Specificity of the Hairpin Ribozyme

SEQUENCE REQUIREMENTS SURROUNDING THE CLEAVAGE SITE*

(Received for publication, June 15, 1999)

Mercedes Pérez-Ruiz‡, Alicia Barroso-delJesus‡§, and Alfredo Berzal-Herranz¶

From the Instituto de Parasitología y Biomedicina "López-Neyra," Consejo Superior de Investigaciones Científicas, Ventanilla 11, 18001 Granada, Spain

Substrate sequence requirements of the hairpin ribozyme have been partially defined by both mutational and in vitro selection experiments. It was considered that the best targets were those that included the N↓GUC sequence surrounding the cleavage site. In contrast to previous studies that failed to evaluate all possible combinations of these nucleotides, we have performed an exhaustive analysis of the cleavage of 64 substrate variants. They represent all possible sequence combinations of the J2/1 nucleotides except the well established G↓1. No cleavage was observed with 24 sequences. C↓1 variants showed little or no cleavage, whereas U↓2 substrates were all cleavable. The maximal cleavage rate was obtained with the AGUC substrate. Cleavage rates of sequences HGUC (H = A, C, or U), GGUN, GGGR (R = A or G), AGUu, and UGUA were up to 5 times lower than the AGUC one. This shows that other sequences besides NGUC could also be considered as good targets. A second group of sequences WGGG (W = A or U), UGUK (K = G or U), MGAG (M = A or C), AGUA, and UGGA were cleaved between 6 and 10 times less efficiently. Furthermore, the UGCU sequence of a non-cleavable viral target was mutated to AGUC resulting in a proficiently cleavable substrate by its cognate hairpin ribozyme. This indicates that our conclusions may be extrapolated to other hairpin ribozymes with different specificity.

The hairpin ribozyme belongs to the group of small trans-acting catalytic RNAs. They are considered promising candidates for the development of specific tools for RNA inactivation. Numerous studies aimed at the inactivation of targeted RNAs by the hairpin ribozyme have been carried out with varying success. Among other factors, the extension of cleavage of selected targets would be affected by the sequence at the region containing the cleavage site within the substrate (J2/1 in Fig. 1). Results based on both in vitro selection and mutational experiments carried out with the hairpin catalytic motif derived from the (-)sTRSV1 established the substrate sequence for optimal cleavage as 5’-RNY↓GHY-3’ (1) (where N is any nucleotide; R is A or G; Y is C or U; and H is A, C, or U as defined by the International Union of Biochemistry). Although in vitro selection strategies allow the analysis of a large number of sequences, the authors did not evaluate every possible sequence combination. Indeed, we have observed that substrate sequences following this consensus are not cleavable by the hairpin ribozyme.° Similarly, Hampel and co-workers (2) defined the N↓GUC consensus as the optimal sequence for trans-cleavage by the hairpin ribozyme. However, sequence restrictions for these nucleotides surrounding the cleavage site were mostly determined by the effect of individual mutations. Therefore, although each consensus defined could be acceptable, it might lead to an under or overestimation of the proficiency of certain sequences to be used as substrates for the hairpin ribozyme. In this work we wanted to investigate the substrate sequence requirements at the J2/1 region to redefine sequence specificity of the hairpin ribozyme. Our results show that other sequences different from NGUC are also efficiently cleaved by the ribozyme. In addition, sequences fitting the established consensus (1) cannot be used as substrates for the hairpin catalytic motif. These results are of great importance in targeting RNA.

EXPERIMENTAL PROCEDURES

Construction of DNA Templates and RNA Synthesis—Oligodeoxyribonucleotides were synthesized on an Oligo 1000 DNA Synthesizer (Beckman Instruments), purified by electrophoresis on a 10–15% 7 M urea polyacrylamide gel, visualized by UV shadowing, excised out, and eluted overnight at 37°C in 500 mM ammonium acetate, 0.1% SDS, and 1 mM EDTA. RNAs were recovered by sequential extraction with phenol and chloroform-isooamyl alcohol, and ethanol precipitated in 0.3 M sodium acetate, pH 5.2. Partially degenerate oligonucleotide for the synthesis of 64 substrate variants (GAGGA TCCTT TAAAC AGNNC GTTC; mutJ2/1) was converted to double-stranded DNA after annealing to E-T7 (GAGAA TTCTA ATACG AACTCA CTATA) and was cloned into the EcoRI-BamHI site of pUC19 to generate the pT7-MUTJ2/1 plasmid series. Transcription of each of the 64 DraI-digested pT7-MUTJ2/1 plasmids yielded the corresponding 17-nt-long RNA substrate variant, carrying an extra GCG sequence at the 5’-end.

Substrate sequences RWTGACCA GGTAA TATAC CACCA CGTCG TTCTC TCGT GCACCA CTATA GTGAG TCGTA TTA were used as templates for the synthesis of wild-type and TAR ribozymes, respectively, prior to annealing to T7p (TAATAC GACT CACTA TA). Similarly, UGCU-TAR (GGCT GACGT GCGCG GTGGA TGTCT TAGCT CATTATA ATC) and AGUC-TAR (GGCTT AGACT GTTGG CCAT ATGTA ATCGG ATAGT ATC) were used to obtain short UGUC-TAR (17 nt) and AGUC-TAR (14 nt) RNA substrates, respectively. Long UGCU-TAR substrate was obtained by transcription of the BamHI-digested pG3TAR (3). pG3TAR was polymerase chain reaction-amplified to generate the DNA template for the synthesis of long AGUC-TAR substrate. Sense and antisense primers for the polymerase chain reaction were T7p and 3’-TAR (GAATATCGTTAGCTATGACCGTGGTGTG).

* This work was supported by Grant PB-96-0825 from the Spanish Dirección General de Enseñanza Superior and Grant 98/112-00 from the Fundación “la Caixa” (to A. B.-H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Contributed equally to this work.

§ Supported by a fellowship from the Spanish Ministerio de Educación y Cultura.

¶ To whom correspondence should be addressed. Tel.: 34-958-80-51-87; Fax: 34-958-30-33-23; E-mail: Alba@ipb.csic.es.

1 The abbreviations used are: (-)sTRSV, negative strand of the satellite RNA associated with the tobacco ring spot virus; TAR, trans-activation response element; nt, nucleotide; Rz, ribozyme.

2 M. Pérez-Ruiz and A. Berzal-Herranz, unpublished data.
Substrate Specificity of the Hairpin Ribozyme

RESULTS AND DISCUSSION

Pre-steady-state kinetics were carried out, as described under “Experimental Procedures,” to determine the ability of the hairpin ribozyme to cleave all possible variants. We performed an exhaustive analysis of the ability of the hairpin ribozyme to cleave 64 of 216 possible sequence combinations of the four nucleotides surrounding the cleavage site (Fig. 1). In this study, position +1 was fixed to a guanosine residue, as it has been shown to be essential for catalytic activity (4, 5). Substrate variants were synthesized from the DraI-digested pTT-MUTJ2/1 plasmid series and assayed against the in vitro synthesized trans-cleaving hairpin ribozyme corresponding to the 50-nt-long sequence of the (-)sTRSV catalytic motif (Fig. 1) (6, 7).

Cleavage Rates—Pre-steady-state kinetics were carried out, as described under “Experimental Procedures,” to determine and compare the ability of the hairpin ribozyme to cleave each substrate variant. The percentage of cleavage as a function of time was fitted by nonlinear regression to a single exponential equation as follows (SigmaPlot 4.0 software; Fig. 2 shows two examples).

\[ y = y_0 + a(1 - \exp^{-bt}) \]  

(Eq. 1)

Where \( y \) is the fraction cleaved, \( a \) is the amplitude, \( x \) is the time, and \( b \) is the reaction rate constant (min\(^{-1}\)).

Forty of the 64 substrate variants were cleavable by the hairpin ribozyme (Fig. 3), and cleavage fragments were of the expected sizes for all of them (Fig. 2 and data not shown). Under our conditions we did not observe cleavage at other phosphodiester linkage than the one between nucleotides +1 and −1 (Fig. 1) in any of the substrates assayed. Substrate variant carrying the sequence AGUC in loop J2/1, which corresponds to the sequence in (-)sTRSV (wild-type in this work) showed the highest cleavage rate (shown by an asterisk in Fig. 3). The catalytic activity of the hairpin ribozyme against the other substrate variants was compared with the cleavage rate of the wild-type one. Kinetic parameters obtained for other substrate variants indicated that other sequences besides AGUC could also be considered as good targets for trans-cleavage (Fig. 3). In addition, it is worth noting that all U→2 sequences were cleavable. Here we discuss only those substrates processed by the hairpin ribozyme with cleavage rates up to 10 times lower than the wild-type one. This group comprises 19 different sequences, and in general, cleavage of these substrates also reached high amplitude values. These variants have been further classified into two subgroups; variants HGUC (H = A, C, or U), GGUN (N = any nucleotide), GGR (R = A or G), AGUU, and AGUA with a rate ranging from 1 to 5 times lower than the wild-type substrate fall within the first subgroup (shown on a black background in Fig. 3). Amplitude values ranged from 0.75 to 1, except in the case of substrates with GGUC and UGUA sequences whose amplitudes were 0.49 and 0.34, respectively. The second subgroup consists of substrate variants WGGG (W = A or U) with amplitude values of 0.87 and 0.85; UGUK (K = G or U) with amplitudes of 0.75 and 0.83; MGAG (M = A or C) with amplitudes of 0.62 and 0.53; and AGUA and UGGA with amplitude values of 0.43 and 0.38,
respectively (shown on a gray background in Fig. 3). These variants were all cleaved with rates between 5 and 10 times lower than the AGUC substrate.

As it has been shown that the best theoretical cleavage sites are not always accessible to ribozymes in complex molecular contexts (full-length RNAs) (8), it would be of great interest to have at our disposal alternative sequences for targeting purposes. We can conclude from our work that, besides the AGUC (or wild-type) sequence, variants fitting the NGUC consensus can be efficiently cleaved by the (-) sTRSV hairpin ribozyme. In addition, we have identified some more variants that were cleaved with comparable rates to the previous ones, and therefore they should be considered for in vivo targeting of mRNAs as well. Indeed, variants GGUU and UGUA were cleaved with higher rates than CGUC or GGUC, but they have been systematically excluded in the search for targeting sites (9).

Noncleavable Substrates—We failed to detect any cleavage products in 24 sequences after 1 h of reaction (Table I). Those substrate variants have been excluded from Fig. 3. Some general rules for sequence requirements can also be extrapolated from the noncleavable substrate variants. It was remarkable that sequences containing a C\(^3\) showed little or no cleavage. In general, no detectable cleavage products were seen for substrates containing NGCN (Table I). AGCA, AGCC, and UGCC sequences constituted exceptions and showed cleavage products, though their cleavage rates were between 20 and 27-fold lower than for the wild-type substrate, and the maximal extent of cleavage ranged between 0.18 and 0.25 (Fig. 3). Other noncleavable substrates were UGAN, CGAW, NGGU, and CGGC (Table I). Interestingly, under our experimental conditions, sequences completely compatible with the established consensus (1) were noncleavable (e.g. CGAU or UGAN).

### Extent of Cleavage

The analyses presented here also provided information about the maximal extent of cleavage of each substrate variant. This feature is described by the amplitude of the reaction. The final cleavage percentage is highly variable (from 0.14 to 1.0; Fig. 3). Low amplitudes (<0.40) usually correspond to rather poor cleavage rates, whereas the highest cleavage rates exhibited high amplitudes as well. The UGUA sequence constitutes the only exception, because it is cleaved with one of the highest rates (0.502), but the amplitude is only about 0.34 (Fig. 3). A common consensus for the 12 sequences with lower amplitudes (<0.40) cannot be extrapolated. How-

---

### Table I
**Noncleavable substrates**

| Position -1 | A | C | G | U |
|-------------|---|---|---|---|
| A           | A | C | G | U |
| C           | CGA | C | G | U |
| G           | CGAU | CCGA | CCGC | CCGG |
| U           | UGAA | UGCA | UGAC | UGAG |

---

### FIG. 3.
Cleavable substrates. The left side of the figure shows the kinetic parameters for the cleavage reaction. A graphical display of the data is also shown on the right. Cleavage rates have been normalized to the maximum value for the graphical representation. Sequences shown on a black background correspond to the ones cleaved between 1 and 5 times less efficiently than the wild type, which is indicated with an asterisk. Sequences shown on a gray background correspond to the ones cleaved between 5 and 10 times less. Data included in the figure for each parameter represent the average of at least three independent experiments ± S.D. The r\(^2\) determination of goodness of fit exceeded 0.99 in all cases except for five substrate variants in which it was 0.985. The standard error for the fitted parameters was less than 10%. Interexperimental error was within 20%.

---

**TABLE I**

| SEQUENCE | RATE (min\(^{-1}\)) | AMPLITUDE |
|----------|---------------------|-----------|
| AGAA     | 0.047 ± 0.003       | 0.701 ± 0.015 |
| AGAC     | 0.046 ± 0.004       | 0.772 ± 0.026 |
| AGAG     | 0.085 ± 0.004       | 0.618 ± 0.017 |
| AGAU     | 0.022 ± 0.004       | 0.388 ± 0.037 |
| AGCA     | 0.022 ± 0.003       | 0.259 ± 0.019 |
| AGCC     | 0.031 ± 0.005       | 0.187 ± 0.013 |
| AGGU     | 0.046 ± 0.010       | 0.779 ± 0.007 |
| AGGC     | 0.022 ± 0.005       | 0.207 ± 0.028 |
| AGGG     | 0.066 ± 0.004       | 0.372 ± 0.018 |
| AGUA     | 0.092 ± 0.013       | 0.437 ± 0.003 |
| AGUC     | 0.667 ± 0.021       | 0.655 ± 0.013 |
| AGUG     | 0.037 ± 0.001       | 0.385 ± 0.007 |
| AGUU     | 0.128 ± 0.007       | 1.018 ± 0.026 |
| CGAC     | 0.018 ± 0.001       | 0.719 ± 0.027 |
| CGAG     | 0.101 ± 0.003       | 0.635 ± 0.011 |
| CGGA     | 0.016 ± 0.003       | 0.640 ± 0.037 |
| CGGG     | 0.023 ± 0.002       | 0.896 ± 0.029 |
| CGUA     | 0.039 ± 0.003       | 0.832 ± 0.021 |
| CGUC     | 0.318 ± 0.006       | 0.913 ± 0.011 |
| CGUG     | 0.011 ± 0.003       | 0.586 ± 0.014 |
| CGUU     | 0.025 ± 0.001       | 0.918 ± 0.014 |
| GGGA     | 0.146 ± 0.008       | 0.766 ± 0.018 |
| GGGG     | 0.029 ± 0.003       | 0.285 ± 0.013 |
| GGGG     | 0.243 ± 0.016       | 0.691 ± 0.023 |
| GGUC     | 0.175 ± 0.008       | 1.056 ± 0.032 |
| GGUG     | 0.204 ± 0.018       | 0.495 ± 0.023 |
| GGUU     | 0.515 ± 0.007       | 0.921 ± 0.019 |
| GUGC     | 0.149 ± 0.010       | 0.995 ± 0.059 |
| GUGA     | 0.027 ± 0.007       | 0.229 ± 0.024 |
| UGGG     | 0.094 ± 0.003       | 0.384 ± 0.007 |
| UGUC     | 0.017 ± 0.001       | 0.761 ± 0.042 |
| UUGG     | 0.110 ± 0.005       | 0.849 ± 0.007 |
| UUGA     | 0.502 ± 0.031       | 0.340 ± 0.015 |
| UGAC     | 0.545 ± 0.014       | 0.750 ± 0.032 |
| UGAG     | 0.098 ± 0.002       | 0.752 ± 0.007 |
| UGUA     | 0.072 ± 0.001       | 0.833 ± 0.005 |
| UGUC     | 0.098 ± 0.002       | 0.752 ± 0.007 |
| UGUU     | 0.072 ± 0.001       | 0.833 ± 0.005 |
ever, it is important to point out two conclusions. First, this group does not include sequences with U at position +2 with the only exception of the previously mentioned UGUA variant. Second, the four GGAN sequences belong to this group with amplitudes ranging from 0.14 to 0.38.

All the experiments were performed under single-turnover conditions, so it is very unlikely that substrate intermolecular associations could be the reason for the wide amplitude variability. On the other hand, after a careful sequence analysis, we can affirm that intramolecular structures (like hairpins) are not limiting the availability of substrate. Therefore, as all sequences share identical helix 1 and helix 2 regions (Fig. 1), all variants should have similar intrinsic capability of binding to the ribozyme. It has been proposed that the fraction of substrate that cannot be cleaved could be trapped into a nonactive conformation of the ribozyme-substrate complex (10). The final extent of cleavage would depend on the magnitude of the misfolded subpopulation. It is clear then, that the J2/1 region is somehow participating in the correct folding of the complex. Although some work has been carried out to elucidate this matter, the role of these nucleotides in the three-dimensional structure of the substrate-ribozyme complex still remains unclear (11). A high resolution structural analysis would be required to elucidate the role of this region and to explain the observed differences. The data presented here support both catalytic and structural roles for the J2/1 nucleotides.

**Catalytic Activity of Hairpin Ribozymes with Different Specificity—**All the variants analyzed in this work shared common sequences at both ribozyme binding regions (H1 and H2 in Fig. 1). We wanted to test whether these conclusions might be of general application. A 14-nt sequence that fulfilled the proposed substrate consensus (1) was previously identified at the 3′-end of the human immunodeficiency virus, type 1 TAR region (Fig. 4). Attempts to cleave this sequence by the corresponding hairpin ribozyme had been unsuccessful.2 Interestingly, the J2/1 sequence of this putative target was UGCU, which corresponds to one of the noncleavable sequences identified in this work (Figs. 3 and 4A). We tested whether the specific hairpin ribozyme (Rz-TAR) could cleave a mutated 92-nt-long human immunodeficiency virus, type 1 RNA carrying an AGUC sequence at the corresponding J2/1 region (long AGUC-TAR substrate, Fig. 4). The AGUC sequence yielded the highest cleavage rate in the analysis presented above (Fig. 3). Significant cleavage of the AGUC-TAR substrate was observed, whereas no cleavage products of the UGCU-TAR substrate were detected under the same conditions (Fig. 4B). These results indicate that the J2/1 nucleotides might determine the availability of an RNA target to be cleaved by the hairpin ribozyme. However, we could not rule out the possibility that other factors could be affecting the ability of the ribozyme to cleave the UGCU-TAR. Thus, the three nucleotide change might bring about a conformational change of the RNA structure altering the accessibility of the target region. This feature was evaluated by targeting two short RNA molecules carrying either the UGCU or the AGUC-TAR sequence (Fig. 4C). As expected, no cleavage products were detected for UGCU-TAR, whereas significant cleavage was observed for the AGUC variant. These data suggest that the results presented above might be, at least qualitatively, applied to other ribozyme-substrate systems different from the one derived from the (-)sTRSV.

**Conclusions—**Our results indicate that the substrate sequence requirements for the hairpin ribozyme at the J2/1 region established for cis-cleaving may not be applicable to trans-cleaving ribozymes. Furthermore, substrate specificity is not restricted to NGUC. We have identified a number of sequences with cleavage efficiencies comparable to NGUC. The hairpin ribozyme could cleave up to 40 different sequences, which is of great interest for ribozyme targeting as the putative good targets may not be accessible. Equally important is the identification of 24 sequences that cannot be cleaved by this catalytic motif.

**Acknowledgments—**We thank V. Augustin for excellent technical assistance.

**REFERENCES**

1. Berzal-Herranz, A., Joseph, S., Chowrirra, B. M., Butcher, S. E., and Burke, J. M. (1993) EMBO J. 12, 2567-2574
2. Anderson, P., Monforte, J., Tritz, R., Nesbitt, S., Hearst, J., and Hampel, A. (1994) Nucleic Acids Res. 22, 1096-1100
3. Pérez-Ruiz, M., Sievers, D., García-López, P. A., and Berzal-Herranz, A. (1998) *Antisense Nucleic Acid Drug Dev.* **9**, 33–42
4. Chowrim, B. M., Berzal-Herranz, A., and Burke, J. M. (1991) *Nature* **354**, 320–322
5. Berzal-Herranz, A., Joseph, S., and Burke, J. M. (1992) *Genes Dev.* **6**, 129–134
6. Feldstein, P. A., Buzayan, J. M., and Bruening, G. (1989) *Gene (Amst.)* **82**, 53–61
7. Hampel, A., and Tritz, R. (1989) *Biochemistry* **28**, 4929–4933
8. Yu, Q., Pecchia, D. B., Kingsley, S. L., Heckman, J. E., and Burke, J. M. (1998) *J. Biol. Chem.* **273**, 23524–23532
9. Hampel, A. (1998) *Prog. Nucleic Acid Res. Mol. Biol.* **58**, 1–39
10. Esteban, J. A., Banerjee, A. R., and Burke, J. M. (1997) *J. Biol. Chem.* **272**, 13629–13639
11. Earnshaw, D. J., Masquida, B., Muller, S., Sigurdsson, S. Th., Eckstein, F., Westhof, E., and Gail, M. J. (1997) *J. Mol. Biol.* **274**, 197–212
12. Feng, S., and Holland, E. C. (1988) *Nature* **334**, 165–167