Nucleic Acids Research, 2005, Vol. 33, No. 5 ***1671–1677***
doi:10.1093/nar/gki312

A systematic analysis of the silencing effects of an active siRNA at all single-nucleotide mismatched target sites

Quan Du*, Håkan Thonberg, Jue Wang[1], Claes Wahlestedt and Zicai Liang

Center for Genomics and Bioinformatics, Karolinska Institutet, 171 77 Stockholm, Sweden and [1]Division of Molecular Toxicology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm 171 77, Sweden

Received December 20, 2004; Revised February 14, 2005; Accepted March 2, 2005

ABSTRACT

The specificity of small interfering RNA (siRNA)-mediated gene silencing is a critical consideration for the application of RNA interference (RNAi). While the discovery of potential off-target effects by siRNAs is of concern, no systematic analysis has been conducted to explore the specificity of RNAi. Here, we present a study where a functionally validated siRNA (siCD46) was examined for silencing specificity on all possible 57 permutated target sites, each carrying a single-nucleotide mutation that would generate a mismatch when paired with siRNA antisense strand. We found that it was not only the position of the mismatched base pair, but also the identity of the nucleotides forming the mismatch that influenced silencing. Surprisingly, mismatches formed between adenine (A) and cytosine (C), in addition to the G:U wobble base pair, were well tolerated and target sites containing such mismatches were silenced almost as efficiently as its fully matched counterpart by siCD46. Northern blots showed that the silencing of fusion genes harboring the mutated target sites involved target mRNA degradation. This study provides direct evidence that the target recognition of siRNA is far more degenerative than previously considered. This finding is instrumental in the understanding of RNAi specificity and may aid the computational prediction of RNA secondary structure.

INTRODUCTION

The discovery that small interfering RNA (siRNA) can silence gene expression through sequence-specific cleavage of the cognate RNA transcript has led to the rapid adoption of RNA interference (RNAi) as a technology for analyzing gene function in mammalian cell culture and tools for drug target validation. There is also high expectation for siRNA as a tool for *in vivo* investigation and as a platform for therapeutic development (1). Target recognition by siRNA was initially thought to be a highly sequence-specific process mediated by the antisense strand (or so-called guide strand) of siRNA duplexes (2), and a single-nucleotide mismatch to the target was reported to abolish the gene silencing effect. This view was further strengthened by the assessment of RNAi specificity using genome-wide expression profiles (3,4). The optimistic view was, however, challenged when significant off-target effects were observed in carefully designed microarray experiments. These studies showed that genes with partial sequence similarities to a siRNA were also down-regulated significantly (5). While off-target effects of siRNAs have been widely discussed, systematic analysis of such effects has been missing. Several mutational analyses have been performed to explore the specificity of RNAi, and found that the terminal nucleotides usually did little to affect the silencing efficacies, whereas some central mutations did abolish the silencing activities of the tested siRNAs (1,5,6). However, in these cases, the conclusions were compromised by the fact that the siRNA sequences, rather than the target sequences, were mutated. As we now know, the efficacy of an siRNA is actually governed by at least two factors, the ability to enter the RNA-induced silencing complex (RISC) and the ability to recognize the target sequences (being either the perfectly matched target or mutated sites) (7). In cases where the siRNA sequences were mutated, it became uncertain whether the loss of silencing activity was caused by alterations in the RISC entry step or in the target recognition step.

Understanding off-target effects is not only important for siRNA design and interpretation of the actual experimental results, but also crucial for the development of siRNAs as drug candidates. In order to systematically explore the specificity of

*To whom correspondence should be addressed. Tel: +46 8 5248 6381; Fax: +46 8 327934; Email: quan.du@cgb.ki.se
Correspondence may also be addressed to Zicai Liang. Tel: +46 8 5248 6371; Fax: +46 8 327934; Email: zicai.liang@cgb.ki.se

RNAi, we chose to investigate the silencing effects of a proven functional siRNA on all 57 permutations of its wild-type target site where each of the mutated sites can form a different single-nucleotide mismatch when paired with siRNA antisense strand. Our results demonstrate that target sites carrying single-nucleotide mutations are silenced to varying degrees and that the silencing efficiency is governed by both the position and the identity of the mismatched base pair.

MATERIALS AND METHODS

Plasmid construction and siRNA target site modification

A modified version of the previously reported siRNA validation vector siQuant (8) was used in this study. Modification consisted of inserting an in-frame ATG start codon before the original *firefly* luciferase gene. The wild-type target site of siCD46 siRNA, corresponding to nucleotides 604–622 of the human CD46 gene (XM_036622), was cloned between the new start codon and the original start codon of *firefly* luciferase gene by PCR. Degenerate oligonucleotides were used to construct 57 different mutated target sites (Table 1). In brief, the PCR products amplified by one of the degenerate forward primers and the site reverse primers (5′-AGTGAGATCTCACAGCCCATGGTGC-3′) were restricted by BglII, gel purified and self-ligated to construct the expression vectors containing various mutated target site of siCD46. The fusion constructs containing wild-type and mutated target sites of siNPY siRNA were prepared exactly according to a previous protocol (8) with the oligonucleotides listed in Table 2. All clones used in this study were verified by sequencing. All DNA oligonucleotides were purchased from biomers.net GmbH (Ulm, Germany). RNA oligonucleotides were obtained from Dharmacon Research (Lafayette, CO). The siRNA duplex was prepared by mixing complementary sense and antisense strand RNA at equal concentration of 50 μM in water. The mixture was then incubated in boiling water for 1 min and cooled overnight to allow formation of siRNA duplex. The quality of the RNA duplexes was assessed on PAGE gel. The sequences of siCD46 are sense, 5′-CTTATTGGAGAGAGCACGA-3′; and guide strand, 5′-TCGTGCTCTCTCCAATAAG-3′. The sequences of the siNPY are sense, 5′-TGAGAGAAAGCACAGAAAA-3′; and guide strand, 5′-TTTTCTGTGCTTTCTCTCA-3′.

Table 1. Oligonucleotides used for constructing mutated target sites of siCD46

Wild-type	5′-TGTGAGATCTCACTTATTGGAGAGAGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-1	5′-TGTGAGATCTCADTTATTGGAGAGAGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-2	5′-TGTGAGATCTCACVTATTGGAGAGAGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-3	5′-TGTGAGATCTCACTVATTGGAGAGAGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-4	5′-TGTGAGATCTCACTTBTTGGAGAGAGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-5	5′-TGTGAGATCTCACTTAVTGGAGAGAGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-6	5′-TGTGAGATCTCACTTATVGGAGAGAGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-7	5′-TGTGAGATCTCACTTATTHGAGAGAGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-8	5′-TGTGAGATCTCACTTATTGHAGAGAGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-9	5′-TGTGAGATCTCACTTATTGGBGAGAGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-10	5′-TGTGAGATCTCACTTATTGGAHAGAGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-11	5′-TGTGAGATCTCACTTATTGGAGBGAGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-12	5′-TGTGAGATCTCACTTATTGGAGAHAGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-13	5′-TGTGAGATCTCACTTATTGGAGAGBGCACGATTGGGCCC-GGCGCCATGG-3′
SITE-14	5′-TGTGAGATCTCACTTATTGGAGAGAHCACGATTGGGCCC-GGCGCCATGG-3′
SITE-15	5′-TGTGAGATCTCACTTATTGGAGAGAGDACGATTGGGCCC-GGCGCCATGG-3′
SITE-16	5′-TGTGAGATCTCACTTATTGGAGAGAGCBCGATTGGGCCC-GGCGCCATGG-3′
SITE-17	5′-TGTGAGATCTCACTTATTGGAGAGAGCADGATTGGGCCC-GGCGCCATGG-3′
SITE-18	5′-TGTGAGATCTCACTTATTGGAGAGAGCACHATTGGGCCC-GGCGCCATGG-3′
SITE-19	5′-TGTGAGATCTCACTTATTGGAGAGAGCACGBTTGGGCCC-GGCGCCATGG-3′

Degenerate nucleotides were shown according to the IUB convention (B = C, G or T; D = A, G or T; H = A, C or T; and V = A, C or G). The regions encoding the target sites were underlined.

Transfection and dual-luciferase assay

Human embryonic kidney cells (HEK293) were grown in DMEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 100 U/ml penicillin and 100 μg/ml streptomycin (all from Life Technologies, GIBCO) and seeded into 24-well plates at $\sim 1 \times 10^5$ cells/ well density 1 day before the transfection. The cells were transfected with use of Lipofectamine2000 (Invitrogen) at ~50% confluence. The fusion constructs (0.17 μg) were co-transfected with 0.017 μg pRL-TK, either with or without siRNA (13 nM final concentration). The cells were harvested after 24 h, according to the dual-luciferase assay manual (Promega) in passive lysis buffer, and the luciferase activities were measured on a fluorometer (Novostar, BMG Labtechnologies GmbH, Germany). The *firefly* luciferase signal was normalized to the *renilla* luciferase signal for each individual well, and the silencing efficacy of each construct was calculated by normalization to a control that was not treated with siRNA. All experiments were performed in triplicates and repeated three times.

Northern blot assay

Twenty-four hours after transfection with fusion constructs, total RNA was harvested from HEK293 cells with RNeasy Mini kit (Qiagen) according to the manufacturer's instruction. Total RNA, 10–12 μg, was separated by electrophoresis in an ethidium bromide-containing agarose–formaldehyde gel. The intensities of the 18S and 28S rRNA bands were checked under ultraviolet light to verify that all samples were loaded equally and that no RNA degradation had occurred. The *firefly* luciferase cDNA was labeled with [α-^{32}P]dCTP using Ready-To-Go DNA labeling kit (Amersham Pharmacia Biotech). Hybridization and stringent washing were performed as described previously (9), and the signals were detected by phosphorimaging on Typhoon 9400 (Amersham). The β-actin cDNA probe was used as the loading control.

Table 2. Oligonucleotides used for constructing mutated target sites of siNPY

NPY08G-F	5′-GATCTCAATGAGAGAGAGCACAGAAAACGGGCC-3′
NPY08G-R	5′-CGTTTTCTGTGCTCTCTCTCATTGA-3′
NPY08T-F	5′-GATCTCAATGAGAGATAGCACAGAAAACGGGCC-3′
NPY08T-R	5′-CGTTTTCTGTGCTATCTCTCATTGA-3′
NPY08C-F	5′-GATCTCAATGAGAGACAGCACAGAAAACGGGCC-3′
NPY08C-R	5′-CGTTTTCTGTGCTGTCTCTCATTGA-3′
NPY13A-F	5′-GATCTCAATGAGAGAAAGCAAAGAAAACGGGCC-3′
NPY13A-R	5′-CGTTTTCTTTGCTTTCTCTCATTGA-3′
NPY13G-F	5′-GATCTCAATGAGAGAAAGCAGAGAAAACGGGCC-3′
NPY13G-R	5′-CGTTTTCTCTGCTTTCTCTCATTGA-3′
NPY13T-F	5′-GATCTCAATGAGAGAAAGCATAGAAAACGGGCC-3′
NPY13T-R	5′-CGTTTTCTATGCTTTCTCTCATTGA-3′
NPY16G-F	5′-GATCTCAATGAGAGAAAGCACAGGAAACGGGCC-3′
NPY16G-R	5′-CGTTTCCTGTGCTTTCTCTCATTGA-3′
NPY16C-F	5′-GATCTCAATGAGAGAAAGCACAGCAAACGGGCC-3′
NPY16C-R	5′-CGTTTGCTGTGCTTTCTCTCATTGA-3′
NPY16T-F	5′-GATCTCAATGAGAGAAAGCACAGTAAACGGGCC-3′
NPY16T-R	5′-CGTTTACTGTGCTTTCTCTCATTGA-3′

Degenerate nucleotides were shown according to the IUB convention (B = C, G or T; D = A, G or T; H = A, C or T; and V = A, C or G). The regions encoding the target sites were underlined.

RESULTS

siCD46 showed significant silencing effects on a majority of single-nucleotide mismatched targets

To study the effect of mismatches between siRNA and its mRNA target on RNAi silencing activity, we constructed fusion luciferase reporter plasmids carrying all possible single-nucleotide mutations of the target site for a functionally validated siRNA, siCD46 (10). This allowed us to effectively and systematically assess the impact of any single mutation in an easy read-out format, i.e. the dual-luciferase assay.

A total of 57 mutated constructs, with three different constructs for each position of the 19 wild-type nucleotides of the target site, were prepared and analyzed by the above-mentioned methodology (Figure 1A). This is a unique approach because, rather than modifying the siRNA molecule itself, we modify the target site, assuring that the siRNAs are not differently unwound or/and loaded into the RISC. The silencing efficiency of siCD46 was explored 24 h after co-transfection with fusion reporter plasmid and pRL-TK, the control plasmid. When a reporter with wild-type target site was used, the remaining luciferase activity from siRNA-treated samples were found to be ∼8% levels of the untreated samples, which corresponds to a silencing efficiency of 92% (Figure 1B).

When we analyzed the impact of the single-nucleotide mutations of the target site on the silencing efficiency, we found that 23 out of the 57 mutated constructs (40%) caused expression knockdown by >70% relative to the control, whereas 19 constructs (∼33%) reduced the luciferase signal by 40–70%. Only 15 constructs resulted in the expression silencing of <40% (Figure 1B). In other words, 73% of the 57 tested fusion reporters, each of which carries a mutated target site that forms a mismatch with its cognate siRNA, were significantly silenced by siCD46. Only a minority of single-nucleotide mismatches between the target site and siRNA can abolish the silencing activity of the siRNA. This result indicates that the off-target effects of a siRNA can be very broad, and the level of off-target effects variable.

In such a systematic mutational analysis, it is inevitable that stop codons are introduced in the transcripts. Therefore, we adapted the double ATG strategy, having an ATG immediately before and after the inserted target site. In our previous experiments, we have demonstrated that the second ATG could be used for the production of functional luciferase when an in-frame stop codon was inserted between the two ATG sites. This observation was now further confirmed in 2 of the 57 clones in which in-frame stop codons were introduced. For these clones, the expression of luciferase was in the normal range (data not shown), and the responses of these fusion constructs toward siCD46 siRNA were all in reasonable accordance with their counterparts that did not contain stop codons. These results indicated that the introduction of stop codons in the current experimental system did not have a profound effect on the assessment of siRNA silencing activity.

Mismatch tolerances of siCD46 are position-dependent

As shown in Figure 1B, siCD46 responded differently against mismatches introduced along the length of the target site. The siRNA tolerated mismatches without significant changes in the silencing activity. Based on the silencing efficacies of various mismatches caused by single-nucleotide variation, the target site of siCD46 can be divided into low, medium or high tolerance regions. The low tolerance region comprises position 5–11 of the target site. Mismatches at these positions abolish most of the suppression activities. This is consistent with the previous reports where perfect base pairing in the central region of the target site was found to be critical for the silencing activities, and siRNA is highly sensitive to mismatches in this region (1,5,6). However, a more detailed picture was presented here and our data extend the low tolerance region further toward the 5′ end of the target site, covering positions from 5 to 11. It should be stressed that even in this 'low tolerance region', some mutations were well tolerated. For example, certain nucleotide mutations at position 5, 7, 8 and 11 were found to be tolerated fairly well and the expression of the fusion gene was repressed >40%. These high tolerance mismatches in the low tolerance region could be very important for RNAi mechanism studies and its applications.

The medium tolerance region includes positions 3, 4 and 12–17, and represents mutations that are fairly well tolerated. Expression of fusion transcripts carrying substitutions at these positions was modestly, albeit significantly, repressed. Surprisingly, silencing efficiencies for mutations at position 13 (13-T) and position 14 (14-A) were similar to the wild-type sequence. This indicates that it is not only the position of mismatches, but also the identity of the substituted nucleotide that determines silencing efficacies. Finally, the high tolerance region consists of the two terminal nucleotide positions at both ends, being positions 1, 2, 18 and 19. None of the single-nucleotide substitution at these positions seemed to affect the silencing activities.

Mismatch tolerances of siCD46 are nucleotide-dependent

As mentioned above for the mismatched target sites 13-T and 14-A, mismatched nucleotides at particular positions affected silencing efficiency of siCD46 very differently (Figure 1C–F). For example, silencing efficiencies vary from 0 to 65% for

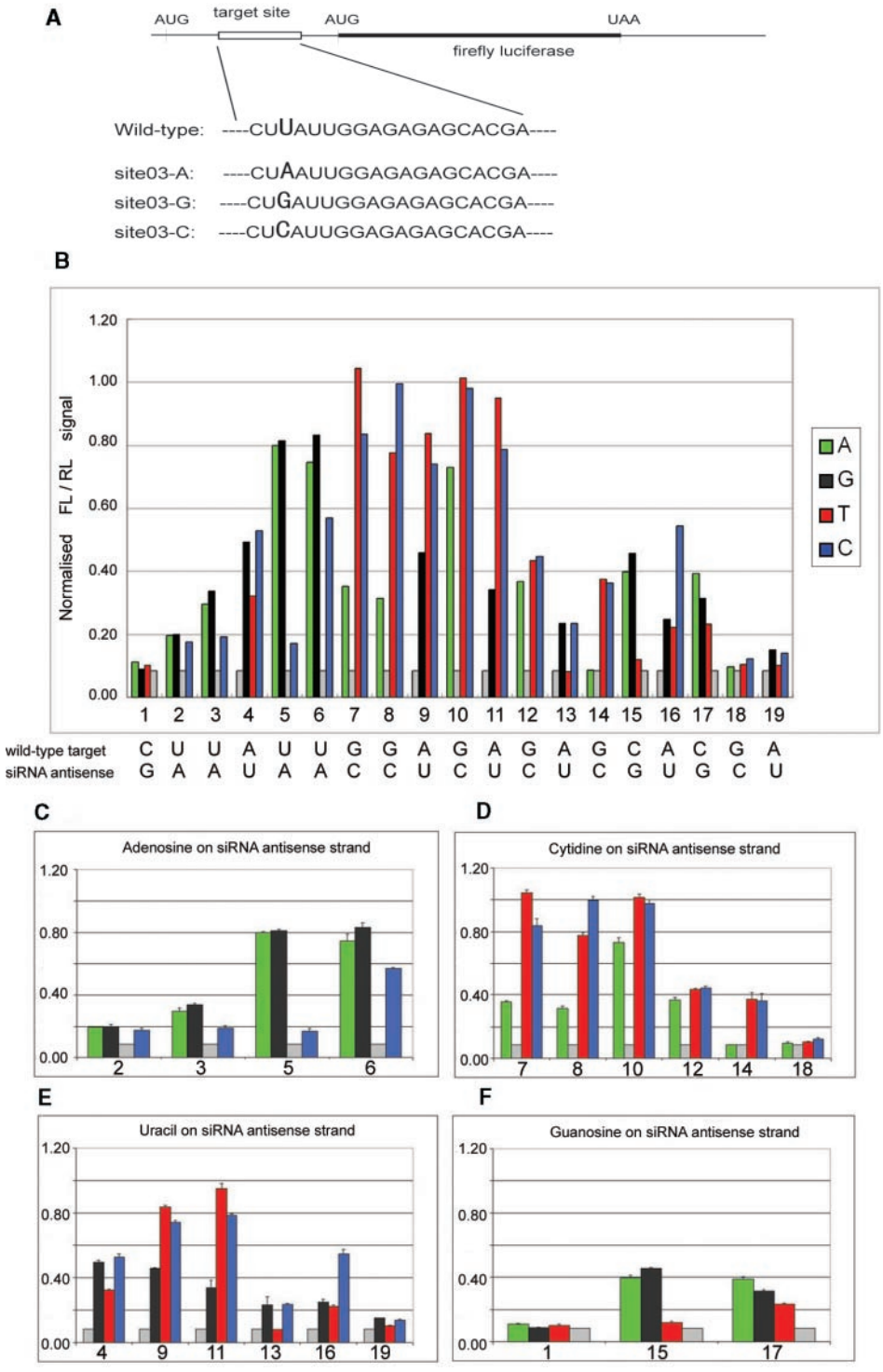

Figure 1. Systematic analysis of how single-nucleotide mismatches between siCD46 and its target sites affect silencing effect of siCD46. The results showed that such effects are dependent on both the position of the mismatches and the bases involved. (**A**) Schematic diagram of the fusion firefly luciferase reporter used. Shown to the left is the inserted in-frame AUG start codon, followed by a box representing the siRNA target site, and then as a thick line, the luciferase gene with authentic AUG codon. The target sequences mutated at position 3, as well as the wild-type sequence, were shown as examples. (**B**) Silencing efficiencies of siCD46 siRNA on mismatched fusion reporters plotted against the position of the mismatches (numbered from the start of the target site) and the nature of mismatched nucleotides. The target sequence and the guide strand (antisense) sequence of the siCD46 siRNA were given below the diagram. In (C–F), the silencing efficiencies of siCD46 were plotted in groups divided according to the base identity of the nucleotide on the siRNA antisense strand that the mutated nucleotides of the target sites paired with. (**C**) Mutated target sites paired with adenine on siRNA antisense strand. (**D**) Mutated target sites paired with cytosine on siRNA antisense strand. (**E**) Mutated target sites paired with uracil on siRNA antisense strand. (**F**) Mutated target sites paired with guanine on siRNA antisense strand. The wild-type matches were expressed as gray bars and the mutations were expressed as colored bars: green, mutated to adenine; black, mutated to guanine; red, mutated to uracil, and blue, mutated to cytosine. The numbers on *x*-axis represent the positions of 19mer target site numbered from the 5′ end. Results were average values of assays in triplicates and all experiments were repeated three times.

the mismatched target sites at position 7. This phenomenon reveals a correlation between the silencing efficiency and the nature of the mismatch for a given position. This was viewed in detail by examining the silencing profiles of siCD46 on mutated sites that introduce mismatches between the target site and all adenines on the siCD46 guide strand (Figure 1C). Surprisingly, cytosines were consistently the best-tolerated mismatch for the guide strand adenine. In agreement with this, a cytosine on the guide strand of the siRNA was found to tolerate mismatches with mutated adenine target sites in all the cases (Figure 1D).

In Figure 1E, the mutated target sites carrying either guanine or uracil, forming G:U or U:U when base pair with the uracil of the guide strand of the siRNA, were best tolerated. For the low tolerance region (positions 9 and 11), G:U mismatch was somewhat better tolerated than the U:U base pair. For the other sites, the differences between a U:U and a G:U base pair did not seem remarkable (Figure 1E). Finally, for guide-strand siRNA positions with guanines, uracils were the best-tolerated mismatch base. The G:U wobble base pair has previously been shown to be well tolerated for both RNA degradation and translation repression in antisense oligonucleotide–mRNA interaction, mRNA secondary structure formation and RNAi (11). Our observations here provide a quantitative view of how well the G:U wobble base pair behaves in comparison with a native G:C or A:T base pair in RNAi.

Other mismatches (A:A and A:G base pair) were much less tolerated. Comparable suppression efficacies were observed for A:A and A:G mismatch at the same position, but quite different repression efficacies were observed for the same mismatch at different positions.

In summary, in addition to confirming that G:U wobble base pairings are well tolerated for RNAi at most of the positions in the mRNA–siRNA pairing, we further identified that A:C mismatches are also very well tolerated in siRNA–mRNA interactions. Although effects on gene silencing efficiency were dependent on the positions of such base pairs, in most cases, both A:C or G:U wobble base pairings would allow the siCD46 to knockdown the reporter signals by >60%. This information, if can be generalized, might dramatically change our understanding of the spectrum of the off-target effects by siRNA.

Similar mismatch-tolerances are observed for siNPY305

To assess whether the mismatch tolerance profiles observed with siCD46 could be applicable to other siRNAs, we extended the investigation to include another siRNA. We chose a functionally validated siRNA against rat neuropeptide Y (siNPY305) and used the same experimental approach. The wild-type target site of siNPY305 and nine mutated target sites each carrying a single-nucleotide mutation were cloned into a fusion reporter plasmid (Figure 1A). The mutations were made on three positions of the siNPY305 target site: 8, 13 and 16. As a confirmation, the wild-type target site was efficiently silenced 95% by siNPY305. At position 8, the G:U wobble base pair was tolerated fairly well, whereas mismatches creating either U:U or C:U base pairing compromised the silencing activities (Figure 2). Mutations at positions 13 and 16

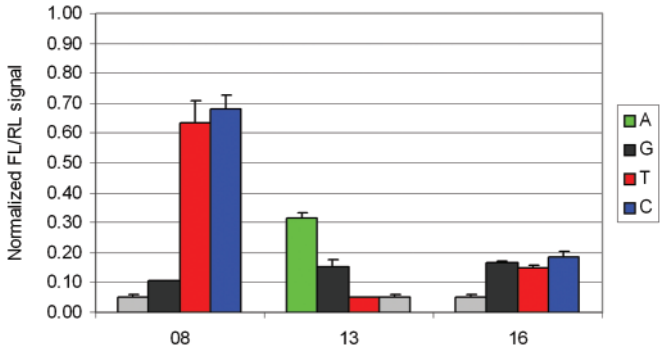

Figure 2. The effects of single-nucleotide mismatches between mutated target sites and siNPY305 on the silencing efficiencies. Fusion *firefly* luciferase constructs bearing single-nucleotide mismatched target sites at position 8, 13 and 16 were co-transfected into HEK293 cells with *renilla* luciferase expression vector, with or without siNPY305. Mutation sites and mutated bases were indicated on the horizon axis. The wild-type matches were expressed as gray bars and the mutations were expressed as colored bars: green, mutated to adenine; black, mutated to guanine; red, mutated to uracil, and blue, mutated to cytosine. Results were average values of assays in triplicates and all experiments were repeated three times.

were found to have only marginal effects on the silencing activity of siNPY305, regardless of the identity of the mismatch introduced, confirming that these positions do have high or medium tolerance for mismatches (Figure 2). Again, the most tolerated mismatch is a G:U wobble base pair. Overall, these data recapitulated the observed tolerance pattern using siCD46 and suggest that the observation might be generally applicable to other siRNAs.

Mismatched base pairing with the target causes mRNA degradation

Since some siRNAs might function as microRNA when imperfectly base paired with the mRNA targets, causing translation arrest rather than mRNA degradation (12–14), we performed northern blot analysis to investigate the nature of the silencing observed with siCD46 on mutated target sites. HEK293 cells were transfected with six different fusion reporter plasmids carrying the following representative mutations: 8-A, 11-G, 13-T, 13-G, 16-C and 16-T, with or without siCD46. The transfected fusion constructs were selected from both of the low and medium tolerance regions. Northern blots showed that the introduced siRNA resulted in a significant decrease in the mRNA level for all of the tested constructs (Figure 3). We thus concluded that the tested siRNA function as siRNA even when single-nucleotide mismatched target sites were introduced.

DISCUSSION

RNAi was initially considered to be a sequence-specific gene silencing mechanism that is dependent on a perfect match between the guide strand of the siRNA and the mRNA target sequence (1). It was, thus, further proposed that siRNA offers a way to silence disease genes in an allele-specific manner (15). Recently, however, specificity issues of RNAi have been raised by additional studies showing that various mismatches between the guide strand of the siRNA and the mRNA target

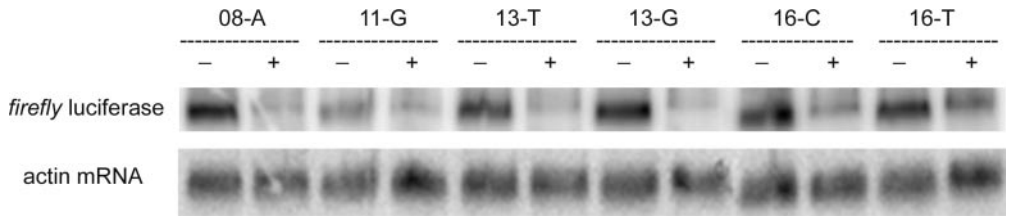

Figure 3. siCD46 induces degradation of mRNA harboring target sites with single-nucleotide mismatches. Reporter plasmids harboring mutated site 8-A, 11-G, 13-T, 13-G, 16-C and 16-T were co-transfected into cells with (+) or without (−) siCD46 and the reporter mRNA was evaluated by northern blot. The siRNA treatment markedly decreased the fusion gene mRNA level relative to the siRNA-untreated control. Bottom panel showed the expression of β-actin visualized by a β-actin cDNA probe as loading control.

site could be tolerated without abolishing the silencing effect of the siRNA (6,16,17). One common conclusion out of several such investigations indicated that central mismatches between the siRNA guide strand and the mRNA target are more critical than mismatches occurring toward either the 3′ or 5′ end. The studies differ from one another in assigning how mismatches located toward the 3′ or 5′ end would affect the silencing activity. We suspect that to some extent this discrepancy is owing to the fact that the majority of mismatch-tolerance studies of RNAi introduced mutated nucleotides into the siRNA rather than into the target mRNA. Since studies of RNAi mechanism revealed that the efficacy of siRNA could also be governed by how efficient the antisense strand of the siRNA enters into RISC complex, there is an obvious risk that alteration of siRNA sequences could interfere the partition of the antisense strand of the siRNA into RISC, thus affecting the silencing efficiency in a way that does not relate to target site recognition. The current method of introducing mutations into the target mRNA while maintaining the siRNA sequence unchanged allowed us to examine the target recognition step of RNAi without such potential complications. Of course, alteration of the target sequence may lead to the alteration of mRNA secondary structure, but the likelihood of this being a serious problem seems low in this case.

In addition to confirming the low tolerance of mismatches located centrally, we have been able to provide a detailed tolerance profiles for a siRNA on all possible single-nucleotide mismatched target sites. Although the results are from only two siRNAs and it is not certain that all of the findings can be generally applicable, the high levels of mismatch tolerance found at positions 1, 2, 18 and 19 of the target site and the significant level of tolerance at positions 3, 4, 12–17 should be of concern to those working with high-throughput, or long-term (stable) RNAi as well as with siRNA drug development efforts. What makes this issue particularly difficult is the unexpected tolerances with A:C or C:A mismatches at some positions, a phenomenon that was not observed in previous studies.

Apart from the standard Watson–Crick base pairings, a large collection of specific base–base interactions (non-canonical base pairings) have been predicted and frequently observed in crystal and NMR structures of RNA molecules (18). It is now widely accepted that such non-standard interactions can stabilize the secondary as well as the tertiary structures of RNA (19,20). A recent study demonstrated the tolerance for G:U wobble base pairing between siRNA oligonucleotides and its mRNA target (21), a finding that was confirmed in our study. In addition to the reported G:U wobble base pairing, we also found that A:C mismatched base pairs were well tolerated for siRNA-mediated gene silencing at positions, whereas other mismatches could affect the silencing activity most dramatically, e.g. at positions 5, 7 and 8 of the target site. Currently, it is unclear whether the A:C mismatch tolerance is only applicable to RNAi process or it is also applicable to the RNA secondary structure formation.

ACKNOWLEDGEMENTS

The authors thank Dr Liam Good and Ola Larsson for comments on the manuscript. This work was supported by a grant from WCN to Z.L. and C.W. Funding to pay the Open Access publication charges for this article was provided by a grant from WCN to Z.L. and C.W.

Conflict of interest statement. None declared.

REFERENCES

1. Elbashir,S.M., Martinez,J., Patkaniowska,A., Lendeckel,W. and Tuschl,T. (2001) Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J*., **20**, 6877–6888.
2. Bass,B.L. (2000) Double-stranded RNA as a template for gene silencing. *Cell*, **101**, 235–238.
3. Semizarov,D., Frost,L., Sarthy,A., Kroeger,P., Halbert,D.N. and Fesik,S.W. (2003) Specificity of short interfering RNA determined through gene expression signatures. *Proc. Natl Acad. Sci. USA*, **100**, 6347–6352.
4. Chi,J.T., Chang,H.Y., Wang,N.N., Chang,D.S., Dunphy,N. and Brown,P.O. (2003) Genomewide view of gene silencing by small interfering RNAs. *Proc. Natl Acad. Sci. USA*, **100**, 6343–6346.
5. Jackson,A.L., Bartz,S.R., Schelter,J., Kobayashi,S.V., Burchard,J., Mao,M., Li,B., Cavet,G. and Linsley,P.S. (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nat. Biotechnol*., **21**, 635–637.
6. Amarzguioui,M., Holen,T., Babaie,E. and Prydz,H. (2003) Tolerance for mutations and chemical modifications in a siRNA. *Nucleic Acids Res*., **31**, 589–595.
7. Schwarz,D.S., Hutvagner,G., Du,T., Xu,Z., Aronin,N. and Zamore,P.D. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, **115**, 199–208.
8. Du,Q., Thonberg,H., Zhang,H.Y., Wahlestedt,C. and Liang,Z. (2004) Validating siRNA using a reporter made from synthetic DNA oligonucleotides. *Biochem. Biophys. Res. Commun*., **325**, 243–249.
9. Thonberg,H., Zhang,S.J., Tvrdik,P., Jacobsson,A. and Nedergaard,J. (1994) Norepinephrine utilizes alpha 1- and beta-adrenoreceptors synergistically to maximally induce c-fos expression in brown adipocytes. *J. Biol. Chem*., **269**, 33179–33186.
10. Xu,Y., Zhang,H.Y., Thormeyer,D., Larsson,O., Du,Q., Elmen,J., Wahlestedt,C. and Liang,Z. (2003) Effective small interfering RNAs and

phosphorothioate antisense DNAs have different preferences for target sites in the luciferase mRNAs. *Biochem. Biophys. Res. Commun.*, **306**, 712–717.
11. Saxena,S., Jonsson,Z.O. and Dutta,A. (2003) Small RNAs with imperfect match to endogenous mRNA repress translation. Implications for off-target activity of small inhibitory RNA in mammalian cells. *J. Biol. Chem.*, **278**, 44312–44319.
12. Doench,J.G., Petersen,C.P. and Sharp,P.A. (2003) siRNAs can function as miRNAs. *Genes Dev.*, **17**, 438–442.
13. He,Z. and Sontheimer,E.J. (2004) ''siRNAs and miRNAs'': a meeting report on RNA silencing. *RNA*, **10**, 1165–1173.
14. Olsen,P.H. and Ambros,V. (1999) The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev. Biol.*, **216**, 671–680.
15. Abdelgany,A., Wood,M. and Beeson,D. (2003) Allele-specific silencing of a pathogenic mutant acetylcholine receptor subunit by RNA interference. *Hum. Mol. Genet.*, **12**, 2637–2644.
16. Boutla,A., Delidakis,C., Livadaras,I., Tsagris,M. and Tabler,M. (2001) Short 5′-phosphorylated double-stranded RNAs induce RNA interference in *Drosophila*. *Curr. Biol.*, **11**, 1776–1780.
17. Holen,T., Amarzguioui,M., Wiiger,M.T., Babaie,E. and Prydz,H. (2002) Positional effects of short interfering RNAs targeting the human coagulation trigger tissue factor. *Nucleic Acids Res.*, **30**, 1757–1766.
18. Nagaswamy,U., Voss,N., Zhang,Z. and Fox,G.E. (2000) Database of non-canonical base pairs found in known RNA structures. *Nucleic Acids Res.*, **28**, 375–376.
19. Hermann,T. and Patel,D.J. (1999) Stitching together RNA tertiary architectures. *J. Mol. Biol.*, **294**, 829–849.
20. Nagaswamy,U., Larios-Sanz,M., Hury,J., Collins,S., Zhang,Z., Zhao,Q. and Fox,G.E. (2002) NCIR: a database of non-canonical interactions in known RNA structures. *Nucleic Acids Res.*, **30**, 395–397.
21. Amarzguioui,M., Brede,G., Babaie,E., Grotli,M., Sproat,B. and Prydz,H. (2000) Secondary structure prediction and *in vitro* accessibility of mRNA as tools in the selection of target sites for ribozymes. *Nucleic Acids Res.*, **28**, 4113–4124.