RNA Structures

DOI: 10.1002/ange.200601311

The Roles of Hydrogen Bonding and Sterics in RNA Interference**

*Alvaro Somoza, Jijumon Chelliserrykattil, and Eric T. Kool**

RNA interference (RNAi) has become one of the most important new tools for biological research in the past decade.[1–5] This methodology is used widely for "knocking down" (regulating) expression of specific genes in cell cultures, and it can be carried out conveniently by using synthetic RNA oligonucleotides (short interfering RNAs or siRNAs) that are complementary to a segment of a desired messenger RNA target. When the appropriate double-stranded 21 mer RNAs (the sense and antisense strands) are added to a cell culture, they are taken up by the cellular RNA-induced silencing complex (RISC), which presents the separated antisense ("guide") strand for binding and subsequent cleavage of the target mRNA.[1–5,6] One of the most useful features of this approach is that the resulting mRNA cleavage occurs with sequence selectivity[7,8] so that one gene can often be knocked down to low levels of activity with little effect on the rest of cellular gene expression.[9,10]

Recent RNA-interference studies with mismatched target RNAs have demonstrated sequence selectivity (at the single-nucleotide level) at many positions on the standard 21-nucleotide probe length.[8] The origins of this selectivity are not known; selectivity may arise from base-pair hydrogen bonding, which contributes to selective hybridization,[11] or from steric complementarity of nucleobases, which is important in replication by DNA polymerases.[12] Furthermore, selectivity could come chiefly from the RNA itself, or could be modulated by the RISC complex. Beyond this, it is not clear why responsiveness to mismatches varies along the siRNA strand.[8] Although recent structural studies of Piwi/Argonaute/Zwille (PAZ) domains[13] and *Argonaute* proteins[14] (parts of the RISC complex) have led to models of RNA cleavage, no structures of the RISC complex bound to siRNA and mRNA are yet available.

Herein, we describe experiments that give insights into the origins of RNAi activity and selectivity. We have evaluated this with a nonpolar, non-hydrogen-bonding ribonucleoside isostere (rF; Figure 1)[15,16] that we have incorporated into a siRNA guide strand in place of natural uridine. We found that this analogue can maintain near-wild-type activity in human cells at a number of positions in the strand and importantly, we observe that it can retain and even enhance sequence selectivity.

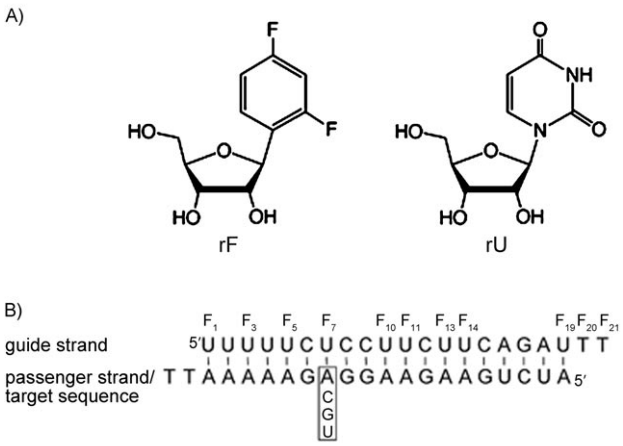

Figure 1. Modified RNA structures and sequences used in this study. A) 2,4-Difluorobenzene ribonucleoside rF, a uridine nonpolar isostere, is shown next to uridine (rU). B) Sequences of siRNA duplexes used in RNAi experiments. Sites of rF substitutions in various experiments are marked with "F_n". The "guide"-strand sequence (antisense to the mRNA) is above; "passenger" strands (corresponding to the target mRNA and mutants) are below. The shown sequence corresponds to nucleotides 501–519 in *Renilla* luciferase mRNA.

The analogue rF contains difluorobenzene, a uracil isostere, in place of the earlier-used difluorotoluene deoxyriboside (dF) as a thymidine mimic in DNA.[17] This nonpolar structure serves as a probe for the importance of hydrogen bonding and electrostatics in the RNA context.[15,16,18] RNA hybridization studies have shown that rF pairs have little or no inherent selectivity and are destabilizing to the RNA duplex[15] consistent with its nonpolar properties; this is also consistent with the behavior of dF in DNA.[19] In separate experiments, we incorporated rF into eleven different positions in place of natural rU along an RNA guide strand that was complementary to a luciferase reporter gene in an A-rich site (Figure 1).

Thermal denaturation studies of synthetic RNA duplexes showed that rF is destabilizing to 21 mer double-stranded siRNAs when it is placed near the center (see the Supporting Information). However, this destabilization lessened near the duplex ends (positions 1, 3, and 19), and virtually no destabilization was seen at unpaired positions 20 and 21. This is similar to the positional effects of mismatches on the helix stability of RNA duplexes,[20] and is reminiscent of the behavior of dF in DNA.[11] To measure selectivity, studies were also carried out with RNAs mismatched at a central position (position 7), where destabilization by rF is high. Results confirmed (see the Supporting Information) that this nonpolar nucleoside gives no pairing selectivity for adenine.[15,16,18,19] By comparison, natural uridine showed substantial pairing selectivity for adenine.

We then carried out separate RNA-interference studies in HeLa cells with the eleven modified guide-RNA strands

[*] Dr. A. Somoza, Dr. J. Chelliserrykattil, Prof. Dr. E. T. Kool
Department of Chemistry, Stanford University
Stanford, CA 94305-5080 (USA)
Fax: (+1) 650-725-0259
E-mail: kool@stanford.edu

[**] This work was supported by the US National Institutes of Health (GM072705), and by a Marie Curie OIF Fellowship (6th Framework Program, UE) to A.S.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

containing single rF substitutions. Experiments were carried out in triplicate. The cells were first transfected with dual reporter plasmids that express *Renilla* luciferase and firefly luciferase, the activities of which could be separately quantitated by their luminescence.[21] The synthetic RNAs were complementary to positions 501–519 of *Renilla* luciferase mRNA[22] and not complementary to the firefly mRNA. The effects of the different RNAs on luciferase expression were evaluated after dosing with 0.21–21 ng of RNA in the cell media, and measuring relative luminescence responses after 22 h. The firefly luminescence is an internal control that rules out possible indirect effects of the different siRNAs, such as selective cell toxicity, variable cellular uptake, or differential enzymatic degradation.

Results showed that the rF substitution disrupted *Renilla* luciferase-specific RNA interference activity at two central positions of the guide RNA; surprisingly, however, a number of other positional substitutions retained near-wild-type activity (Figure 2). Substitutions of rF at positions 10 and 11

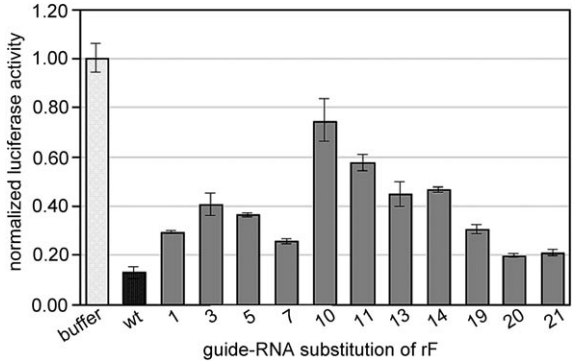

Figure 2. Histogram of gene-specific RNA interference activity for rF-substituted siRNAs at one site in the *Renilla* luciferase mRNA expressed in HeLa cells. The numbers refer to the position of rF in the guide-RNA strand. Data were normalized by an internal control of a noncomplementary firely luciferase gene. Data are for the 2.1-ng RNA amount, and were measured in triplicate; standard deviations are as indicated. See the Supporting Information for the full data. wt = wild type.

caused loss of most of the interference activity even with a higher 21-ng dose of siRNA (see the Supporting Information); these positions flank the expected cleavage site in the target mRNA.[3] In contrast to this, all other substitutions retained substantial activity within 2–4-fold of the natural uracil-substituted strand. Especially remarkable is substitution at position 7, which showed near-wild-type activity despite the strong destabilization that this substitution causes (Figure 3). At the higher (21 ng) dose, guide RNAs containing rF at positions 1, 7, 19, 20, and 21 all showed near-wild-type activity (see the Supporting Information). Intermediate effects were found at positions 3, 5, 13, and 14. These results suggest that canonical hydrogen bonding is not necessary at several positions for the RISC complex to maintain high levels of RNA interference activity with the intended target.

The retention of strong activity at position 7 with the rF analogue, despite its central location in the siRNA, offers a

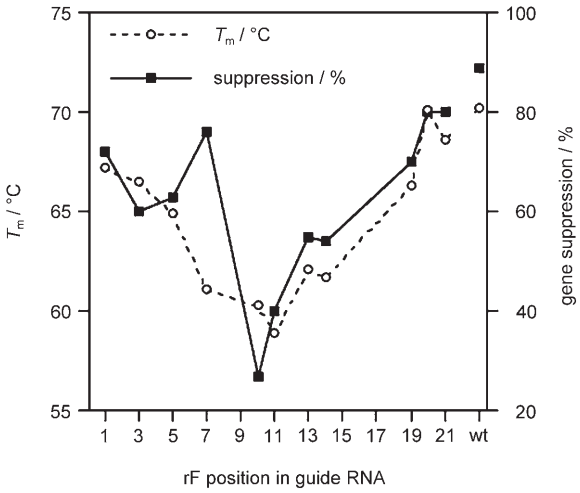

Figure 3. Dual plot comparing helix stability (T_m) and RNA gene knockdown activity (% suppression) with rF substitution along the guide strand of siRNA duplexes at the positions shown. Data at the far right ("wt") are for the natural rU-substituted siRNA. See the Supporting Information for the original T_m data. Suppression data are from 2.1-ng siRNA amounts (see the Supporting Information).

unique opportunity to test the chemical origins of sequence selectivity in RNA interference. The corresponding position in the *Renilla* luciferase mRNA is universally variable in the gly 171 codon. We therefore prepared three new mutant plasmids encoding *Renilla* luciferase with singly mismatched mRNAs. Measurement of activity with these four targets showed that the naturally substituted guide RNA did indeed distinguish single mismatches (Figure 4); the mismatches led to as much as a 2-fold drop in activity at a 2.1-ng RNA dose. This is consistent with a report for a different RNA target in which target mutations at position 7 led to significant selectivity.[8] In our experiments, a U–U mismatch was well tolerated, whereas a U–C mismatch was most disruptive to activity.

Remarkably, experiments with rF (position 7) showed that it also displayed selectivity, and more surprisingly, the level of selectivity was higher than that of the natural base. A selectivity of nearly threefold greater that the F–C mismatch was seen (Figure 4), and the rF-containing strand also showed significant selectivity for cleavage of the G mutant, which rU discriminated poorly. Thus, we conclude that, at least at position 7 of the RISC complex, mRNA-target-sequence selectivity does not require canonical Watson–Crick hydrogen bonding.

These experiments give new insight into the mechanism of RNA interference. We have shown that a nonpolar nucleobase analogue can maintain substantial cellular RNAi activity at nine of eleven sites tested, establishing clearly that canonical Watson–Crick hydrogen bonding is not crucial at all positions. We do note a general correspondence of T_m with mRNA-suppression activity, however (Figure 3), suggesting that not many simultaneous substitutions in guide RNA would be well tolerated. Notably, two positions tested (7 and 10) deviate markedly from this correlation, suggesting considerable changes in RNA–protein interactions across this

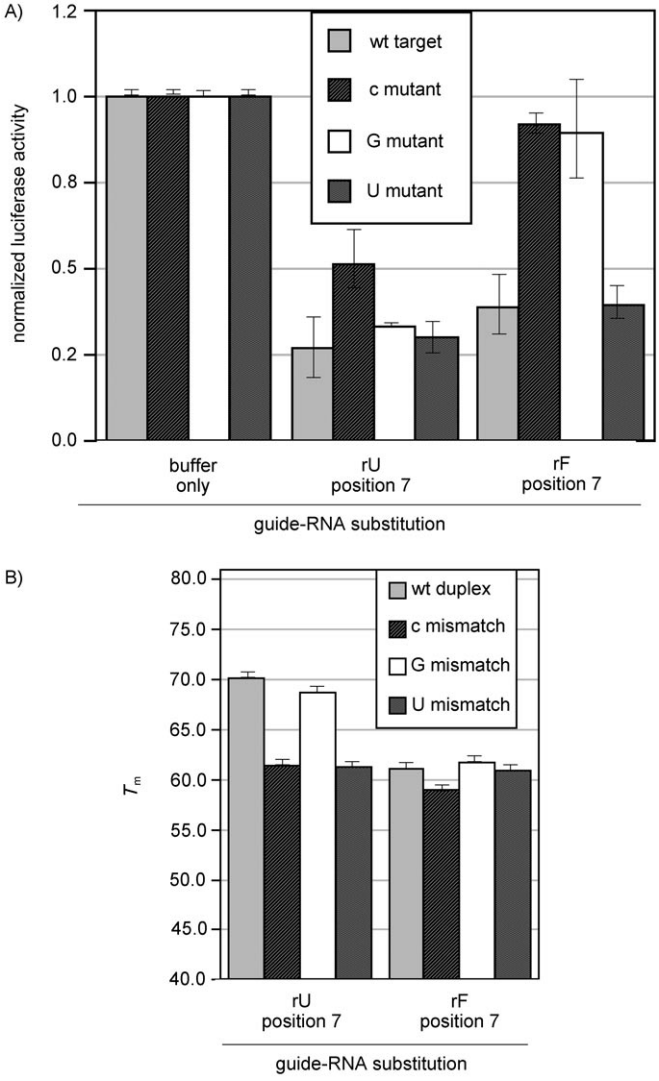

Figure 4. Selectivity of modified siRNA guide strands for singly mismatched RNA targets. The mismatch is opposite position 7 of the guide RNA and involves mutating the corresponding nucleotide in the complementary RNA. A) Histogram of RNAi activity showing rF-enhanced selectivity for adenine relative to mutations at position 513 in the luciferase target mRNA (position 513 is opposite position 7 on the guide RNA). B) Thermal stability histogram showing inherent hybridization selectivity of rU, and lack of selectivity of rF, as measured by thermal denaturation of the guide/passenger duplexes corresponding to the mutations in part (A). Raw data are given in the Supporting Information. T_m (melting temperature) represents the RNA duplex stability.

localized part of the RNA. The overall profile of rF-substitution responses suggests that guide-RNA positions 10 and 11, in particular, may require hydrogen bonding or base-pairing stability with the target (or both) for efficient RNA cleavage activity.

After submission of this manuscript, a report appeared testing difluorotoluene riboside, closely related to the current rF, in siRNA activity.[18] This study found similar positional effects on RNA interference, with the base analogue causing strong disruption of activity at positions 10 and 11, and little or no effect at position 7. The consistency of findings between the two studies, carried out with different targets, suggests that these effects are general for the RISC complex. Sequence selectivity with mismatched targets was not examined in that study.

The current results strongly suggest that at position 7, the sequence selectivity of RNA interference arises not from the selectivity of hydrogen bonding, nor from selectivity caused by the RNA backbone alone. A possible hypothesis is that the selectivity arises instead from an enforced steric selection by the RISC complex (and *Argonaute* in particular[14]). Experiments with base analogues with varied size or shape could test this steric hypothesis explicitly. At this point, it is unknown whether selectivity at other positions depends on hydrogen bonding or not.

Finally, we note with interest the observation that use of rF in place of rU at position 7 appears to enhance sequence selectivity beyond that of the natural base. We suggest that this may be due to the general destabilization of the guide–target duplex by the nonpolar base analogue; it is known that a nearby mismatch can increase specificity of DNA hybridization,[23] perhaps by a similar destabilization mechanism. In any case, the result suggests a more-general practical utility of nonpolar isosteres in RNA interference studies.

Experimental Section

Nucleoside phosphoramidite synthesis: The rF ribonucleoside was prepared as described.[15,16] For automated RNA synthesis, the previously unknown 5-dimethoxytrityl, 2′-*O*-TOM (TOM = [(triisopropylsilyl)oxy]methyl), 3′-*O*-phosphoramidite derivative was prepared. Details are given in the Supporting Information.

Modified oligoribonucleotides: RNA sequences were made following standard protocols for 2′-*O*-TOM-protected phosphoramidites[24] and were purified by polyacrylamide gel electrophoresis and characterized by MALDI-TOF mass spectrometry. Details are supplied in the Supporting Information.

RNA interference studies and mismatched luciferase vectors: The dual reporter Renilla/firefly assay was carried out by using HeLa cells transfected with reporter expression plasmids purchased from Promega (San Luis Obispo, CA). Sequence mutant plasmids were prepared following standard methods and were characterized by sequencing. Details are given in the Supporting Information.

Received: April 4, 2006
Revised: May 16, 2006
Published online: June 27, 2006

Keywords: base pairs · difluorobenzene · isosteres · protein structures · RNA structures

[1] A. Fire, S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, C. C. Mello, *Nature* **1998**, *391*, 806–811.
[2] S. M. Elbashir, J. Martinez, A. Patkaniowska, W. Lendeckel, T. Tuschl, *EMBO J.* **2001**, *20*, 6877–6888.
[3] S. M. Elbashir, W. Lendeckel, T. Tuschl, *Genes Dev.* **2001**, *15*, 188–200.
[4] G. Hutvagner, P. D. Zamore, *Curr. Opin. Genet. Dev.* **2002**, *12*, 225–232.
[5] T. Tuschl, G. Meister, *Nature* **2004**, *431*, 343–349.
[6] M. Manoharan, *Curr. Opin. Chem. Biol.* **2004**, *8*, 570–579.
[7] B. L. Bass, *Cell* **2000**, *101*, 235–238.

[8] Q. Du, H. Thonberg, J. Wang, C. Wahlestedt, Z. Liang, *Nucleic Acids Res.* **2005**, *33*, 1671 – 1677.
[9] J.-T. Chi, H. Y. Chang, N. N. Wang, D. S. Chang, N. Dunphy, P. O. Brown, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6343 – 6346.
[10] D. Semizarov, L. Frost, A. Sarthy, P. Kroeger, D. N. Halbert, S. W. Fesik, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6347 – 6352.
[11] a) B. A. Schweitzer, E. T. Kool, *J. Am. Chem. Soc.* **1995**, *117*, 1863 – 1872; b) E. T. Kool, J. C. Morales, K. M. Guckian, *Angew. Chem.* **2000**, *112*, 1046 – 1068; *Angew. Chem. Int. Ed.* **2000**, *39*, 990 – 1009.
[12] S. Moran, R. X.-F. Ren, E. T. Kool, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10506 – 10511.
[13] a) A. Lingel, B. Simon, E. Izaurralde, M. Sattler, *Nat. Struct. Mol. Biol.* **2004**, *11*, 576 – 577; b) J. B. Ma, K. Ye, D. J. Patel, *Nature* **2004**, *429*, 318 – 322; c) A. Lingel, M. Sattler, *Curr. Opin. Struct. Biol.* **2005**, *15*, 107 – 115.
[14] a) J. J. Song, S. K. Smith, G. J. Hannon, L. Joshua-Tor, *Science* **2004**, *305*, 1434 – 1437; b) J. J. Song, L. Joshua-Tor, *Curr. Opin. Struct. Biol.* **2006**, *16*, 5 – 11.
[15] J. Parsch, J. W. Engels, *Helv. Chim. Acta* **2000**, *83*, 1791 – 1808.
[16] K. M. Guckian, Ph.D. Thesis, University of Rochester (NY), **1998**.
[17] K. M. Guckian, E. T. Kool, *Angew. Chem.* **1997**, *109*, 2942 – 2945; *Angew. Chem. Int. Ed.* **1997**, *36*, 2825 – 2828.
[18] J. Xia, A. Noronha, I. Toudjarska, F. Li, A. Akinc, R. Braich, M. Frank-Kamenetsky, K. G. Rajeev, M. Egli, M. Manoharan, *ACS Chem. Biol.* **2006**, *1*, 176 – 182.
[19] S. Moran, R. X.-F. Ren, S. Rumney, E. T. Kool, *J. Am. Chem. Soc.* **1997**, *119*, 2056 – 2057.
[20] Y. Xu, H.-Y. Zhang, D. Thormeyer, O. Larsson, Q. Du, J. Elmén, C. Wahlestedt, Z. Liang, *Biochem. Biophys. Res. Commun.* **2003**, *306*, 712 – 717.
[21] R. Kierzek, M. E. Burkard, D. H. Turner, *Biochemistry* **1999**, *38*, 14214 – 14223.
[22] G. Grentzmann, J. A. Ingram, P. J. Kelly, R. F. Gesteland, J. F. Atkins, *RNA* **1998**, *4*, 479 – 486.
[23] Z. Guo, Q. Liu, L. M. Smith, *Nat. Biotechnol.* **1997**, *15*, 331 – 335.
[24] a) S. Pitsch, P. A. Weiss, L. Jenny, A. Stutz, X. Wu, *Helv. Chim. Acta* **2001**, *84*, 3773 – 3795; b) S. Jin, C. V. Miduturu, D. C. McKinney, S. K. Silverman, *J. Org. Chem.* **2005**, *70*, 4284 – 4299.

