7560–7569 Nucleic Acids Research, 2009, Vol. 37, No. 22 Published online 8 October 2009
doi:10.1093/nar/gkp835

Profiling of mismatch discrimination in RNAi enabled rational design of allele-specific siRNAs

Huang Huang[1], Renping Qiao[2], Deyao Zhao[3], Tong Zhang[2], Youxian Li[1], Fan Yi[1], Fangfang Lai[1], Junmei Hong[1], Xianfeng Ding[3], Zhenjun Yang[2], Lihe Zhang[2], Quan Du[1,*] and Zicai Liang[1,*]

[1]Institute of Molecular Medicine, Peking University, Beijing 100871, [2]The State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Science, Peking University, Beijing 100083 and [3]School of Life Science, Zhejiang Sci-Tech University, Hangzhou 310018, China

Received August 27, 2009; Revised and Accepted September 18, 2009

ABSTRACT

Silencing specificity is a critical issue in the therapeutic applications of siRNA, particularly in the treatment of single nucleotide polymorphism (SNP) diseases where discrimination against single nucleotide variation is demanded. However, no generally applicable guidelines are available for the design of such allele-specific siRNAs. In this paper, the issue was approached by using a reporter-based assay. With a panel of 20 siRNAs and 240 variously mismatched target reporters, we first demonstrated that the mismatches were discriminated in a position-dependent order, which was however independent of their sequence contexts using position 4th, 12th and 17th as examples. A general model was further built for mismatch discrimination at all positions using 230 additional reporter constructs specifically designed to contain mismatches distributed evenly along the target regions of different siRNAs. This model was successfully employed to design allele-specific siRNAs targeting disease-causing mutations of PIK3CA gene at two SNP sites. Furthermore, conformational distortion of siRNA-target duplex was observed to correlate with the compromise of gene silencing. In summary, these findings could dramatically simplify the design of allele-specific siRNAs and might also provide guide to increase the specificity of therapeutic siRNAs.

INTRODUCTION

RNAi is an evolutionarily conserved process where small interfering RNA (siRNA) specifically represses the expression of target genes (1,2). siRNAs are widely expected to become next generation of biological therapeutics (3,4), and they are initially anticipated to play a major role in treatment of diseases involving single nucleotide polymorphisms (SNPs) where discrimination against single nucleotide variation between wild-type and mutant alleles is demanded (5–8). This dream was then hammered by subsequent reports demonstrating that siRNAs could incur widespread knockdown of unrelated genes, a phenomenon known as 'off-target effects' (9–12). Closer scrutiny of 'off-target effects' of siRNA has however revealed the Janus-like view of siRNA in term of target specificity. On one hand, siRNA does induce weak down-regulations on sites that are apparently only related to the siRNA by matching to the seed region (13). On the other hand, siRNA could discriminate some very closely related target sites with only two or even one nucleotide mismatches (14,15). Nobody knows how siRNA or the RISC complex could harmonize the two apparently conflicting properties of siRNAs.

Due to utmost importance of siRNA specificity in siRNA drug development and the need for generating allele-specific siRNAs, we have created an experimental system to generate insight of siRNA discrimination of mismatched target sites. Using more than 400 reporter plasmids for 20 siRNAs we have revealed a general rule for mismatch tolerance and discrimination. Using this discovery as a guideline, we have constructed the first model of mismatch-tagged, position-specific discrimination of closely related target sites. The model was then successfully used for designing allele-specific siRNA.

MATERIALS AND METHODS

Oligonucleotides and plasmids

DNA oligonucleotides were from Invitrogen (Beijing, China). RNA oligonucleotides were from Genepharma

*To whom correspondence should be addressed. Tel: +86 10 62769861; Fax: +86 10 62769862; Email: quan.du@pku.edu.cn
Correspondence may also be addressed to Zicai Liang. Tel: +86 10 62769862; Fax: +86 10 62769862; Email: liangz@edu.pku.cn

(Shanghai, China) and Proligo Sigma (Paris, France). Plasmid DNAs were extracted using a mini-purification kit (Promega).

RNAi assay

Human embryonic kidney (HEK293) cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 100 U/ml penicillin and 100 μg/ml streptomycin (Life Technologies, Gibco). The cells were seeded into 24-well plates at a density of $\sim 1 \times 10^5$ cells/well one day before transfection. siQuant vector (0.17 μg/well) carrying the target site of tested siRNA was transfected into HEK293 cells at approximately 50% confluence, together with pRL-TK control vector (0.017 μg/well), with or without the siRNA (13 nM). The activity of both luciferases was determined by a fluorometer (Synergy HT, BioTek, USA) before the *firefly* luciferase activity was normalized to *renilla* luciferase for each well. Silencing efficacy of each siRNA was calculated by comparison with a sample without siRNA treatment. All experiments were performed in triplicate and repeated at least twice.

Northern blot assay

Twenty-four hours after transfection with siQuant vector and siRNA, total RNA was harvested from HEK293 cells with Trizol reagent (Invitrogen) according to the manufacturer's instructions. Total RNA (10–20 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 DNA probe was labeled with biotin-dUTP using the Prime-a-Gene labeling system (Promega). Hybridization and stringent washing were performed according to ExpressHy (Clontech), and the signals were detected by Streptavidin IRDye 800CW on an Odyssey infrared imaging system (LI-COR).

Spectroscopy

UV-Vis melting curves (absorbance versus temperature profile) of RNA duplexes were collected by a Varian Cary 300 Bio UV-visible spectrophotometer (Varian, USA) with a cell of 1.0 cm path length. 0.2 OD RNA duplex was used for each measurement. Prior to the measurements, RNA duplexes were denatured and annealed by heating to 90°C for 5 min and slowly cooling to 4°C overnight. The absorbance was measured at 260 nm from 20 to 90°C with a heating rate of 0.5°C per minute in RNase-free buffer (10 mM phosphate buffer, 100 mM NaCl and 0.1 mM EDTA, pH 7). Melting temperature was calculated using the Carry WinUV software package.

Circular dichroism spectra

The circular dichroism (CD) spectrums were measured from 320 to 200 nm at 25°C using a JASCO J-810 spectropolarimeter (AVIV, USA), with a 0.1 cm path length cuvette. 0.2 OD RNA duplex in RNase-free buffer (10 mM phosphate buffer, 100 mM NaCl and 0.1 mM EDTA, pH 7) was used in each measurement. For each RNA duplex, three independent measurements were made.

Computational modeling of the secondary structure of RNA duplexes

Molecular dynamics (MD) simulation was performed using the Sander module of the AMBER8 program. The RNA molecule was neutralized by Na^+ counter ions at 37°C (placed according to xLEaP at favorable electrostatic potential points), surrounded by an 8 Å TIP3P periodic water box. Optimization and MD simulation were performed according to the following procedure: (i) 1000 steps of initial energy minimization with solvate and ions, holding RNA structure fixed; (ii) 2500 steps of restrained energy minimization (no restraints for the minimization of the wild-type RNA structure); *cis*-Watson–Crick base pairing hydrogen bonds were used as constrains for the optimization; (iii) 20 ps molecular dynamics simulation by heating the whole system from 0 to 300°K, while keeping the RNA molecule fixed with weak restraints; and (iv) molecular dynamics equilibration with a time step of 2 fs at 300°K until the RMSD curves became stable.

RESULTS

Position-dependent mismatch discrimination pattern of siRNA

Using a previously established gene silencing assay (16), a siRNA target fragment, perfectly matched or single nucleotide mismatched relative to the siRNA guide strand, was fused in-frame with the *firefly* luciferase gene in a mammalian expression vector. The resulting fusion reporter was then used as an artificial target to evaluate the silencing efficacy of the siRNA on the target. In the present study, target mutations were initially made at three positions, selected from target regions of differing mutation sensitivity (9,15). In total, twenty effective siRNAs were examined for their silencing efficacy on 200 matched or single nucleotide mismatched targets, covering all such mismatches at these positions (Supplementary Data S1). The mismatch formed between the siRNA guide strand and the mutant target is indicated by X:Y, where X stands for the nucleotide in the guide strand and Y stands for the opposite nucleotide in the target.

Surprisingly, when the silencing efficacies were aligned in terms of the position and the identity of the mismatched nucleotide in the siRNA guide strand, consistent discrimination profiles were found in each subgroup (Figure 1, measurement variances in Supplementary Data S2). For example, in the case of four siRNAs having adenine at position 17 of the guide strand (Figure 1, upper left panel), mutation of the target nucleotide from U to C led to an A:C mismatch between the guide strand and the target. Compared with other mutations, mutation from U to C caused the least loss of silencing efficacy for all four siRNAs, while mutation from U to G

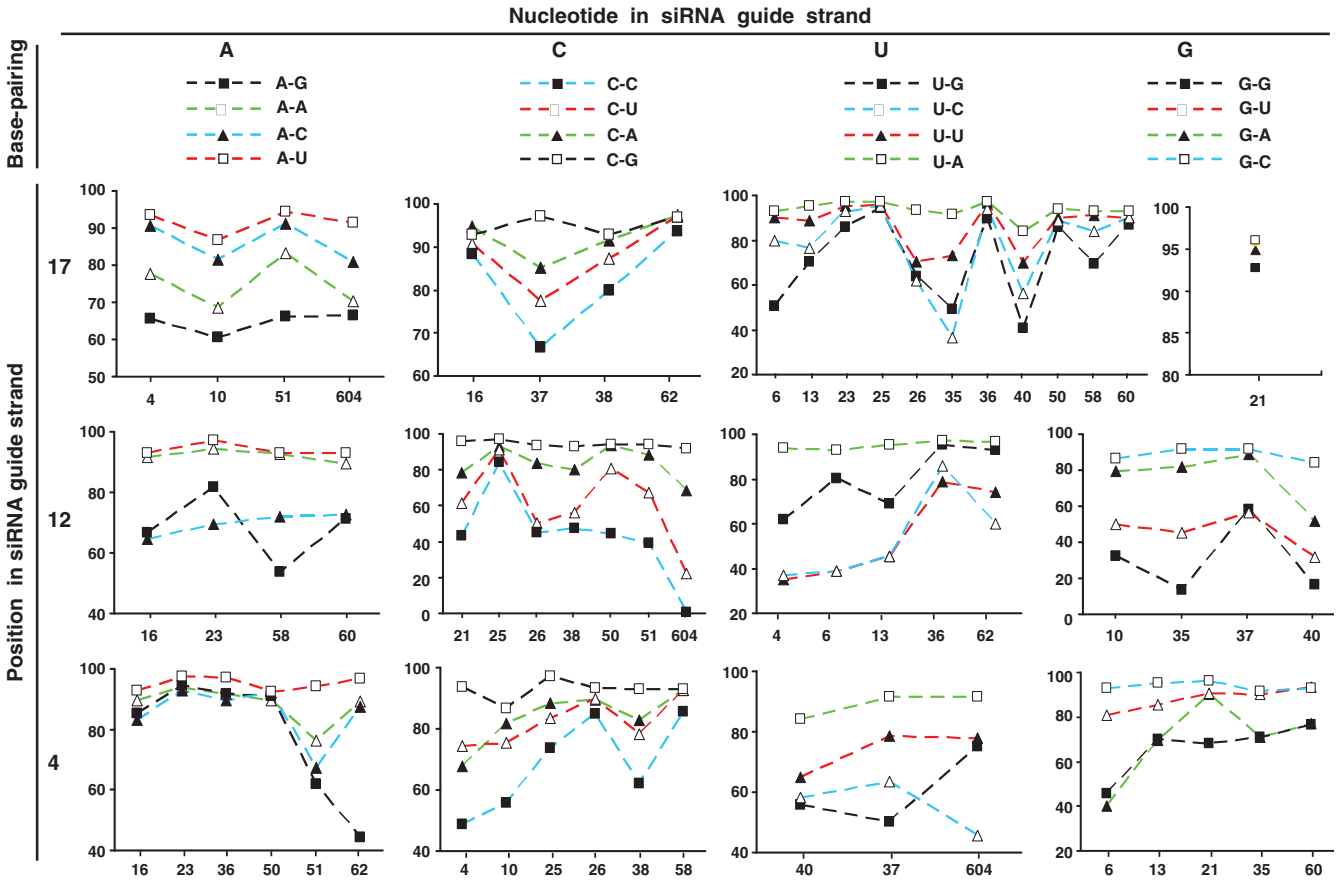

Figure 1. Mismatch discrimination pattern at three positions. Gene silencing tolerance of single point target mutations was studied by examining the silencing efficacies of 20 siRNAs on their perfectly matched or single nucleotide mismatched targets at three guide positions—4, 12 and 17. The silencing efficacies were aligned in terms of position and then the identity of the mismatched nucleotide in the siRNA guide strand. Horizontal axis, siRNA; vertical axis, silencing efficacy displayed as percentage.

(forming an A:G mismatch) resulted in the most significant loss of silencing activity. Therefore, for siRNAs having adenine at this position, the tolerance order was C > A > G; for siRNAs with cytosine at this position, the order was A > U > C; and for siRNAs with uracil, the order was U > C > G.

When mutations were made on the target nucleotide opposite to position 12 of the guide strand, the tolerance order was A > C = G for adenine, A > U > C for cytosine, G > C = U for uracil, and A > U > G for guanine in the guide strand. Unlike position 17, which is located in a region of low sensitivity, position 12 is in a region of high sensitivity approximating the cleavage site. Accordingly, relatively low tolerance was found at this position.

Position 4 was selected from the seed region of the siRNA guide strand; this position is implicated in off-target gene silencing mediated by translation repression (9,10). When mutations were made here, the tolerance order was A > C = G for A in the guide strand, A > U > C for C in the guide strand, U > G = C for U in the guide strand, and U > A = G for G in the guide strand.

From the layered gene silencing efficacy at these positions, several revealing observations were made. First, a unique tolerance profile was identified at each tested position, demonstrating significant difference in the gene silencing of the most tolerant and the less tolerant target mutations (Supplementary Data S3). For example at position 12, the A:A mismatch had virtually no effect on the silencing efficacy of all five siRNAs having adenine, however, the same mismatch significantly compromised gene silencing of the four siRNAs with adenine at position 17 (15% reduction on average). Second, dramatically different tolerance levels occurred if mismatched nucleotides were swapped between the guide strand and the target. For example at position 17, both A:C and C:A mismatches were well tolerated and mediated strong gene silencing; equivalent tolerance did not occur at position 12. While an A:C mismatch at position 12 was largely tolerated (11% reduction on average), silencing efficacy of the siRNAs was significantly compromised by a C:A mismatch (25% reduction on average). In addition, tolerance of the U:G wobble base pair was not as great as that of the G:U wobble base pair in general. Third, the order of position-specific tolerance was independent of the sequence context of the siRNA or the target.

To rule out potential bias due to the cell type used in the assay, the experiments were repeated with Hela cells,

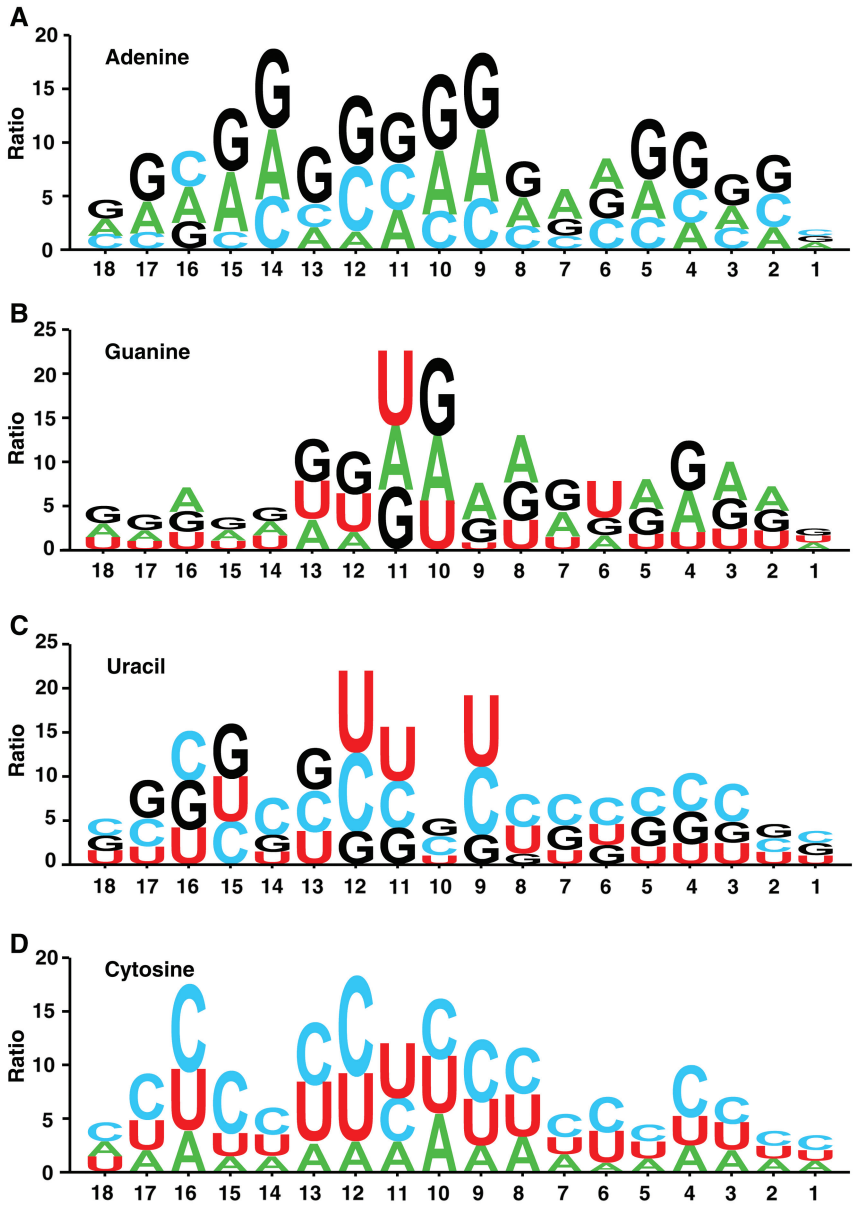

Figure 2. The mismatch discrimination profile across the whole target positions. A general discrimination profile was presented in logo according to the nucleotide in siRNA guide strand: (**A**) adenine; (**B**) guanine; (**C**) uracil; (**D**) cytosine. Mismatches formed between the guide strand and single-nucleotide mismatched targets were represented by the symbols of the mismatched target nucleotide. The height of the symbol was made proportional to the average ratio that was calculated by normalizing the remaining luciferase activity of mismatched target to perfectly matched counterpart, indicating the extent of interference with gene silencing. Larger size of the symbol indicates greater interference and less tolerance of the mismatched target. At each position, the symbols were sorted in terms of height from bottom to top, and the overall height of the symbols indicates the discrimination property of the position.

revealing similar tolerance profiles. Together, these data demonstrated that the tolerance profiles are a common property of RNAi processing.

Building a general model of single nucleotide mismatch discrimination profile of siRNA

Taking into consideration of its potential significance in allele-specific siRNA design, other target positions were further examined. For each given identity of the nucleotide on a target position, at least two siRNAs were examined using the same experimental setting. The resulting gene silencing data were aligned accordingly (Figure 2, Supplementary Data S4). As expected, a unique discrimination profile was found at individual positions across the whole target site, demonstrating the generality of the mismatch discrimination in RNA interference.

To the best of our knowledge, this is the most comprehensive data set regarding single point target mutations to date. The most striking finding was that, depending on the nucleotide in the siRNA guide strand, a most tolerant and a least tolerant mismatch, as well as a few most sensitive

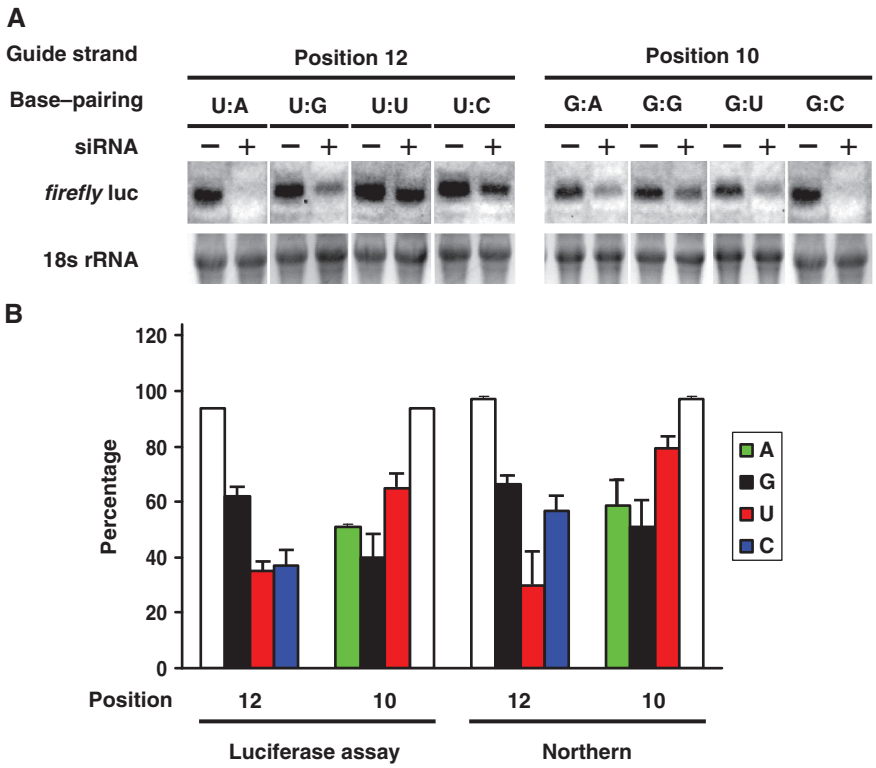

Figure 3. Silencing of single point mismatched target is mediated by target cleavage. Northern blot was performed with mRNA transcripts harboring perfectly matched or single nucleotide mismatched target sites of siRNA-4. Reporter plasmids mutated at positions 12 and 10 were co-transfected into cells with (+) or without (−) the siRNA, and then the amount of reporter mRNA was evaluated using fluorophore-labeled probe. (**A**) siRNA treatment markedly decreased the fusion gene mRNA level relative to the control without siRNA. 18s rRNA was included as loading control. (**B**) The amount of mRNA transcript was quantified and compared with the gene silencing efficacy. Both the northern and luciferase assay were normalized to perfectly matched target.

positions occurred. Given A in the guide strand, the most tolerant mismatch was the A:C (purine:pyrimidine) mismatch that resulted in 69% reduction of reporter gene activity on average, and the least tolerant mismatch was A:G (purine:purine, 52% reduction on average). The most sensitive positions were 9, 10, 12, 14 and 15 (Figure 2A). For G in the guide strand, the most tolerant mismatch was G:U (purine:pyrimidine, 72% reduction) and the least tolerant was G:G (purine:purine, 64% reduction); the most sensitive positions were 8–13 (Figure 2B). For U in the guide strand, the most tolerant mismatch was U:G (purine:pyrimidine, 62% reduction) or U:U (pyrimidine:pyrimidine, 62% reduction) and the least tolerant was U:C (pyrimidine:pyrimidine, 49% reduction); the most sensitive positions were 9, 11–13 and 15–16 (Figure 2C). For C in the guide strand, the most tolerant mismatch was C : A (purine : pyrimidine, 80% reduction) and the least tolerant was C : C (pyrimidine:pyrimidine, 62% reduction); the most sensitive positions were 10–13 and 16 (Figure 2D). In general, purine:pyrimidine mismatches were more tolerated than purine:purine or pyrimidine:pyrimidine mismatches. Other interesting observations included: (i) distinct sensitivity profiles were found depending on the nucleotide in the siRNA guide strand. The mismatches formed by G were generally more tolerated than mismatches formed by A in the guide strand (Figure 2A and B). (ii) In contrast to previous suggestions that perfectly complementary around the cleavage site is required in RNAi (17–20), our data showed that these were not always the most sensitive positions. For example, mismatches at position 10 were much more tolerated than at positions 9 and 11 (Figure 2C). (iii) Generally low sensitivity was found in the seed region of the siRNA guide strand.

To investigate the nature of the silencing in our study, northern blot assay was performed. The results showed that the amount of mutated target transcript was significantly reduced by the introduction of the siRNA, indicating that the gene silencing was mediated by target cleavage rather than translational repression (Figure 3A). Furthermore, mRNA quantification by northern blot perfectly matched the gene silencing efficacy measured by luciferase activity (Figure 3B).

Rational design of allele-specific siRNA using the mismatch discrimination model as a guide

The remarkable discrimination profile revealed in our studies indicated that, for a given mismatch, some positions are more sensitive (less tolerant) than others. If a mismatch was placed at its most sensitive position, selective gene silencing between the perfectly matched and the single nucleotide mismatched target could be achieved.

Table 1. Recommended positions for allele-specific siRNA design

Mutation type	Nucleotide on siRNA	Discriminate (least tolerant)	Positions in guide (strand)
A→C	G	A	8, 9, 10, 11, 16
A→G	C	A	10, 11, 12, 16
A→T	A	A	9, 10, 14, 15, 16
C→A	U	C	11, 12, 13, 14, 16
C→G	C	C	9, 10, 12, 13, 16
C→T	A	C	9, 10, 11, 12, 14
G→A	U	G	10, 13, 15, 16
G→C	G	G	10, 11, 12, 13
G→T	A	G	9, 10, 11, 12, 14
T→A	U	U	9, 11, 12, 15
T→C	G	U	10, 11, 12, 13
T→G	C	U	10, 11, 12, 13, 16

This discrimination is of great value to the silencing of SNP disease alleles, in which a siRNA that is designed to perfectly match the mutant allele will also pair imperfectly with the wild-type allele, and form a single nucleotide mismatch between them. By placing the mismatch in its most sensitive positions (Table 1), the siRNA might mediate effective silencing of the mutant allele, while off-target silencing of the wild-type allele is minimized.

To establish proof of concept, single nucleotide mutated alleles of the PIK3CA (phosphoinositide-3-kinase, catalytic, alpha polypeptide) gene were further studied. PIK3CA encodes the p110α catalytic subunit that regulates essential signaling pathways important for neoplasia (21,22). This gene is frequently mutated in human tumors in the form of a single point mutation, thus providing us with a model system to evaluate the selective gene silencing strategy (23). The two most frequent SNP sites derived from human tumor samples were selected for study: one causes a G to A transition (mutation 1633), and the other an A to G transition (mutation 3140). Based on our earlier data, four allele-specific siRNAs were designed to target mutation 1633 and three were designed to target mutation 3140, placing the mismatched nucleotide at their most sensitive positions in the siRNA guide strand (Figure 4A). To evaluate the selective gene silencing, fusion reporter genes carrying wild-type or mutant target were made as described above. Cell assays successfully identified three allele-specific siRNAs against mutation 1633 (1633-10, 13, 16) and one against mutation 3140 (3140-11) (Figure 4B and C). While pronounced gene silencing occurred with perfectly matched mutant alleles, silencing of mismatched wild-type alleles was insignificant, displaying high selectivity in allele-specific gene silencing. The discrimination was further observed over siRNA concentrations from 0.02 to 80 nM, indicating that it is not a concentration-dependent phenomenon (Figure 4D and E).

Duplex conformation change is correlated with mismatch discrimination

In light of a recent study showing that a mismatch between the siRNA guide strand and the target RNA strand causes marked structural change and distorts its helix conformation (24–26), we quantitatively investigated two crucial features of the siRNA-target duplex, helix conformation and thermodynamic stability. Studies were performed with duplex RNAs formed by the guide strand of siRNA-604 and its matched, or single nucleotide mismatched targets, at position 6, 9, 10, 12, 15 or 17 (Supplementary Data S5). At these positions, the perfectly matched duplexes, the most tolerant duplexes (A:C or C:A), as well as the least tolerant duplexes (A:A or C:C) were examined.

To quantify the conformational distortion caused by the mismatches, an experimental CD spectrum assay and computing-based MD simulations were performed. Consistent with previous studies (25,26), a distorted helix conformation was observed in mismatched duplexes when compared with perfectly matched counterparts (Figure 5B). More importantly, matched patterns were revealed between the silencing efficacies and the deviations in negative Cotton values of the CD spectra exhibiting conformational distortion of the helix structure (Figure 5A and B). This finding was further corroborated by MD simulations (Figure 6), suggesting a correlation between the compromised gene silencing and the degree of conformational distortion. Taken together, our data indicated that greater conformational distortion caused more interference with gene silencing activity, thus providing an explanation of why some target mutations are more tolerated than the others.

Besides conformational change, the influence on the thermodynamic stability of the RNA duplexes was also examined in terms of melting temperature. It was interesting to note that changes in melting temperature correlated with the positions but not the identities of the mismatches (Figure 5A and C). For example, at position 15, A : A and A : C mismatches had little effect on melting temperature compared with the wild-type counterpart. The silencing efficacy of the A : A mismatch was, however, decreased by a significant 72%, while the silencing of the A : C mismatch was decreased by only 9%. The apparent inconsistency between the changes of melting temperature and gene silencing efficacies indicated that the perturbed thermodynamic stability of the duplex does not account for the compromised gene silencing.

DISCUSSION

In allele-specific gene therapy, specifically designed siRNAs are expected to selectively reduce the expression of disease-causing alleles without affecting the expression of their wild-type counterparts. To achieve this goal, we performed a large-scale RNAi assay to understand the role of target mutation in off-target gene silencing. The results demonstrated a position-specific discrimination profile across the whole target site, and in particular, identified the most tolerant and the most disruptive single point target mutation at each target position. Furthermore, for any given mismatch between siRNA guide strand and target RNA, the study identified the most sensitive positions for a single point target

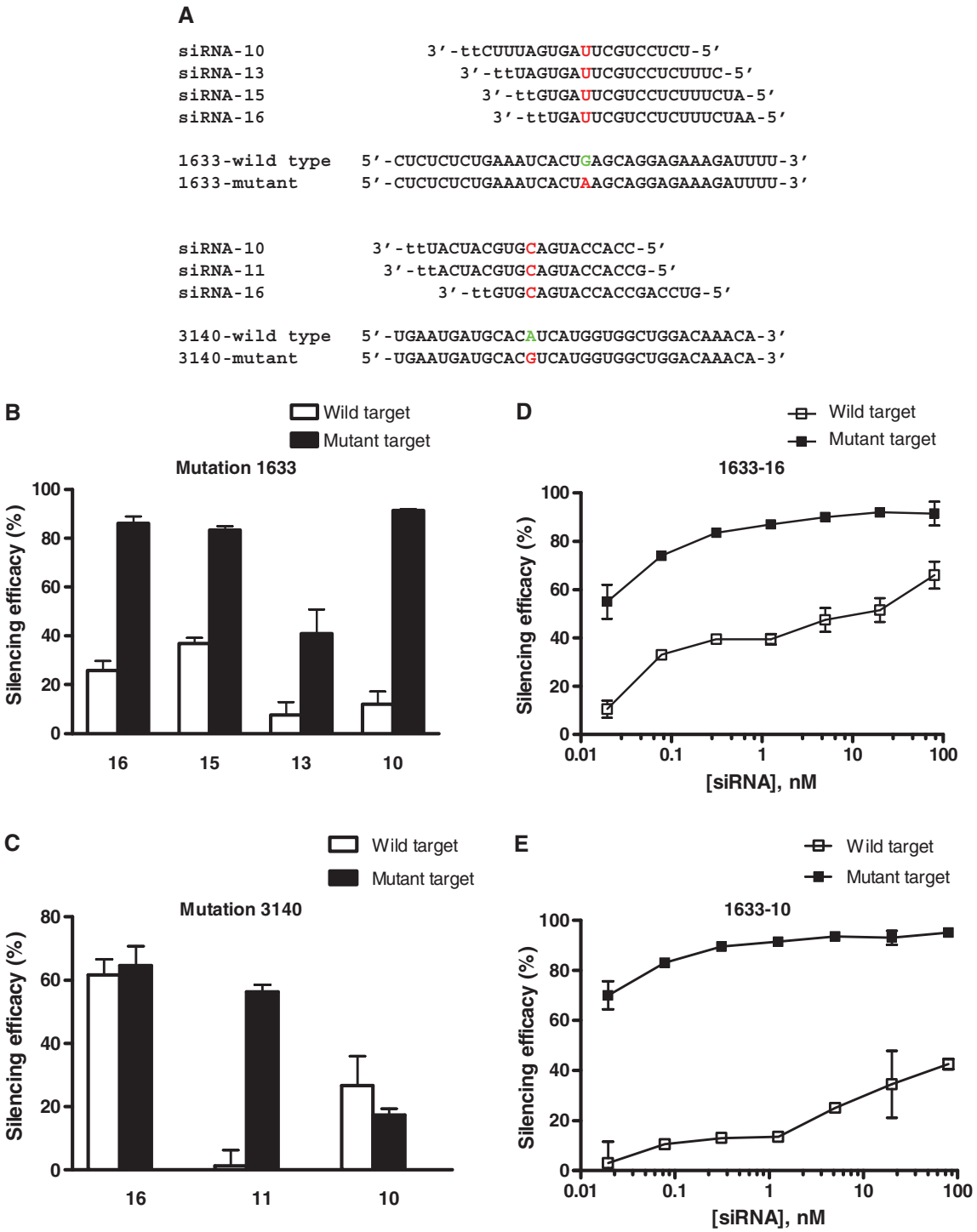

Figure 4. Potent and selective silencing of SNP disease alleles. (**A**) Four siRNAs were designed to target the disease-causing allele at position 1633 of the PIK3CA gene, and three siRNAs were designed to target the disease allele at position 3140. (**B, C**) RNAi assays performed with fusion luciferase reporter having wild-type or target mutation at position 1633 and 3140. (**D, E**) RNAi assays performed with siRNA 1633-10, 16 over a series of siRNA concentrations.

mutation, therefore establishing a solid basis for allele-specific siRNA design.

A consistent tolerance preference was observed for the same mismatch when it was at the same position in the target site, indicating that processing of similarly mismatched substrate duplexes at the same location within the Ago protein likely results in comparable interference with its functionality. In contrast, when the same mismatch was at different target positions, distinct tolerance preferences were found. This indicated that processing of similarly mismatched duplexes at different locations within the protein resulted in distinct levels of interference

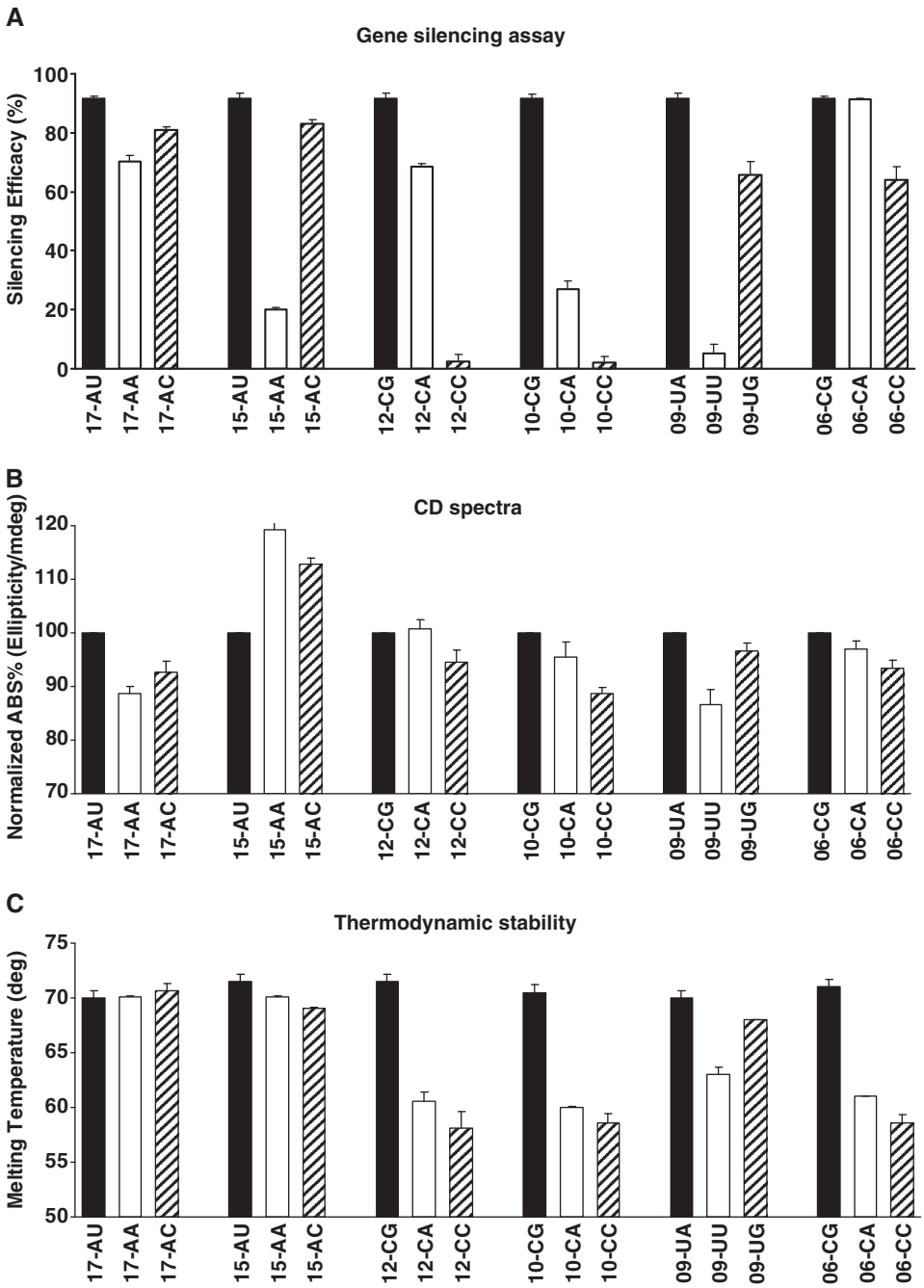

Figure 5. Correlation between gene silencing efficacy and CD spectra or thermodynamic stability**.** Mismatch-derived alteration of conformational distortion or duplex thermodynamic stability was examined using RNA duplexes formed by the guide strand of siRNA-604 and matched or single nucleotide mismatched targets. Conformational distortion of the RNA duplex was evaluated by measuring CD spectra (**B**), and the thermodynamic stability was evaluated by melting temperature (**C**). For comparison, the gene silencing efficacy of the respective target is shown in (A). Mismatched position in siRNA guide strand and base-pairing type are indicated on the horizontal axis.

with its functionality. In search of the mechanism involved in the discriminative gene silencing against differently mutated targets, quantitative assays were performed on siRNA-target duplexes, with the focus on the influence on duplex conformation and thermodynamic stability. In agreement with previous studies (25,26), our data showed that mismatch between siRNA guide strand and target RNA caused a marked change in the normal helix structure. Furthermore, our data indicated that the tolerant mismatches caused less structural distortion than the disruptive mismatches, which explains why some mutations were more tolerated than others. Collectively, these lines of evidence led us to speculate that, depending on its position and identity, a mismatch formed between the siRNA guide strand and the target RNA might introduce a conformational change in the

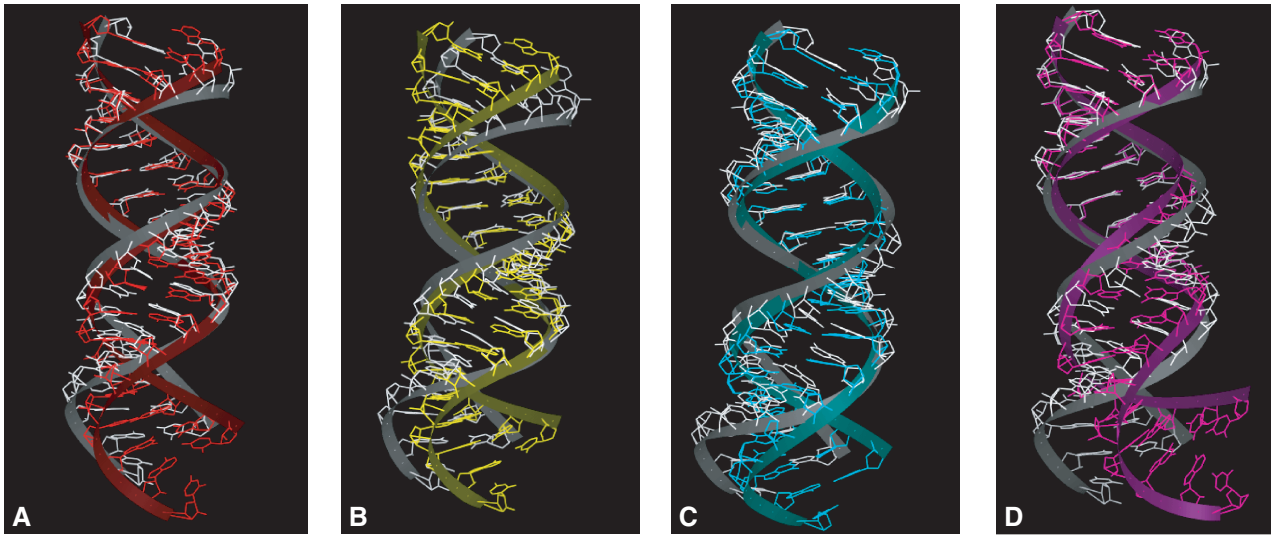

Figure 6. Distorted RNA structures revealed in MD simulations. Predicted structures of RNA duplexes formed by the guide strand of siRNA-604 and matched or single nucleotide mismatched targets. In each panel, structures of perfectly matched duplex (white) and a mismatched duplex (another color) are shown in parallel. RMSD values indicate the structural distortion in mismatched RNA duplexes. (**A**) Most tolerant A:C mismatch at position 15; (**B**) least tolerant A:A mismatch at position 15; (**C**) most tolerant C:A mismatch at position 12; (**D**) least tolerant C:C mismatch at position 12.

siRNA-target duplex, which in turn interferes with the functionality of the Ago protein in off-target gene silencing. The influence on Ago function is determined by the level of the conformational change caused by the target mutation.

As a major potential application of therapeutic siRNAs, allele-specific gene therapy has received increasing attention in recent years. Among the 2.6 million SNP polymorphisms characterized in the human genome, the proportions of the substitutions are A/G, 32.77%; C/T, 32.81%; A/C, 8.98%; G/T, 9.06%; A/T, 7.46%; and C/G, 8.92. Around 65% of SNPs occur between the two purines (A or G) or the two pyrimidines (C or T), which were thus named transition substitutions (27). Targeting these abundant SNP alleles raised a serious challenge to siRNA design. For example, to silence a disease-causing mutation from wild-type A to G, siRNA designed against the mutant G allele has a C in the guide strand. When this siRNA interacts with the wild-type A allele, a C:A mismatch is formed between them. As revealed in our study, the C:A mismatch is well tolerated in the processing of RNAi and causes minimal interference with gene silencing efficacy. To confront this challenge, a G→A and an A→G transition mutation were chosen to test in our study. By placing the mismatches in their most sensitive positions, as identified in the study, the specifically designed siRNA showed great discrimination against these two alleles, indicating that the principle proposed here is generally applicable to allele-specific gene silencing.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors appreciate Drs Iain C. Bruce, Chunmei Cao and Xi Zhang for critically reading this manuscript.

FUNDING

National Natural Science Foundation of China (30873187, 30771085 and 30871385), the Beijing Natural Science Foundation (5092011), the National High-tech R&D Program of China (2006AA02Z104 and 2007AA02Z165), the National Basic Research Program of China (2007CB512100), the National 985 Program and grants from the Department of Education of China (20070001011 and 200800010019). Funding for open access charge: the National High-tech R&D Program of China (2007AA02Z165).

Conflict of interest statement. None declared.

REFERENCES

1. Elbashir,S.M., Harborth,J., Lendeckel,W., Yalcin,A., Weber,K. and Tuschl,T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, **411**, 494–498.
2. Siomi,H. and Siomi,M.C. (2009) On the road to reading the RNA-interference code. *Nature*, **457**, 396–404.
3. Castanotto,D. and Rossi,J.J. (2009) The promises and pitfalls of RNA-interference-based therapeutics. *Nature*, **457**, 426–433.
4. de Fougerolles,A., Vornlocher,H.P., Maraganore,J. and Lieberman,J. (2007) Interfering with disease: a progress report on siRNA-based therapeutics. *Nat. Rev. Drug Discov.*, **6**, 443–453.
5. Dykxhoorn,D.M., Schlehuber,L.D., London,I.M. and Lieberman,J. (2006) Determinants of specific RNA interference-mediated silencing of human beta-globin alleles differing by a single

nucleotide polymorphism. *Proc. Natl Acad. Sci. USA*, **103**, 5953–5958.
6. Klootwijk,R.D., Savelkoul,P.J., Ciccone,C., Manoli,I., Caplen,N.J., Krasnewich,D.M., Gahl,W.A. and Huizing,M. (2008) Allele-specific silencing of the dominant disease allele in sialuria by RNA interference. *FASEB J.*, **22**, 3846–3852.
7. Miller,V.M., Xia,H., Marrs,G.L., Gouvion,C.M., Lee,G., Davidson,B.L. and Paulson,H.L. (2003) Allele-specific silencing of dominant disease genes. *Proc. Natl Acad. Sci. USA*, **100**, 7195–7200.
8. Schwarz,D.S., Ding,H., Kennington,L., Moore,J.T., Schelter,J., Burchard,J., Linsley,P.S., Aronin,N., Xu,Z. and Zamore,P.D. (2006) Designing siRNA that distinguish between genes that differ by a single nucleotide. *PLoS Genet.*, **2**, e140.
9. 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.
10. 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.
11. Scacheri,P.C., Rozenblatt-Rosen,O., Caplen,N.J., Wolfsberg,T.G., Umayam,L., Lee,J.C., Hughes,C.M., Shanmugam,K.S., Bhattacharjee,A., Meyerson,M. *et al.* (2004) Short interfering RNAs can induce unexpected and divergent changes in the levels of untargeted proteins in mammalian cells. *Proc. Natl Acad. Sci. USA*, **101**, 1892–1897.
12. Snove,O. Jr. and Holen,T. (2004) Many commonly used siRNAs risk off-target activity. *Biochem. Biophys. Res. Commun.*, **319**, 256–263.
13. Jackson,A.L., Burchard,J., Schelter,J., Chau,B.N., Cleary,M., Lim,L. and Linsley,P.S. (2006) Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA*, **12**, 1179–1187.
14. Dahlgren,C., Zhang,H.Y., Du,Q., Grahn,M., Norstedt,G., Wahlestedt,C. and Liang,Z. (2008) Analysis of siRNA specificity on targets with double-nucleotide mismatches. *Nucleic Acids Res.*, **36**, e53.
15. Du,Q., Thonberg,H., Wang,J., Wahlestedt,C. and Liang,Z. (2005) A systematic analysis of the silencing effects of an active siRNA at all single-nucleotide mismatched target sites. *Nucleic Acids Res.*, **33**, 1671–1677.
16. 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.
17. Chiu,Y.L. and Rana,T.M. (2002) RNAi in human cells: basic structural and functional features of small interfering RNA. *Mol. Cell*, **10**, 549–561.
18. Doench,J.G., Petersen,C.P. and Sharp,P.A. (2003) siRNAs can function as miRNAs. *Genes Dev.*, **17**, 438–442.
19. He,Z. and Sontheimer,E.J. (2004) "siRNAs and miRNAs": a meeting report on RNA silencing. *RNA*, **10**, 1165–1173.
20. 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.
21. Phillips,W.A., St Clair,F., Munday,A.D., Thomas,R.J. and Mitchell,C.A. (1998) Increased levels of phosphatidylinositol 3-kinase activity in colorectal tumors. *Cancer*, **83**, 41–47.
22. Vivanco,I. and Sawyers,C.L. (2002) The phosphatidylinositol 3-Kinase AKT pathway in human cancer. *Nat. Rev. Cancer*, **2**, 489–501.
23. Samuels,Y., Wang,Z., Bardelli,A., Silliman,N., Ptak,J., Szabo,S., Yan,H., Gazdar,A., Powell,S.M., Riggins,G.J. *et al.* (2004) High frequency of mutations of the PIK3CA gene in human cancers. *Science*, **304**, 554.
24. Song,J.J., Liu,J., Tolia,N.H., Schneiderman,J., Smith,S.K., Martienssen,R.A., Hannon,G.J. and Joshua-Tor,L. (2003) The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat. Struct. Biol.*, **10**, 1026–1032.
25. Wang,Y., Juranek,S., Li,H., Sheng,G., Tuschl,T. and Patel,D.J. (2008) Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature*, **456**, 921–926.
26. Wang,Y., Sheng,G., Juranek,S., Tuschl,T. and Patel,D.J. (2008) Structure of the guide-strand-containing argonaute silencing complex. *Nature*, **456**, 209–213.
27. Zhao,Z. and Boerwinkle,E. (2002) Neighboring-nucleotide effects on single nucleotide polymorphisms: a study of 2.6 million polymorphisms across the human genome. *Genome Res.*, **12**, 1679–1686.