targets revealed the presence of a common sequence motif (YUNR) that specifies U-turns at the same relative positions within the recognition loops (Franch et al., 1999). Subsequently, the U-turn was proposed to be important for the high overall binding rates observed in natural antisense systems and was considered as a general enhancer of binding rates in RNA-RNA interactions (Franch and Gerdes, 2000). The U-turn can either be present in the target RNA as seen in CopT, or in the antisense RNA (e.g. RNAI of ColE1). It seems not to be of importance which of the two interacting RNA molecules carries the U-turn, instead, it is of importance that one of the two RNAs has a U-turn structure (paper IV).

The YUNR-motif (Y=pyrimidine, N=any nucleotide, R=purine) folds into a structure which creates an abrupt change in the direction of the phosphodiester-backbone between the invariant uracil and the N-base stabilized by three non Watson-Crick interactions involving the invariant uracil nucleoside. The nucleobases on the 3’ side of the turning phosphate are presented in an A-helical structure creating an unpaired Watson-Crick surface for rapid interaction with complementary RNA (Franch et al., 1999). Furthermore, the U-turn architecture retracts the phosphodiester backbone within the loop, thereby decreasing the local electronegative potential surrounding the bases presented. This promotes the pairing to the complementary bases by reducing backbone repulsion in the initial recognition step. So, in other words, the U-turn structure facilitates the formation of an embryonic helix and by this speeds up the binding rate.

**Bulges**

Bulges are unpaired bases in an otherwise helical stem. Bulges create instability in a stem-loop structure, by lowering the ΔG of the stem. This facilitates the formation of other structures in complex with the interacting RNA. It has been shown that bulged residues promote the progression of a loop-loop interaction to a stable and inhibitory antisense-target RNA complex, the four-way junction (Kolb et al., 2001b). Without bulges, CopA and CopT do not progress in their binding and arrest at the kissing complex stage. Consequently, inhibition of repA translation is impeded.
**Single stranded regions**

Single stranded regions are present as 5’ tails or as unstructured middle regions. When involved in base-pairing interactions, they stabilize antisense/target complexes and allow the conversion of kissing complex to stable inhibitory complexes, like for example the four-way junction structure of CopA/CopT. Furthermore, there are antisense systems that use unstructured tails to initiate antisense/target binding. They do this via loop-linear RNA pairing. Examples of these systems are Hok/Sok of plasmid R1, RNA-IN/RNA-OUT of Tn10 (Kittle et al., 1989) and RNAII/repB of pLS1.

**Helical junctions**

Branched nucleic acid species, three-way and four-way junctions in particular, constitute important structural elements in RNA and DNA, either as part of the architecture (mainly RNA) or as potential intermediates in biological processes (mainly DNA). The most well known helical junction structure is the DNA Holliday junction, which can be observed during homologous and site-specific recombination. In the hairpin ribozyme a four-way junction determines the stability of the tertiary structure, and structural perturbation of the junction leads to impaired catalytic activity (Lilley, 1999). Also, in the binding pathway of CopA and CopT such a branched structure is the stable end product of the pathway. The formation of such a structure is important for fast kinetics, as it abolishes the necessity to form a complete duplex in order to be able to inhibit mRNA translation. In general, four-way junctions have a high degree of flexibility and can propagate along homologous stretches through dissociation and reassociation of base pairs, called branch migration.

**Methods for studying antisense-target RNA interaction in real time**

Previously, RNA-RNA interactions were studied by indirect methods e.g. band shift assays. However, nowadays real time kinetic methods are available to study interactions between RNA molecules. Real time techniques are ideal for studying antisense and target RNA interactions. The very early interactions in the binding pathways of antisense/target RNAs have been difficult to study, because of high association rates. As soon as an early complex has formed, the subsequent steps are very rapid, and therefore, early intermediates may not build up to an appreciable degree and are not detected. In real time, the pathway can
be studied from the start. Two real time techniques have been used in studies summarized in this thesis: surface plasmon resonance (SPR) and fluorescence energy resonance transfer (FRET). Advantages of SPR are that one can use a wide range of experimental conditions and that only low amounts of interactants are required for the analysis. However, this can also be a drawback, as measurements at high concentrations needed to study the early steps of antisense/target RNA pairing are not feasible.

FRET has several advantages over SPR: the reactants are free in solution, not disturbed by immobilization or mass transfer effects. Furthermore, FRET has in principal no limitation of concentration range. Third, FRET relies on chromophores and their changes in fluorescence in correlation to distance. This allows a high degree of freedom of experimental setup, chromophores can be placed at chosen positions to study putative binding sites.

**Surface Plasmon Resonance**
The BIACORE instrument makes use of the optical surface plasmon resonance (SPR) phenomenon allowing direct, real time kinetic measurements of the interaction of unlabeled biological molecules at surfaces. The BIACORE instrument is set up so that a light source and an optical detection unit are on one side of a glass surface. This surface, called the sensor chip, is coated on the bottom with a thin gold layer. As light from the light source shines on the sensor chip, it is reflected at all angles but one, at which it will be absorbed by the gold film. The identity of the absorbed angle depends on the refractive index of the material close to the nonilluminated side of the gold surface. A ligand, DNA, RNA or a protein, can be immobilized on this side of the sensor chip. An analyte of interest is analysed for its binding to the ligand by letting it pass by the immobilized ligand on the sensor chip. Binding of the analyte to the ligand causes a change in refractive index and will result in a shift in the absorbed angle, which is recorded by the detector and is displayed on a computer screen in the form of a sensogram. The association of the ligand and analyte can be followed in real time, as well as the dissociation.

The kinetic analyses obtained by SPR can provide detailed insights into the mechanisms of complex formation. Although the SPR technique has mainly been used to study protein-protein interactions, also nucleic acids interacting with proteins and nucleic acid-nucleic acid interactions can be studied by using this technique. However,
reports on studies of folded RNA-RNA interactions by SPR are rare. The first report by Duconge et al. (2000) deals with studies of the complex formation between the trans-activation-responsive (TAR) RNA and RNA aptamers. Furthermore, there is a report of Nair et al. (2000) on studies of the HIV TAR kissing hairpin complex and Paper II about SPR studies of a short version of CopA binding to CopT.

**Fluorescence Resonance Energy Transfer**

Fluorescence resonance energy transfer (FRET) is a distance dependent interaction between the electronic excited states of two dye molecules in which excitation is transferred from a donor molecule to an acceptor molecule without emission of a photon. FRET is an important technique for investigation of a variety of biological phenomena that produce changes in molecular proximity. Examples of applications are immunoassays, receptor/ligand interactions, probing interactions of single molecules, structure and conformation of nucleic acids, detection of nucleic acid hybridization, and many more.

To follow the kinetics of a conformational change or a binding reaction, the RNA is labeled with donor and acceptor fluorophores at two sites that are expected to experience a change in distance during the reaction. Using a stopped-flow mixing apparatus attached to a fluorometer the RNAs of interest are mixed rapidly and the change in donor fluorescence is monitored over time. Rate constants can be extracted by fitting the transient fluorescence with the appropriate mathematical functions.

Similar as to the situation for BIACORE, not many papers report studies of RNA-RNA interactions. Rather, RNA conformation and folding of for example ribozymes has been studied by using FRET (Klostermeier and Millar, 2001; Walter, 2001; Tuschl et al., 1994).

However, Rist and Marino (2001) studied the association of the kissing complex of RNAI and RNAII of ColE1. Using a stopped flow device they observed at least three different microenvironments during association of the complex, suggesting a kinetic intermediate in the kissing pathway. Based on this they propose a three state model for the association of the loop-loop kissing complex. The first step is collision followed by formation of a loop-loop helix between the two individual hairpins, which at a slower time scale isomerizes to the final tertiary fold.
THE PRESENT INVESTIGATION

Scope of this investigation

Previous studies suggested that the binding pathway of CopA and CopT takes place in two steps: formation of the kissing complex intermediate followed by full duplex formation. However, recently it was revealed that the initial CopA/CopT loop-loop interaction is disrupted and proceeds to the formation of a more extended kissing complex, instead of a complete duplex. The extended kissing complex is a four-way junction structure with four pair-wise stacked helices: two intramolecular helices A and A’, and two intermolecular helices B and B’. An additional helix C is formed between the 5’ tail of CopA and the complementary part of CopT. It was not clear whether this four-way junction could inhibit repA translation, and therefore, its biological significance was investigated.

Further, formation of the four-way junction structure involves extensive breakage of intramolecular base pairs and the subsequent formation of two intermolecular helices B and B’. The formation of these helices B and B’ seemed to appear in a hierarchical way. The preference of directionality of helix progression towards helix B, or helix B’, was investigated. The reason for this directionality in the 4-way junction structure formation has been elusive for some time. Based on the known tRNA anticodon loop structure and on results from the Sok/Hok antisense/target system, it had been suggested that a U-turn in the loop of CopT might determine the preference for helix B over helix B’ formation. The involvement of the putative U-turn in the loop of CopT in determination of direction of helix progression was investigated.

Furthermore, studies had previously conducted on the importance of loop size (Hjalt and Wagner, 1992) and the role of the bulges in the helical stems of CopA and CopT (Hjalt and Wagner, 1995). The upper stem bulges have a destabilizing effect and are crucial for rapid binding and inhibition. These bulged out nucleotides were assumed to be involved in destabilization of the stem-loop structures of CopA and CopT in order to facilitate inter-strand helix propagation after formation of the kissing intermediate. After elucidation of the 4-way helical junction structure it turned out that these bulged nucleotides are at the position of the junction. Removal of the bulges abolishes the
possibility of building a four-way junction structure (Kolb et al., 2001b).

We addressed the question of to what extent the stability of the upper stem segments between the bulges is important for binding and inhibition. What effect would a disruption of these stem segments have (i.e. a further destabilization)? Is extensive base pairing of the stem so important for binding that it compensates for the topological cost of forming intermolecular helices B and B’?

During the early nineties, the overall binding rate constant ($k_{\text{obs}}$) of CopA and CopT has been determined to be around $10^6 \text{ M}^{-1}\text{s}^{-1}$ (Persson et al., 1988, 1990a, 1990b). Also, the equilibrium dissociation constant ($K_D$) had been determined by gel-shift assays, which is an indirect method (Hjalt and Wagner, 1995). Association and dissociation rate constants for the formation of important intermediates were less well known. The availability of the new methods SPR and FRET tempted us to ask the question whether it was possible to use these techniques to investigate the kinetic parameters of the CopA/CopT binding pathway. Furthermore, in case these techniques proved to be feasible, we would study the most early steps of the binding pathway of CopA and CopT, in an attempt to identify the rate limiting step of the pathway.

CopA and CopT seem to be a perfect RNA couple. No need for a protein at any stage of their binding pathway has been reported; the same appears to be true for any of the other cis-encoded antisense RNAs. However, many of the newly isolated small RNAs require the Hfq protein to be able to exert their function. One feature these small RNAs have in common is that they are trans-encoded. Here I investigated whether as this is a general rule, cis-encoded antisense RNAs do not need Hfq, but trans-encoded antisense RNAs do.

**Extended complex needed for rapid pairing (paper I)**

To determine whether conversion of the initial kissing complex to a four-way junction is required for rapid pairing *in vitro*, different mutants were constructed and investigated. Base pair inversions at different positions in the stem of CopA and compensatory mutations in CopT were introduced and pairing rates on homologous and heterologous wild type and mutant CopA/CopT pairs were analyzed *in
vitro by gel-shift assays. These assays showed that the mismatches in the upper part of the stem, above the lower bulge, decreased binding rates. However, the mutation below the lower bulge showed no effect on binding rate, indicating that intermolecular base pairing between CopA and CopT extends throughout the upper part of the two stems and is needed for high binding rates in vivo.

Preference of directionality (paper I)

To test if there is a preference of directionality of helix progression for helix B or helix B’ after formation of the initial loop-loop contact, mutations were introduced in the upper stems of CopA and CopT. Results from in vitro binding assays, lead-induced hydrolysis and enzymatic structural probing suggested that helix B has to be formed first in order to permit subsequent formation of helix B’ and helix C. This was further confirmed by performing competitive inhibition assays including CopA and CopT and different RNA variants as competitive inhibitors, designed to mimic the formation of helix B and helix B’.

In order to determine the effects of the mutations in vivo, wild-type and mutant CopA/CopT control regions were introduced into plasmids carrying repA-LacZ fusions. RepA-LacZ fusion protein synthesis was measured in the presence or absence of CopA supplied in trans. The mutations located above the lower bulged nucleotide did affect the in vivo inhibition efficiency, in accordance with the conclusions from the in vitro binding experiments, and it was concluded that the formation of helix B, but not helix B’ is required for inhibition of repA translation.

The U-turn determines directionality of helix progression (paper IV)

The asymmetric strand migration of helix B requires stem disruption above the position of the lower bulge and frees the opposing side of the stem for the second intermolecular helix B’ to form. The reason for the observed directionality of the step from the first loop-loop contact to formation of the helixes is not known. It was suggested that the U-turn in the loop of CopT might determine the preference for helix B over
helix B’ formation (Kolb et al., 2000). This was tested by in vivo inhibition efficiency studies and by in vitro binding analysis of two stem-loop II mutants whose RNA structures are shown in Figure 6. A reversed loop sequence on a wild-type CopA stem, removing the U-turn completely, resulted in decreased in vivo inhibition efficiency and thirteen times lower reaction rates in vitro compared to wild type. This suggests that the U-turn plays a role in fast binding kinetics, in agreement with the suggestion that the U-turn acts as an RNA-RNA binding rate enhancer. To investigate the role of the U-turn in determination of the directionality of helix formation, a mutant (flip) with a CopT loop sequence on a CopA stem, transferring the U-turn from CopT to CopA, was investigated. In this mutant, it was shown by a competitive inhibition assay that a change of directionality of helix progression occurs upon swapping the loops of CopA and CopT, suggesting a role of the U-turn in directing the preference of helix formation. This flip mutant was neither affected in inhibition efficiency nor in binding rates, indicating that swapping the U-turn structure from CopT to CopA does not interfere with binding and formation of an inhibitory complex. From this we can conclude that it is important that one of the reactants has a U-turn structure, but that it is unimportant on which of the two is resides.

**Figure 6**: Locations of the mutational changes (highlighted in bold) in stemloop II of CopA and the resulting structures.
The upper stem segment functions as a scaffold (paper IV)

The bases in the upper parts of the helical stems of CopA and CopT are involved in the breaking of intramolecular helices and the formation of helices B and B’ in the formation process of the four-way junction. The importance of these helical stems was investigated by determining the effects of mutations on binding rates and in vivo inhibitory performance. Three mutations were designed (see Figure 6): two mutants contained large loops of either 13 nt (mut1) or 18 nt (mut3). Mut2 had a normal loop size and one large bulge in the helical stem. In vivo repA-lacZ expression assays showed that the mutants with large loop sizes were more severely impaired in inhibition efficiency compared to mut2, which had a similar efficiency as wild type. Mut3 was most severely affected in regulatory performance. Binding rate constants of these mutants measured by gel-shift assays showed a 14-fold impairment of mut1 and a 76-fold impairment of mut3, relative to the wt RNA pair. These results show that the loop size is important (as already had been noticed by Hjalt and Wagner (1992)) and that the upper helical stem in CopA and CopT is probably needed to keep the loop at the appropriate size, so that the YUNR motif can adopt its proper U-turn structure.

The role of Hfq (paper V)

The fact that Hfq is very important for the action of certain trans-encoded antisense RNAs like for example OxyS raises the question if there is any role for Hfq in the interaction of cis-encoded antisense RNAs and their targets. As a model system to study cis-encoded RNA, CopA/CopT was used. MicF/ompF was chosen to study trans-encoded antisense RNA, because an Hfq requirement for its activity had not been investigated. In in vivo ompF-lacZ expression assays with MicF supplied in trans, a 10-fold lower inhibition was observed in hfq mutant cells, compared to wild type. This indicated that Hfq is needed for the function of MicF as well. The observed impaired performance of MicF in hfq cells could be due to two different effects: either MicF is instable in hfq mutant cells, or MicF binding to its target ompF is less effective in absence of Hfq. The stability of MicF in wild type and hfq cells was investigated. The data showed that MicF half-life was decreased by 10-fold in hfq cells, and the steady state level was correspondingly
decreased compared to wild-type cells. This lower stability may account for the entire effect seen in the *ompF-lacZ* expression assays. In *repA-lacZ* expression assays with CopA supplied *in trans*, no difference between wild-type and *hfq* mutant strain was observed, indicating that CopA and CopT do not need Hfq. By southern blot, the role of Hfq in *cis*-encoded antisense regulation of plasmid copy number was investigated in two other plasmids, both of the ColE1 type: pBR322 and pUC19. No differences in copy number were observed in the case of pBR322. However, in the case of pUC19, a 4-fold lower copy number in *hfq* cells in comparison to wild-type was observed. It is not clear if this is due to an artifact or whether the single point mutation that is present in the replication primer RNAII (Lin-Chao et al., 1992) influences the role of Hfq in pUC19 replication.

**SPR studies (paper II)**

The feasibility of SPR for studying the binding of folded RNAs was investigated by using a short version of CopA (CopI), consisting only of stemloop II, and CopT. CopI can only form an extended kissing complex and was chosen to be able to study just binding in the loop-region without any further interactions disturbing the interpretation of the assay. CopI was immobilized to the sensor chip and homologous CopT was used as analyte. Apart from wild-type CopI/CopT, two mutants were studied, a loop mutant (C6) having a loop size of 8 nucleotides of which six are C-residues. The L/U mutant lacks the two bulges of the wild-type RNA due to two base changes in the helical stem. Both mutations had been shown to result in slower binding kinetics *in vitro* and a decrease in antisense RNA efficiency *in vivo* (Hjalt and Wagner, 1992; 1995). *K_D* values obtained by SPR analysis were similar to the values found earlier by gel-shift assays. However, SPR permits the determination of association and dissociation constants and gives thereby more information about the binding process. The C6 RNA pair, e.g., has a very high affinity compared to the wild type pair, indicated by a *K_D* of 6.9x10^{-10}M. Determination of *k_a* and *k_d* showed that this is due to a 20-fold decrease in *k_d* compared to wild-type. In contrast, the L/U pair has a low affinity with a *K_D* of 8.9x10^{-8}M, due to a 20-fold lower *k_a* value compared to the wild-type pair. These experiments showed the feasibility of SPR to study the binding of antisense-target RNA.
The BIACORE-apparatus enabled us to perform kinetic assays at different temperatures up to 40ºC. The association and dissociation of all three CopI/CopT pairs was investigated over a range of temperatures from 5 to 37ºC. Association rate constants were relatively independent of temperature. However, dissociation rate constants changed with temperature in all three CopI/CopT pairs. In general, $k_d$ increased with higher temperature. At low temperatures between 5 to 15ºC, $k_d$ decreased in the C6 and L/U mutations.

Temperature dependence of stable complex formation between CopA and CopT was tested as well. In all three RNA pairs, an increasing binding rate constant was observed at increasing temperatures, until a plateau was reached at or above 37ºC. From a comparison of the results obtained with the CopI/CopT pairs with results from the CopA/CopT pairs it is clear that the $K_D$ values that can be calculated from the temperature assays are not proportional to $k_{obs}$ values. At higher temperatures, dissociation rates increase in the SPR experiments. Overall complex formation rates determined in gel-shift assays increase as well, despite the increased dissociation rates. This shows that the stability of the kissing complex is not the major determinant of the overall rate of stable complex formation.

**FRET studies (paper III)**

In the SPR studies discussed above, only low concentrations of ligand and analyte could be used, due to mass-transfer problems occurring at higher concentrations. Unfortunately, to study the early steps of the CopA/CopT binding pathway, higher ligand and analyte concentrations are needed. FRET, not having the problem of mass-transfer as the reactants are free in solution, seemed more suitable for this kind of experiment. In the FRET studies, CopI/CopT pairs of wild-type and two mutants were investigated: the L/U mutant described above and mutant 104, containing one base change in the 5’-most nucleotide of the CopA loop, resulting in a U-A loop closing base pair, and a loop size of 4 nucleotides. Due to the fact that mutant 104 had not been analyzed by SPR, and that values for the L/U mutant could not be obtained in the FRET assays (because of sterical/distance effects, see paper III), only the values of the wild type CopI/CopT pair could be compared between the two different techniques. In comparison to the SPR technique, FRET gave a similar value for the dissociation rate.
constant but a four- to tenfold higher $k_a$ value. Consequently, the $K_D$ obtained with the different methods differed somewhat. To study early transition states during complex formation, chosen bases in the loops of CopI and CopT were substituted by 2-aminopurine and CopI and CopT were allowed to bind. However, due to the experimental setup it was impossible to measure the signal from the start of the binding reaction.
DISCUSSION

The binding of CopA and CopT

Based on previous data and results presented in this thesis, a tentative binding pathway of CopA and CopT can be presented. Upon encounter in the *E. coli* cell, CopA and CopT initiate binding at their recognition loops and form a kissing complex. Subsequently, directed by the orientation of the putative U-turn structure in the loop of CopT, formation of a four-way junction starts unidirectionally with the formation of helix B. At the same time, the upper stem helices of CopA and CopT are disrupted, and the now open structure allows the formation of helix B’. Helices A and A’ remain as intramolecular helices, and finally helix C is formed between the 5’ segment of CopA and the complementary part of CopT.

Results presented in this thesis indicate that it is not important which of the two interacting RNAs carries the U-turn structure. In contrast to CopA/CopT, in several other antisense RNA systems the U-turn is located in the antisense RNA (Franch et al., 1999), showing that both orientation variants are found in nature. Antisense and target RNAs of plasmids of the *IncI*α, *IncB*, *IncZ* family, as well as those related to R1, have been proposed to form structurally equivalent complexes (Kolb et al., 2001a). All target RNAs have upper stems that are predicted to be, or have been demonstrated to be, unstable due to bulges and internal loops and carry identical recognition loops including the UUGG motif. In the CopA/CopT four-helical junction structure, the crossover point is at the position of the lower bulge in CopA. In the Inc/RepZ complex, this point is located correspondingly (Kolb et al., 2001a). The presence of U-turn structures in antisense or target RNAs is so far circumstantial, since high resolution structures are not yet available. However, the almost invariant loops containing the YUNR sequence motif within a more variable upper stem region is unlikely to be coincidental. Furthermore, mutations abolishing the putative U-turn result in impaired regulation or lower binding rates in RNAI/RNAII of ColE1 (Eguchi and Tomizawa, 1991), Incl/RepZ of Collb-P9 (Asano et al., 1998), and RNA-OUT/RNA-IN of *IS*10 (Kittle et al., 1989). Franch et al. (1999) obtained experimental support for the putative U-turn of *hok* RNA by ethynitrosourea (ENU) probing. ENU is a N-nitroso alkylating reagent which has affinity for the phosphate group oxygens of nucleic acids and maps phosphates not engaged through hydrogen
bonds in tertiary interactions. Similar experiments performed with CopT resulted in an equal accessibility of phosphate group oxygens for ENU throughout the CopT sequence, showing that ENU probing in this case was not informative.

**Real time kinetics studies**

To be able to understand and identify the early steps of the interaction between CopA and CopT, SPR and FRET studies were performed. These real time measurements would permit the identification of the rate-limiting step of the CopA/CopT binding pathway. Both SPR and FRET were found to be feasible as techniques to determine kinetic parameters of this interaction. An important disadvantage of SPR is that the $k_a$ values measured are lower compared to other methods, probably because of mass transfer. Therefore one has to work at very low concentrations, giving small signals with relatively high error values. However, the values obtained with CopI/CopT were similar to the values obtained by other methods, showing that SPR is feasible to study RNA-RNA interactions.

In the studies presented here, the very first initial kissing complex (binding nucleus formation) could not be detected. Nucleus formation is a reversible process and would have a considerably faster on-rate constant. To detect this by SPR would mean that we need a much higher resolution, which is practically impossible. Furthermore, to approach the initial reversible kissing step, high concentrations of the reactants are needed due to expected low stability of the intermediate. For this reason, the SPR technique was not feasible and an alternative real-time approach based on FRET was developed. Due to the experimental set-up, no more information about the initial binding intermediates could be obtained by FRET.

In the FRET studies, two different fluorescent groups were used, one acting as FRET donor (fluorescein) and one as an acceptor (Cy3), positioned at the 5’ ends of helices A and A’. The FRET signal is detected when these fluorophores get close to each other in a four-way helical junction structure. In this way the formation of the four-way helical junction structure is measured and not the initial rate limiting step. To study early interactions of loop sequences 2-aminopurine fluorescence is the choice. Rist and Marino (2001) used 2-aminopurine fluorescence in their FRET studies of the ColE1 antisense and target RNAs. 2-AP probes act as extremely sensitive reporters for RNA-RNA association since the fluorescence of 2-AP is highly quenched when it
is stacked with other bases, but increases as much as 100-fold when fully exposed to solvent (Millar, 1996; Jean and Hall, 2001). In addition, 2-AP can form a base pair with uracil which is isosteric with a Watson-Crick AU base pair. Substitution of 2-AP at adenosine positions in the loops of antisense and/or target RNAs permits FRET measurements during the loop-loop complex binding reaction.

FRET studies on the binding of the L/U mutant CopA/CopT pair did not give FRET signals of any significance, and at the time there was no explanation. However, as shown by Kolb et al. (2000), the L/U mutant RNA pair becomes arrested at the loop-loop kissing intermediate. Due to the absence of bulges in the helical stems of CopA and CopT, the upper stems remain closed, and formation of the four-way junction structure cannot occur. This in turn prevents the side-by-side alignment of helices A and A’ which is needed to obtain a FRET signal in our experimental setup. Since FRET signals decrease with distance, a head-to-head L/U RNA complex fails to give measurable signals.

The role of Hfq in antisense regulation

The very abundant RNA binding protein Hfq has been shown to be required in several small untranslated RNA regulated systems in which the antisense RNA is trans-encoded, such as OxyS, Spot42, and DsrA. In contrast, my investigation suggests that cis-transcribed antisense RNA systems (at least in the three cases tested; paper V) appear to act independently of Hfq.

Why would Hfq play such an important role in the action of trans-encoded antisense RNAs but not of cis-encoded antisense RNAs? A possible explanation could be the fact that cis-encoded RNAs have complete complementarity, and an optimal presentation of the binding initiation site in the form of a structurally defined loop. In contrast, trans-encoded RNAs are only partly complementary to their targets and may have complex structures or single stranded regions which contain the start point for duplex formation. Therefore, it is conceivable that Hfq might aid in the opening the structures to create an accessible initiation site for duplex formation. Zhang et al. (2002) observed changes in the OxyS secondary structure due to Hfq binding.

A second possibility is an Hfq role as facilitator. It has been suggested that the primary role of Hfq is to act as a chaperone to aid pairing between trans-encoded RNAs and targets (Zhang et al., 2002).
How Hfq acts is still elusive. Hfq could facilitate RNA-RNA interactions by a number of different mechanisms. As already mentioned, Hfq binding could lead to the exposure of RNA regions that are important for pairing. Hfq could enhance RNA-RNA interactions by one Hfq hexamer binding two RNAs. In another scenario, two Hfq hexamers bind an RNA molecule each and bring them together. In in vitro studies, Hfq has been shown to form more complex compared to in vivo. It is not known if this is due to a more stable complex in vitro or an increase in association. Hfq is known to be tightly associated with ribosomes, and here Hfq could increase the interaction of the regulatory RNA with the target mRNA by bringing the regulatory RNA into the proximity of the ribosome.

Hfq binds to many different RNAs, not only trans-encoded antisense RNAs. Hfq has a preference for A/U rich sequences, which are very common in RNA molecules. Furthermore, there is a lack of sequence homology between the known Hfq substrates (Zhang et al., 2002) and Qβ, the first target of Hfq known. So, how does Hfq differentiate between which RNA molecules to bind or not to bind? Brescia et al. (2003), reporting on the binding of Hfq to DsrA, suggests that Hfq recognizes a higher-order structural element of DsrA. This might be a general rule, as all RNAs will contain elements of secondary structure, of which Hfq might bind a specific one. Further identification of the RNA structural elements to which Hfq binds will provide insight into its function.

The observation in paper V that Hfq has variable effects on trans-encoded antisense RNAs is interesting. In a comparison of wild-type and hfq mutant cells, the stability of OxyS is unaffected in hfq mutant cells, which also had been observed by others (Zhang et al., 2002). In contrast, MicF stability is strongly decreased in hfq cells, suggesting that Hfq acts as a stabilizing factor. Why does Hfq have such a different effect on trans-encoded antisense RNAs? As Hfq is a pleiotropic regulator it is not clear whether the effect on MicF is direct, or whether it is through some other molecule that stabilizes MicF. In vitro binding studies of MicF and Hfq will confirm or disprove complex formation. In addition, such experiments would also address possible effects of Hfq on the binding kinetics of MicF and ompF mRNA.

The increased interest of the scientific community in small non-coding RNAs has resulted in the isolation of many chromosomally
encoded small RNAs. The fact that non-coding RNAs have been isolated from eukaryotes, archa and bacteria indicates that small regulatory RNAs are everywhere and that they are here to stay. From the antisense RNAs isolated many years ago, encoded by the prokaryotic accessory elements, there is still much to learn, and this knowledge can be applied to the newly isolated RNAs. The fact that these known antisense RNAs make use of loop structures presenting the recognizing bases by specific structures can be used to design algorithms in silico to find possible targets for the newly isolated non-coding RNAs.
CONCLUDING REMARKS AND FUTURE PROSPECTS

CopA and CopT are the perfect RNA couple. They have evolved together as a genetic unit so that they form structures enabling fast binding kinetics and excellent regulatory performance. Binding rates reach levels up to the maximum of what is known for antisense RNAs and tRNA anticodon-anticodon interactions and are so fast that the binding kinetics hardly can be measured by the real time methods we have tested. The involvement of proteins in the binding reaction of CopA and CopT has so far not been shown and is probably not required.

These features make CopA and CopT an interesting RNA pair to study. Much can be learned from this natural system and the knowledge acquired can be applied in the design of artificial systems, like antisense RNAs as therapeutic agents.

However, many questions about CopA, CopT and the copy number regulation of plasmid R1 are left. For example, little is known about the antisense/target RNA binding pathways *in vivo*, and it is unknown whether the structures observed *in vitro* exist in “real life”. *In vivo* structural probing experiments are under way and will hopefully shed light on this question.

Furthermore, to get a better understanding of plasmid R1 as a biological system, it would be interesting to determine the intracellular concentrations of the various players involved in copy number control. A model of the R1 regulatory system in the cell has been formulated by Paulsson and Ehrenberg (2000).

Finally, determination of the crystal structure of CopT would tell us whether the putative U-turn conforms structurally to the known, tRNA U-turn, and crystals of the four-way junction complex may throw new light on the exact geometry of the complex between CopA and CopT.
ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to the entire staff at the Department of Microbiology at SLU and the Department of Cell and Molecular Biology at UU. Former and present colleagues have created a pleasant and stimulating working atmosphere, and I am grateful for all the support and encouragement you have given me throughout the years.

In particular I would like to mention:
Professor Gerhart Wagner, always an optimist both in time and in experimental outcome. Thanks for your patience all those years and for sharing your knowledge with us. It cannot be easy to have to work with knasbollar and muppar every day! Professor Leif Kirsebom, for doing a great job as the leader of the department. Professor Emeritus Kurt Nordström for still being an active member of the department and for sharing your knowledge with all of us.

The former and present members of the “GW-group”: Lotta, always having a good story ready; Hilde, for being such a nice colleague and a good friend for a long time, Good Luck with your family!; Pelle, for being such a nice guy; mushroom Petter, what was the name of this drug, buffalomicin?; Jörg, for helping out sometimes and for giving some good advices, Magnus, surviving everyone as a PhD-student, soon its your turn, and then it’s partytime! Yes! Good Luck!; Salme, for spreading some excitement over SLU by lighting some fires; Monica, what do we do without you? Thanks for being such a nice friend! I miss our great laughs in the lab; Klas, thanks acting like a dictionary the last week, for all the help with my computer problems and all our nice chats; Helene, for helping out with Gerharts computer problems; and the newcomers, Fabien, Johan and Henrik, sorry Henrik, still no Hfq, I feel bad........

Fredrik Söderbom, first teacher, then colleague PhD-student and now lunch-buddy. Anders Aspegren for sharing your experiences with job-searching. Santanu, for always being ready to help out with protocols and your knowledge, Janne, for answering so many questions about thesis stuff and for being such a nice guy. Maybe you are sometimes too kind? Dairmaid for being a great teacher. Solveig R, for great friendship, its a pity that our 3-H experiments never gave anything, but it was a nice collaboration! Lotta T., for a really good time in Madison and for being such a nice friend. Per, also for a nice time in Madison, and a good flight to Seattle. Thanks to Nina, Mirjana, Patti, Linda,
Tobias, Pernilla, Mathias, Ulrika, Solan, Anna-Chey, Ema and Fredrik for all the nice chats in the lab and lunchroom. Special thanks to Solan for dealing with the bosses when it was needed. Further, I would like to thank the administrators Linda and Christina for keeping the department running and keeping track of the money. Christer and Eva, for all the last-minute reparations and all the media.

I wish to thank my friends back home in The Netherlands, Marjolein (good luck with your dissertation!) and Walter, Marike and Rolf, Trix and Helene, Jan, Wilma and Kees, for your friendship and keeping contact. My friends in Sweden, Pieter, Anne and little Hanna; Johannes, Jeanette and Simon.

My dear Pa, Ma, Siemon, Erin and Kaitlin, for always showing your interest, for your support and your caring.

My family in law in and around Uppsala, especially Karin, thanks for all the baby-sitting and cleaning. I guess we would not have a life without you. Thanks to Lars for your hospitality in Kålbo, and for letting me cut your apple trees. It is really relaxing, this cutting and digging. The Storvreta-clan: Eva, Thomas, Elsa, Emil and Lille bror, Anita, Erik and Klara, Kenneth and Merit, Sara and Björn, Filip, and Oliver, thanks for all the nice parties and cosy Midsummer days. Olle, Emma, and Jennie, you are the most important in the world.

Thanks for your support, for showing me that there is another, more important life outside the lab and putting me with my feet on the ground when its needed. May our USA-dream come true!
REFERENCES


