The cleavage specificity of RNase III

Lothar Krinke⁺ and Daniel L. Wulff^{*}

Department of Biological Sciences, State University of New York at Albany, Albany, NY 12222, USA

Received April 23, 1990; Revised and Accepted July 23, 1990

ABSTRACT

We determined sites in λ cII mRNA that are cleaved by RNase III in the presence of λ OOP antisense RNA, using a series of OOP RNAs with different internal deletions. In OOP RNA-cII mRNA structures containing a potential region of continuous double-stranded RNA bounded by a non-complementary unpaired region, RNase III cleaved the cII mRNA at one or more preferred sites located 10 to 14 bases from the 3'-end of the region of continuous complementarity. Cleavage patterns were almost identical when the presumptive structure was the same continuously double-stranded region followed by a single-stranded bulge and a second short region of base pairing. The sequences of the new cleavage sites show generally good agreement with a consensus sequence derived from thirty-five previously determined cleavage sequences. In contrast, four 'non-sites' at which cleavage is never observed show poor agreement with this consensus sequence. We conclude that RNase III specificity is determined both by the distance from the end of continuous pairing and by nucleotide sequence features within the region of pairing.

INTRODUCTION

The *Escherichia coli* enzyme RNase III was first isolated on the basis of its specificity for double-stranded RNA in vitro (1). Closely spaced, often staggered single strand breaks at intervals of approximately 15 bases are produced by the enzyme using high-molecular-weight double-stranded RNA substrates (1–3). Since the product molecules are of nearly uniform length it is argued that the enzyme exhibits little or no sequence specificity in acting on high-molecular-weight double-stranded RNA (4), and that the major determinant of a site of cleavage may be the distance from the end of the double-stranded molecule. Length, however, would not seem to be the sole determinant since the product molecules are not all of precisely the same length (2).

A number of cellular functions have been discovered for RNase III, including processing of *E. coli* ribosomal RNA precursor molecules (5–8), and processing the polycistronic transcript of the bacteriophage T7 early region (5,9,10). RNase III processing sites flank the bacteriophage T3 S-adenosylmethionine hydrolase gene (11), and are found in the *E. coli* *rplJL-rpoBC* operon (12), and the mRNA transcripts for the *E. coli* *metf2* (13), and *pnp*

and *rpsO* genes (14,15). The transcript coding for the *E. coli* RNase III gene also has a cleavage site for RNase III (16). The bacteriophage λ *int* gene is regulated by RNase III (17–21), and RNase III processing sites occur elsewhere on λ transcripts (22–24). The complex of *sar* antisense RNA with *ant* mRNA in bacteriophage P22 is cleaved by RNase III (25). The RNA molecules cleaved by RNase III all are capable of forming double-stranded regions in the range of 20 to 40 bp, often with a short internal unpaired region in the vicinity of the RNase III cleavage sites. In these cases a limited homology among RNase III-sensitive sites has been detected (7).

RNase III is in some cases responsible for gene regulation by antisense RNA, such as the regulation of λ cII gene expression by λ OOP antisense RNA (26,27), and the regulation of the *E. coli* *copT* gene by the *copA* antisense RNA (28). OOP RNA is a 77 base transcript that can form a perfect double helix with its target mRNA molecule, a major rightward transcript of λ that extends hundreds of nucleotides to either side of the complementary region. OOP RNA results in RNase III cleavage in the target cII-O mRNA at 13 nucleotides from the 3'-end of the region of complementarity with OOP RNA. When an OOP RNA with an internal deletion opposite this cleavage site is paired with wild-type cII-O mRNA, cleavage no longer occurs at the normal site, and cleavage is detected instead at a new site 22 bases to the 5'-side of the initial cleavage site (27). This new cleavage site is located 13 bases from the end of a region of continuous complementarity between the deleted OOP RNA and cII-O mRNA. This raises the question of whether, in attacking a continuously double-stranded region of considerable length, RNase III always cleaves at precisely 13 nucleotides from the 3'-end. In this paper we determine the cleavage sites for antisense RNA molecules with internal deletions that extend progressively further into the molecule. We show that cleavage does not always occur at precisely 13 nucleotides from the 3'-end of the complementary region, but rather occurs at one or more of several preferred sites located 10 to 14 nucleotides from the 3'-end of the complementary region.

MATERIALS AND METHODS

Bacteria and plasmids

E. coli bacterial strains producing cII-O mRNA are UC6171 (F⁻ *galK*⁻ *hfl*-1 *str*⁻ *sup*^o) (29) and its isogenic RNase III⁻ derivative UC5993 (F⁻ *galK*⁻ *hfl*-1 *rnc105* *glyA::tn5* *str*⁻ *sup*^o),

* To whom correspondence should be addressed

⁺ Present address: Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA

lysogenic for λ int-6 cIts857 cro27 Pam3 (26). *E. coli* bacterial strains producing cII-cat mRNA are UC6650 and its isogenic RNase III⁻ derivative UC6659. The strain UC6650 is a derivative of UC6171 that has inserted into its *galK* gene a λ DNA fragment containing the λ temperature sensitive cIts857 repressor gene and the λ N gene under control of the strong repressible λ p_L promoter followed by a *kanR* gene controlled by its own promoter, and inserted into its *lacZ* gene a fragment of DNA in which the λ p_L promoter governs expression of the λ cII gene and a chloramphenicol transacetylase (*cat*) gene (Fig. 1).

To insert the λ cIts857, λ N, and *kanR* sequences into the *E. coli* *galK* gene, the desired DNA fragments were first assembled into a plasmid between two half *galK* gene DNA fragments. This plasmid (pMEM11) was introduced into a *gal*⁺ host and a *Gal*⁻ derivative was isolated that had incorporated plasmid DNA into the host chromosome by homologous recombination. The inserted DNA was then transferred to strain UC6171 by P1 transduction, selecting for kanamycin resistance. The resulting strain, UC6444, was Kan^r, Amp^s, immune to λ at 32° and sensitive to λ at 42°. The strain UC6444 grows normally at 42°.

To insert the λ p_L, λ cII and *cat* sequences into the *lacZ* of UC6444, the desired DNA fragments were first assembled into a plasmid between two half *lacZ* gene sequences that are missing small portions of the extreme 5'-end, 3'-end and middle regions of the gene. In this plasmid (pMEM49) the λ p_L and λ cII gene fragments are bounded by unique *Hind*III and *Bam*HI sites, allowing any *Hind*III-*Bam*HI DNA fragment to be inserted in their place. A *Sal*I site is located between the *cat* and *lacZ* gene sequences, allowing the entire insert to be replaced by any *Hind*III-*Sal*I DNA fragment. The DNA inserted into the *lacZ* sequences in the plasmid was introduced into the chromosome by homologous recombination, followed by elimination of the plasmid through growth at 42° and P1-mediated transduction to UC6444, selecting for Cm^r resistance at 42°. The resulting strain is Amp^s and Lac⁻, and Cm^r at 42° and Cm^s at 32°.

Figure 1. Structure of the *E. coli* strain UC6650. A 3,601 bp λ DNA fragment (base pairs 38104–34503, numbered as in ref. 30) that contains a λ temperature sensitive cIts857 repressor gene and the λ N gene is inserted into the *E. coli* *galK* gene of strain UC6171 (29). This is followed by a *kan* (kanamycin resistance) cassette of approximately 1,500 bp that includes the *kan* promoter (Pharmacia, Inc.). A p_L-cII-cat fusion is inserted into the *lacZ* gene. A 248 bp λ p_L fragment (bp 25715–35468, numbered according to ref. 30) with short oligonucleotide fragments on each side is joined a λ cII gene fragment (bp 38351–38668), followed by a *cat* gene fragment containing the *cat* ribosome recognition region but no independent promoter. The 5'-side of the *cat* gene fragment is derived from a CAT *Bam*HI cartridge from the plasmid pCM4 (Pharmacia, Inc.), and the 3'-side from a CAT *Sal*I cartridge from the plasmid pCM1 (Pharmacia, Inc.). Nucleotide sequences at the junctions are as follows (/ = junction between two fragments): *lacZ*/oligonucleotide/p_L, 5'-ACTCGGG/AAGCTTGGCTGCAGGTGCA-CG/GATCTCT-3'; p_L/oligonucleotide/cII, 5'-GCATTGG/TTAAAAATTAA-GGAGGAAT/TACTTACA-3'; cII/*cat*, 5'-CTGGATC/CGAGATT-3'; *cat*/*lacZ*, 5'-TCGTCGA/CGCGTTC-3'. Small segments of the *galK* and *lacZ* genes are deleted (bp 763–789 of the *galK* and bp 1322–1352 of the *lacZ* coding sequences, numbered as in refs. 31 and 32). Arrows indicate the orientation of genes and promoters.

To make the RNase III⁻ strain UC6659 from the RNase III⁺ strain UC6650 that produces cII-cat mRNA, a kanamycin sensitive mutant of UC6650 was isolated after mutagenesis with ultraviolet light. This strain was then used as a recipient for P1 transduction from a *mc*-105 *glyA::tn5* strain, selecting for Kan^r. A RNase III⁻ transductant was isolated on the basis of its characteristic small colony size. It was shown to retain the other essential phenotypic features of the parent strain, and its genotype was subsequently confirmed by the failure to find RNase III-dependent bands in primer extension experiments.

The OOP DNA plasmids pLK38, pLK39, pLK40, pLK50 and pLK51 are described in ref. 26, in which the last three plasmids appear as 'Δ11', 'Δ22', and 'Δ20', respectively. The remaining OOP DNA plasmids were made using recombinant DNA techniques. The deletion junctions of all plasmids with internal deletions in the OOP region were verified by DNA sequence determination.

Isolation of cellular RNA and RNase III enzyme reactions

Cellular RNA was isolated from isogenic RNase III⁺ and RNase III⁻ strains as described in ref. 27. Purified cellular RNA (30 μg) from RNase III⁻ strains was incubated with 3 μl purified RNase III (a gift of Prof. H. Robertson) in 100 μl RNase III buffer (100 mM NH₄Cl, 10 mM magnesium acetate, 20 mM TrisCl (pH 7.5)) for 1 hour at 37° (1). The RNA was phenol extracted, ethanol precipitated, washed with 70% ethanol, air-dried, and subjected to primer extension analysis.

Figure 2. The locations of deletions and RNase III cleavage sites. The figure also indicates at which sites a 5'-end was observed with each OOP DNA plasmid, using either cII-O mRNA or cII-cat mRNA. Because of direct repeats at end-points of deletions, the deletion boundaries of pLK40 may be shifted one or two bases to the left; the deletion boundaries of pDW204 one, two or three bases to the left; the deletion boundaries of pDW207 one base to the left; and the deletion boundaries of pLK50 one to six bases to the right. This may affect pairing with cII-O mRNA, which is capable of forming double-stranded complexes with OOP RNA on either side of the deletion, but not with cII-cat mRNA, which can form a double-stranded complex with OOP RNA only to the left of the deletion. (Exceptions are pLK50 and pDW204, which may form alternate complexes with cII-cat mRNA as well).

Primer extension analysis

Primer extension analysis was performed as described in ref. 27. For analysis of *cII-cat* mRNA, we used a DNA oligonucleotide

Figure 3. Potential stem structures between *cII-O* mRNA and an OOP RNA with an internal deletion, and between *cII-cat* mRNA and the same OOP RNA.

that was complementary to a 17 base segment of *cat* mRNA that was the same distance from *cII* gene sequences in *cII-cat* mRNA as the *O* primer used in ref. 27 for primer extension analysis of *cII-O* mRNA.

RESULTS

Hydrolysis of *cII-O* mRNA in the presence of various OOP RNAs

We isolated cellular RNA after λ prophage induction of ten pairs of isogenic RNase III⁺ and RNase III⁻ strains, each of which contained a heat-inducible λ prophage and a multicopy OOP DNA plasmid with a different internal deletion in the OOP region. These plasmids overproduce OOP RNA (26,27). Two additional pairs of isogenic RNase III⁺ and RNase III⁻ strains were used as controls, one of which contained a multicopy wild-type OOP DNA plasmid, and the second of which contained a multicopy OOP DNA plasmid that lacked the *p_O* OOP promoter. We then used primer extension analysis of these various strains to determine the 5'-ends in *cII-O* mRNA in the region of

Figure 4. Primer extension analysis to determine RNase III cleavage sites on *cII-cat* mRNA. Total RNA from various heat-induced strains was hybridized to a DNA primer complementary to *cat* mRNA in a region 58–74 bases from the start site of OOP RNA, extended with reverse transcriptase and run on a sequencing gel. The same gel contained DNA sequencing reactions (labeled G, A, T, and C) using the same primer with single-stranded *cII-cat* DNA cloned into M13. The generated DNA sequence is the complement of the *cII-cat* RNA cleaved by RNase III. The RNase III⁺ strain UC6650 (odd numbered lanes) and the isogenic RNase III⁻ strain UC6659 (even numbered lanes) carry a λ *clts857 N* insert in the *E. coli galK* gene, a *p_L* *cII-cat* insert in the *E. coli lacZ* gene (Fig. 1), and a multicopy OOP DNA plasmid. The plasmids are pLK38 (deleted *p_O* promoter, lanes 1–2); pLK39 (wild-type OOP DNA, lanes 3–4); and the deletion plasmids pLK40 (lanes 5–6), pLK50 (lanes 7–8), pLK51 (lanes 9–10), pDW199 (lanes 11–12), pDW204 (lanes 13–14), pDW205 (lanes 15–16), pDW206 (lanes 17–18), pDW200 (lanes 19–20), pDW207 (lanes 21–22), and pDW208 (lanes 23–24). The extents of the OOP DNA deletions in these plasmids are indicated in Fig. 2.

Table 1: Compilation of RNase III processing sites.

Substrate RNA	Sequence*	Matches**	Ref.
λ late RNA top	cC.AGUGCU.CU.Uu.CC/G.UUGUGCuGAau	13	23
λ late RNA bottom	ua.UCuGCA.CA.ACaGG/u.AAGAGCAUUGa	12	23
λ <u>N</u> gene top	cU.GcCACA.CA.CCaCC/A.AAGCUaaCUGa	9	24
λ <u>N</u> gene bottom	gC.AGaaGG.CU.UU.GG/G.GUGUGUGaUAc	9	24
λ <u>sib</u> top	uu.ugucuu.CU.UG.CG/C.UAAUuuUUUGU	11	21
λ <u>sib</u> bottom	AC.AGAGcA.UU.AG.CG/C.AAGgugauuuu	12	21
λ <u>cIII</u> top	uc.GAGGCG.UU.uU.Uc/guUAUGUaUAAaU	7	22
λ <u>cIII</u> bottom	gc.gguUUU.AccGC.AU/A.ccAAuAACGCU	5	22
λ <u>cII</u> gene	UC.UGAGGU.CA.UU.AC/U.GGAUCUAUCAA	9	26
λ <u>cII</u> gene upstream	CC.CGGCGG.CA.AC.CG/A.GCGUUCUGAAC	6	33
<u>copA</u>	GU.CGAAUU.GU.UG.GC/U.AUACGGUUUAA	9	28
<u>copT</u>	UU.AAACCG.UA.UA.GC/C.AACAAUUCGAC	9	28
<u>E. coli</u> 23s RNA top	UA.ACCUUA.CA.ACgCC/G.AAGGUGUUUUG	9	8
<u>E. coli</u> 23s RNA bottom	UG.AAACAU.CU.UC.GG/G.UUGUGAGGUUA	10	8
<u>E. coli</u> 16s RNA top	UGcAGUGCU.CA.CA.CA/G.AUUGUCUGAUG	11	8
<u>E. coli</u> 16s RNA bottom	UA.UCAGAC.AA.UC.UG/U.GUGGGCACUCG	11	8
<u>E. coli</u> <u>metf2</u> top	UC.GUAGUGgga.UU.UG/A.AAAAauccuuc	9	13
<u>E. coli</u> <u>metf2</u> bottom	ag.aguccU.UU.UU.CA/A.AUaUACUGUGA	12	13
<u>E. coli</u> <u>rnc</u> top	UC.AGguCU.GU.UU.CG/U.GUGCUGaattg	10	16
<u>E. coli</u> <u>rnc</u> bottom	tg.ccaCGG.CA.CA.CG/A.AACAGcgttgg	8	16
<u>E. coli</u> <u>pnp</u> top	GG.UuAGGg.uu.gU.CA/U.UAGUCGCGagg	5	14
<u>E. coli</u> <u>pnp</u> bottom	uc.gCGCGG.CU.AA.UG/A.gAGgCUUuACC	6	14
<u>E. coli</u> <u>dicB</u> R1	tc.agaCCA.UA.ACgCG/a.AAGCGCAUGGC	10	34
<u>E. coli</u> <u>dicB</u> R2	aC.CGUGCG.CU.Uc.CG/G.UUGUGGcaacc	9	34
<u>E. coli</u> <u>dicB</u> R3	aa.AGUGAA.UU.UU.GU/G.AUGUGGUGaau	11	34
<u>E. coli</u> <u>dicB</u> R4	ag.gCACUG.CA.UC.AC/A.AAAUUCAUUGu	11	34
T7 R1	aa.GGUGCC.CU.Uu.au/g.aUAUUCaCUAa	14	9
T7 R2	cG.AGUGGC.CU.Uu.au/g.aUUAUCACUUu	14	9
T7 R3	cg.GGUGGC.CU.Ga.au/A.GGUACGAUUua	8	9
T7 R4	aG.AGUGGC.CU.Uu.au/g.aUUGaCCUUCU	12	9
T7 R5a	AU.aGGgac.AA.UG.cU/U.AAGGUCGCUCu	13	9,10
T7 R5b	gG.AGUGGC.CU.UA.Gu/C.AUUuaaCCaAU	11	9,10
T3 R1 (pred.)***	AG.cGUGGC.CU.Uu.au/g.aUAUUCaCUUA	14	11
T3 R2 (pred.)***	UG.CGUGGC.CU.Uu.au/g.aUUAACAUUCu	13	11
T4 10*5s	aG.AgAGcc.uu.UG.UU/A.AUaaucccaaa	12	35

complementarity to OOP RNA. Our confidence in this method is based on previous experiments using RNA probes, which confirm the results of the primer extension experiments with both OOP RNA and the OOP deletion plasmid pLK40 (27). In each RNase III⁺ strain that produced OOP RNA or a variant of OOP RNA we found one or more 5'-ends that were not present in the isogenic RNase III⁻ strain (Fig. 2; original data is similar in form to Fig. 4, below). With each strain the 5'-end(s) were present in a region that could pair with the OOP RNA produced by the resident multicopy OOP DNA plasmid. Wild-type OOP RNA produced an RNase III-dependent 5'-end in *cII-O* mRNA at 13 bases from the 3'-end of the region in which *cII-O* mRNA is complementary to OOP RNA as found previously (27).

The plasmids pLK50 and pLK51 produce OOP RNAs with internal deletions of 22 and 20 bases, respectively. At least the first 23 bases of these shortened OOP RNAs are available for uninterrupted base pairing with *cII-O* mRNA. With both plasmids a single 5'-end is observed at the precise location as with wild-type OOP RNA, 13 bases from the end of complementarity with the 5'-end of OOP RNA (Fig. 2).

In the remaining eight OOP DNA plasmids with internal deletions, only 9 to 11 bases at the 5'-ends of the mutant OOP RNAs are available for uninterrupted base pairing with *cII-O* mRNA. RNase III-dependent cleavage of *cII-O* mRNA is never observed in the short region complementary to the 5'-ends of these mutant OOP RNAs, but new apparent cleavage sites are observed within the much longer region of *cII-O* mRNA that is complementary to the 3'-ends of the mutant OOP RNAs, at 10–13 bases from the deletion end points (Fig. 2). (An ambiguity exists in assigning the end points of three deletions because of direct repeats at the deletion end points. The 10–13 base range is obtained by assigning deletion end points so as to minimize the variability.) Three new apparent cleavage sites in *cII-O* mRNA were identified using these eight OOP DNA plasmids, although only one site was found as a major band with any one plasmid (Fig. 2). The 3'-ends of the deletions in these eight OOP RNAs occur at seven different sites over a span of ten bases. If the distance from the end of continuous complementarity to the cleavage site were always the same, then not three but seven new sites should have been found.

Hydrolysis of *cII-cat* mRNA in the presence of various OOP RNAs

All of the OOP RNAs with internal deletions can potentially form double-stranded RNA with *cII-O* mRNA on either side of the deletion, with the unpaired *cII-O* mRNA opposite the deletion extending outwards as a bulge (Fig. 3). We wished to determine if the pattern of RNase III cleavage would be the same when only a single double-stranded region could form, with unpaired

OOP RNA and mRNA extending outwards from this single paired region. We constructed a pair of isogenic RNase III⁺ and RNase III⁻ strains which contained, inserted into the *E. coli lacZ* gene, a 250 bp λ *p_L* promoter fragment fused to the λ *cII* gene, followed by a chloramphenicol resistance (*cat*) gene (Fig. 1). The break-point between λ *cII-O* sequence and *cat* sequence lies in the intercistronic region between the λ *cII* and *O* genes. This fusion strain lacks the OOP promoter and first 7 bases of the OOP sequence. The strain also carries, inserted into its *galK* gene, a λ fragment containing the λ *N* gene and *cIts857* repressor gene, as well as a kanamycin resistance gene controlled by its own promoter. The strain is derepressed for *cII* gene and *cat* gene expression at 42°, but is repressed at 32°. Primer extension experiments, using RNA isolated after temperature induction of these strains, revealed a series of 5'-ends that were found only in RNase III⁺ strains and only in the presence of a multicopy OOP DNA plasmid.

The *cII-cat* mRNA produced by these strains can pair with all but the first seven bases of wild-type OOP RNA, and *cII-cat* mRNA should form a single paired region with wild-type OOP RNA, with the first seven OOP RNA bases extending outwards from the paired region (Fig. 3). In primer extension experiments with wild-type OOP RNA, a single 5'-end is observed at a site not seen with *cII-O* mRNA, 13 bases from the 3'-end of the region in which *cII-cat* mRNA is complementary to OOP RNA (Figs. 2 and 4).

The OOP deletion plasmid pLK50, which has 16–22 bases at its 5'-end available for uninterrupted pairing, results in poor cleavage by RNase III, with minor bands appearing at the sites at which *cII-O* mRNA and *cII-cat* mRNA are cleaved in the presence of wild-type OOP RNA. (The 16–22 range reflects the uncertainty in the position of the bulge in the double-stranded complex because of a seven base direct repeat sequence at the boundaries of the deletion in pLK50. See ref. 26.) With plasmid pLK51, which has 18 bases at its 5'-end available for uninterrupted pairing, a major band is observed at the same site as with this plasmid and *cII-O* mRNA. In this case the cleavage site is only six bases from the end of the region of complementarity.

All the other OOP deletion plasmids (with the possible exception of pDW204) can form only one region of pairing with *cII-cat* mRNA (Fig. 2). Primer extension experiments with these plasmids give almost identical results with *cII-cat* mRNA as had been previously observed with *cII-O* mRNA (Figs. 2 and 4). No 5'-ends are observed with these plasmids in the *cII-cat* mRNA strains that were not previously observed (at least as a minor band) in the *cII-O* mRNA strains. The distance between a deletion end-point and the location of an apparent RNase III-dependent cleavage site in *cII-cat* mRNA varies between 10 and 14 bases (Fig. 2).

*Capital letters indicate paired bases; small letters unpaired bases.

Dots indicate positions at which "bulge" nucleotides are present in some sequences. The "bulge" nucleotides are indicated in place of the dots where they are present.

**Number of matches to consensus sequence (Tables 2 and 3).

***Predicted sequence.

RNase III cleavage sites generated by purified RNase III in vitro

To show that the 5'-ends observed in the above experiments indeed represent sites at which RNase III cleaves, and are not created by other nucleases that trim the RNA molecules after initial RNase III cleavage events elsewhere, we treated uncut cellular RNA, isolated from the RNase III⁻ strains that overproduce the various OOP RNAs, with purified RNase III in vitro. Primer extension experiments, using *cII-cat* mRNA treated with RNase III in vitro, revealed cleavage at all of the sites that had been observed in vivo, but with some variations in the relative amounts of the different fragments (Fig. 2). Data for *cII-O* mRNA was similar and is not shown. These experiments indicate that the 5'-ends that were detected in the the RNase III⁺ strains in vivo are the direct result of RNase III cleavage.

DISCUSSION

We have examined how various internal deletions introduced into a natural antisense RNA molecule alter RNase III-dependent cleavage of a target mRNA molecule. When a region of antisense RNA opposite a cleavage site is deleted, cleavage no longer occurs at the natural site, and new cleavage sites in the mRNA were identified at 10 to 14 nucleotides from the 3'-ends of regions of continuous complementarity with the mutant antisense RNAs. While distance from the edge of a continuously paired region

is an important determinant of cleavage, it is not *all*-important, and other factors play a role in determining cleavage sites. These other factors could include nucleotide sequence in the double-stranded region, or possibly features of tertiary structure due to the unpaired region beyond the region of continuous complementarity (Fig. 3). However structures that are identical in their double stranded regions but very different in their adjacent unpaired regions, in that one has two free ends and the other has a bulge followed by a second short region of potential pairing (Fig. 3), are cleaved in almost an identical manner by RNase III. Under these conditions minor differences are observed in the relative amounts of cleavage at different sites, but no sites are detected with one structure that are not also seen with the other. We therefore suspect that sequence differences within the double-stranded region, rather than the configuration of the adjacent single-stranded regions, are the crucial determinants of cleavage specificity. (However one could object that these structures have not actually been shown to exist.) In the eight base region between cleavage sites #1 and #4 (Fig. 2), we have determined the location of four cleavage sites. The failure to detect ever cleavage at the other four nucleotide junctions within this region leads us to conclude that they are non-cleavage sites. If sequence determines the six sites at which cleavage may occur (#1 through #6, Fig. 2) and the four sites at which cleavage never occurs (the four phosphodiester bonds between sites #1 and #2, Fig. 2), then the sequence rules for determining specificity would appear not to be simple.

Table 2: Compilation of RNase III consensus sequence.

Base	Number of occurrences (from Table 1) at position number*:												
	1	2	3	4	5	6	7	8	9	10	11	12	13
G	12	20	9	34	11	36	16	14	5	2	4	9	15
C	11	20	9	14	6	17	23	18	37	3	10	23	27
A	22	14	37	16	24	9	19	20	10	30	12	13	20
U	25	16	15	6	29	8	12	18	18	35	44	25	8
Consensus:	$\hat{A}$ $\hat{U}$	N	A	G	$\hat{A}$ $\hat{U}$	G	N	N	C	$\hat{A}$ $\hat{U}$	U	N	N

*An inverse symmetry is assumed. At '1' is tabulated the sum of occurrences at positions #1 and the complement of #26; at '2' is tabulated the sum of occurrences at positions #2 and the complement of #25; etc. Cleavage occurs between positions #14 and #15.

Table 3: Comparison of *cII* secondary sites and non-sites with the consensus sequence.

Name*	Sequence	Matches to consensus
Consensus	$\hat{U}$ NAG $\hat{A}$ U $\hat{G}$ GNNC $\hat{A}$ U $\hat{U}$ NNN/NA $\hat{A}$ U $\hat{G}$ GNNC $\hat{A}$ U $\hat{U}$ CUN $\hat{A}$	
Site 1	CGAGCGUUCUGAAC/AAAUCCAGAUGG	8
Site 2	GUUCUGAACAAAUC/CAGAUUGGAGUUC	7
Site 3	UUCUGAACAAAUCC/AGAUGGAGUUCU	5
Site 4	UCUGAACAAAUCCA/GAUGGAGUUCUG	9
Site 5	AUGGAGUUCUGAGG/UCAUUACUGGAU	10
Site 6 (wild-type)	UCUGAGGUCAUUAC/UGGAUCUAUCAA	9
Non-site 1	GAGCGUUCUGAACAA/AAUCCAGAUGGA	4
Non-site 2	AGCGUUCUGAACAA/AUCCAGAUGGAG	4
Non-site 3	GCGUUCUGAACAAA/UCCAGAUGGAGU	3
Non-site 4	CGUUCUGAACAAA/CCAGAUGGAGUU	3

*The consensus sequence is derived from sequences in Table 1, as tabulated in Table 2. Cleavage sites are those shown in Fig. 2. Non-sites are the four potential sites between Sites 1 and 2 at which cleavage is never observed, numbered from right to left in the sequence shown in Fig. 2.

To investigate the sequence specificity of RNase III, we first compiled sequences of known natural RNase III cleavage sites (Table 1). Generally, with the exceptions of some T3, T4 and T7 sites, RNase III cleaves both strands of a double-stranded target and creates a two-base 3'-overhang. (In our own studies cleavage was examined only in the cII-O strand; RNase III-dependent cleavage of OOP RNA has not been investigated.) In aligning these sequences, we considered both strands separately because they usually have different sequences due to mismatched bases and single-stranded regions. We considered bases in double-stranded conformations, and also mismatched bases in otherwise double-stranded regions and bases in completely single-stranded regions because these bases can be in helical or pseudohelical conformations. Bulges of one base were not considered in the alignment. Because both strands of most sequences are cleaved, any consensus sequence should have an inverse dyad symmetry centered between the cleavage sites on the two strands of the target RNA. (It does not necessarily follow that the enzyme recognizes an inversely symmetric sequence. However, this is implied by the postulate that the two cleavage events are each catalyzed by one subunit of the dimeric enzyme (3).) In Table 2 we tabulate the number of occurrences of each base at each position, pooling data from the two sides of the postulated center of inverse symmetry, and derive a weak consensus sequence. This sequence is an extension and slight modification of previously noted homologies (7,23). (We accepted the designation A,U at some positions not only because of the tabulated frequencies, but also because RNA binding proteins that bind to double-stranded RNA may make specific contacts only in the minor groove (36), where a binding protein may distinguish A-U and U-A base pairs from G-C and C-G base pairs, but probably cannot distinguish A-U from U-A base pairs (37).) The thirty-five sequences tabulated in Table 1 each show eight to fourteen matches to the consensus sequence (a perfect score is sixteen) with the exception of the λ cIII and *E. coli* *pnp* sites, which show only seven and five (cIII), and six and five (*pnp*) matches. However λ cIII mRNA is cleaved poorly *in vivo* (22).

We next tabulated the matches to the consensus sequence of the five new cleavage sites found in these studies, all of which are presumably not naturally occurring sites. Four of the five sites, Sites 1, 2, 4 and 5, showed between seven and ten matches to the consensus sequence (Table 3), generally within the range of matches in naturally occurring sites (Table 1). Site 3, which showed only five matches to the consensus sequence, appears from our experimental data to be a weaker site than the others (Fig. 2). However, the four 'non-sites' at which cleavage was never observed show only three to four matches to the consensus sequence. Our consensus sequence is therefore of strong predictive value, and can guide future studies on RNase III specificity using point mutations.

Our results show that the specificity of RNase III depends on both distance from the end of a double-stranded region, which determines the region in which the enzyme will cleave, and the detailed sequence in the region of cleavage, which determines the precise sites at which cleavage may occur.

ACKNOWLEDGMENTS

We thank Professor Hugh Robertson of the Rockefeller University for a generous gift of purified RNase III. The *E. coli* strains UC6650 and UC6659 and the plasmids pMEM11 and

pMEM49 were constructed by M.E. Mahoney. This research was supported by grant GM28370 from the National Institutes of Health.

REFERENCES

- Robertson, H.D., Webster, R.E., and Zinder, N.D. (1968). *J. Biol. Chem.* 243, 82-91.
- Robertson, H.D., and Dunn, J.J. (1975). *J. Biol. Chem.* 250, 3050-3056.
- Robertson, H.D., and Barnay, F. (1979). In *Proceedings of the 12th FEBS Congress*. Pergamon Press, Oxford, pp. 285-295.
- Robertson, H.D. (1977). In Vogel, H.J. (ed.), *Nucleic acid-protein recognition*, Academic Press, Inc., New York, pp. 549-568.
- Dunn, J.J., and Studier, F.W. (1973). *Proc. Natl. Acad. Sci. USA* 70, 3296-3300.
- Robertson, H.D., Pelle, E.G., and McClain, W.H. (1980). In Söll, D., Abelson, J., and Schimmel, P. (ed.), *Transfer RNA: biological aspects*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp. 107-122.
- Gegenheimer, P., and Apirion, D. (1981). *Microbiol. Rev.* 45, 502-541.
- Bram, R.J., Young, R.A., and Steitz, J.A. (1980). *Cell* 19, 393-401.
- Dunn, J.J., and Studier, F.W. (1983). *J. Mol. Biol.* 166, 477-535.
- Saito, H., and Richardson, C.C. (1981). *Cell* 27, 533-542.
- Hughes, J.A., Brown, L.R., and Ferro, A.J. (1987). *Nucl. Acids Res.* 15, 717-729.
- Barry, G., Squires, C., and Squires, C.L. (1980). *Proc. Natl. Acad. Sci. USA* 77, 3331-3335.
- Régnier, P., and Grunberg-Manago, M. (1989). *J. Mol. Biol.* 210, 293-302.
- Régnier, P., and Portier, C. (1986). *J. Mol. Biol.* 187, 23-32.
- Portier, C., Dondon, L., Grunberg-Manago, M., and Régnier, P. (1987). *EMBO J.* 6, 2165-2170.
- Bardwell, J.C.A., Régnier, P., Chen, S.-M., Nakamura, M., Grunberg-Manago, M., and Court, D.L. (1989). *EMBO J.* 8, 3401-3407.
- Belfort, M. (1980). *Gene* 11, 149-155.
- Schindler, D., and Echols, H. (1981). *Proc. Natl. Acad. Sci. USA* 78, 4475-4479.
- Guarneros, G., Montanez, C., Hernandez, T., and Court, D. (1982). *Proc. Natl. Acad. Sci. USA* 79, 238-242.
- Gottesman, M., Oppenheim, A., and Court, D. (1982). *Cell* 29, 727-728.
- Schmeissner, U., McKenney, K., Rosenberg, M., and Court, D. (1984). *J. Mol. Biol.* 176, 39-53.
- Altuvia, S., Locker-Giladi, H., Koby, S., Ben-Nun, O., and Oppenheim, A.B. (1987). *Proc. Natl. Acad. Sci. USA* 84, 6511-6515.
- Daniels, D.L., Subbarao, M.N., Blattner, F.R., and Lozeron, H.A. (1988). *Virology* 167, 568-577.
- Steege, D.A., Cone, K.C., Queen, C., and Rosenberg, M. (1987). *J. Biol. Chem.* 262, 17651-17658.
- Liao, S., and McClure, W. (1988). In Cech, T. (ed.), *Molecular Biology of RNA*, UCLA Symp. Mol. and Cell Biol., New Series, Alan R. Liss, Inc., New York, Vol. 94, pp. 289-297.
- Krinke, L., and Wulff, D.L. (1987). *Genes Dev.* 1, 1005-1013.
- Krinke, L., and Wulff, D.L. (1990). *Genes Dev.*, in press.
- Blomberg, P., Wagner, E.G.H., and Nordström, K. (1990). Submitted to *EMBO J.*
- Fien, K., Turck, A., Kang, I., Kielty, S., Wulff, D.L., McKenney, K., and Rosenberg, M. (1984). *Gene* 32, 141-150.
- Daniels, D.L., Schroeder, J.L., Szybalski, W., Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.F., Peterson, G.B., and Blattner, F.R. (1983). In Hendrix, J., Roberts, J., Stahl, F., and Weisberg, R. (eds.), *Lambda II*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp. 519-676.
- Debouck, C., Riccio, A., Schumperli, D., McKenney, K., Jeffers, J., Hughes, C., and Rosenberg, M. (1985). *Nucl. Acids Res.* 13, 1841-1853.
- Kalnins, A., Otto, K., Rüther, U., and Müller-Hill, B. (1983). *EMBO J.* 2, 593-597.
- Krinke, L., and Wulff, D.L. (1990). Submitted for publication.
- Faubladier, M., Cam, K., and Bouché, J.P. (1990). *J. Mol. Biol.* 212, 461-471.
- Pragai, B., and Apirion, D. (1981). *J. Mol. Biol.* 153, 619-630.
- Rould, M.A., Perona, J.J., Söll, D., and Steitz, T.A. (1989). *Science* 246, 1135-1142.
- Seeman, N.C., Rosenberg, J.M., and Rich, A. (1976). *Proc. Natl. Acad. Sci. USA* 73, 804-808.