A role for human Dicer in pre-RISC loading of siRNAs

Kumi Sakurai¹², Mohammed Amarzguioui¹³, Dong-Ho Kim¹⁴, Jessica Alluin¹, Bret Heale¹²⁵, Min-sun Song¹, Anne Gatignol⁶, Mark A. Behlke⁷ and John J. Rossi¹²,*

¹Department of Molecular and Cellular Biology, Beckman Research Institute, City of Hope, 1450 East Duarte Road, Duarte, CA 91010, ²Irell and Manella Graduate School of Biological Sciences, City of Hope, 1450 East Duarte Road, Duarte, CA 91010, USA, ³The Biotechnology Centre of Oslo, Gaustadalleen 21, 0349 Oslo, Norway, ⁴Genolution Pharmaceuticals, Inc. Songpa-gu, Seoul, 138-040, Korea, ⁵Human Genetics Unit, MRC, Western General Hospital, Crew Road, Edinburgh Scotland, UK, ⁶Virus-cell Interactions Laboratory, Lady Davis Institute for Medical Research, McGill University, Montréal, Canada and ⁷Integrated DNA Technologies, Inc. 1710 Commercial Park, Coralville, IA 52241, USA

Received February 27, 2010; Revised August 19, 2010; Accepted September 7, 2010

ABSTRACT

RNA interference is a powerful mechanism for sequence-specific inhibition of gene expression. It is widely known that small interfering RNAs (siRNAs) targeting the same region of a target-messenger RNA can have widely different efficacies. In efforts to better understand the siRNA features that influence knockdown efficiency, we analyzed siRNA interactions with a high-molecular weight complex in whole cell extracts prepared from two different cell lines. Using biochemical tools to study the nature of the complex, our results demonstrate that the primary siRNA-binding protein in the whole cell extracts is Dicer. We find that Dicer is capable of discriminating highly functional versus poorly functional siRNAs by recognizing the presence of 2-nt 3' overhangs and the thermodynamic properties of 2–4 bp on both ends of effective siRNAs. Our results suggest a role for Dicer in pre-selection of effective siRNAs for handoff to Ago2. This initial selection is reflective of the overall silencing potential of an siRNA.

INTRODUCTION

RNA interference (RNAi) is an evolutionarily conserved process mediated by a multi-component ribonucleoprotein complex called the RNA induced silencing complex (RISC) (1–2). RNAi can be described as having at least two well-defined steps: the initiation step, where the ribonuclease (RNase) III enzyme Dicer processes dsRNAs into 21–22-nt-long duplexes (3), and the effector step, in which Argonaute 2 (Ago 2), a core endonuclease of the RISC, executes RNAi (4–6). Two classes of small regulatory RNAs—small interfering RNAs (siRNAs) and microRNAs (miRNAs)—are shown to associate with RISC as RNAi trigger molecules. SiRNAs, 21–23 nt fully complementary double-stranded RNAs (dsRNAs) with 2-nt 3' overhangs, guide sequence-specific degradation of complementary messenger RNAs once incorporated into RISC (7–10). In contrast, miRNAs form imperfect duplexes with the target mRNAs, most often in the 3'-UTR, and the miRNA-containing RISC blocks translation or leads to destabilization of the targeted message (11–18).

The RNAse III family member Dicer, along with co-factor RNA-binding proteins, processes long dsRNAs and pre-miRNAs into the functional 21–23-nt siRNA or miRNA duplexes. The functional roles of Dicer have been most well studied in Drosophila where two distinct enzymes are involved in parallel pathways for small regulatory RNA biogenesis. Dicer-2 (Dcr-2) processes long-double-stranded precursors and generates siRNAs (19,20) while Dicer-1 (Dcr-1) processes pre-miRNAs into mature miRNAs (20). Importantly, these two enzymes utilize different co-factors for their respective functions. Dcr-1 is associated with the dsRNA-binding protein Loquacious (Loqs) (21,22) whereas Dcr-2 is associated with the dsRNA-binding protein R2D2 (19,23). In contrast, mammals possess only a single species of Dicer in association with the dsRNA-binding protein TRBP, which participates in both si- and miRNA biogenesis (24,25). Given this fundamental difference between human and fly RNAi pathways, we sought to characterize the mechanism by which highly functional siRNAs are selected in humans.

*To whom correspondence should be addressed. Tel: +1 626 301 8360; Fax: +1 626 301 8271; Email: jrossi@coh.org

© The Author(s) 2010. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
In humans, the HIV-1 TAR RNA-binding protein (TRBP) has been characterized as one of the factors that assemble Dicer-generated small RNAs into RISC by interacting directly with Dicer and Ago2 (26–29). TRBP has been shown to have versatile roles in small regulatory RNA biogenesis. TRBP has high-sequence similarity to Drosophila Loqs and mammalian PACT (30), and is required for both si- and miRNA silencing (27,29). The Dicer–TRBP interaction is RNA independent (29). TRBP is required for the recruitment of Ago2 to the siRNA–Dicer complex (27) and is proposed to participate in coupling of the RNAi initiation and effector steps. Other studies have shown that MOV10 (31,32) and RNA Helicase A (33) are additional factors that associate with human RISC, forming an active complex. However, little is known about their modes of action in RNAi.

The selection of the guide strand which serves as a sequence-specific RNAi trigger largely determines the efficacy of RNAi (3). Effective siRNAs most often have a thermodynamically unstable 5’ guide strand end (34), and often have asymmetric loading of the guide versus passenger strands (35). In Drosophila, the orientation of the guide strand of the siRNA is determined by the Dcr-2/R2D2 heterodimer siRNA binding, with Dcr-2 binding at the 5’-end of the guide strand and R2D2 binding the 5’-end of the passenger strand (36). R2D2 is also sensitive to the 5’-end stability and the presence of 5’ phosphates on the siRNA (36).

In comparison to observations with the siRNAs, which often show a wide range of potency (37–39), Dicer processing of its substrates during the initiation step of RNAi was shown to result in better programming of RISC (9,19,28). Furthermore the polarity of Dicer processing preferentially defines which strand serves as the guide (40). Conversely, it is known that Dicer processing is dispensable for RISC assembly and siRNA-mediated cleavage of target-RNA transcripts when siRNAs were exogenously introduced into cells lacking functional Dicer (41–43). Interestingly, in Drosophila, an organism which expresses both siRNAs and miRNAs as regulatory small RNAs, the Dcr-2/R2D2 heterodimer was shown to act as a gatekeeper promoting the incorporation of siRNAs for assembly with Ago2 while disfavoring miRNAs as loading substrates for Ago2 (44). The sorting process depends upon the structure of the small RNA duplex and not upon Dcr-2 cleavage.

Aside from the appropriate selection of the guide strand for incorporation into RISC, there is the concern of off-targeting in which either the sense or antisense strand binds to non-targeted sequences triggering inhibition of expression of these transcripts (45–48). Therefore in order to enhance the efficacy of RNAi, it is essential to better understand the mechanism of siRNA selection and strand incorporation in the RNA-silencing pathway.

Here we show that siRNAs associate with a high-molecular weight complex in HEK293 and HCT116 whole cell extracts. The degree of this association varies from siRNA to siRNA, in some cases dramatically between two siRNAs differing by only a single base pair. We observe that Dicer is the core siRNA-duplex-binding component of the whole cell extract complex. Dicer preferentially recognizes the 2-nt 3’ overhangs and thermodynamically unstable ends. Since Dicer can bind to a siRNA duplex in either orientation, either strand of the duplex can be selected as the guide strand based on the thermodynamic end properties of the duplex. Thus, when siRNAs are delivered exogenously to cells, human Dicer (most likely in the form of a complex with TRBP and Ago2) serves as the primary sensor for the selection of highly functional siRNAs leading to handoff to Ago2. This selective process is based upon the presence of ribose 2-nt 3’ overhangs and the thermodynamic end properties of the siRNAs. When siRNAs cannot interact with Dicer, such as in Dicer null cells, they can bypass the Dicer-binding step and directly enter RISC. Nevertheless when Dicer is present, it serves as a primary sensor for siRNA selection when the siRNAs contain an appropriate structure.

MATERIALS AND METHODS

Synthetic siRNAs

All siRNAs used in this study were synthesized by IDT (Coralville, Iowa). The siRNAs were purified using high-performance liquid chromatography. The sequences are listed in Figures 1 and 3.

Cell extracts

HEK293 cells were grown to confluency in 10-cm dishes, harvested and washed with 1x PBS. The cell pellet was re-suspended in 0.5 ml of buffer D (20mM HEPES, pH. 7.9, 0.2mM EDTA, 0.5mM DTT, 50mM KCl, 10% Glycerol, 0.2mM PMSF) and sonicated three times for 20 s on ice with a Sonifier 450 (Branson, Danbury, CT) at a setting of 60. The supernatants were collected after 15 min of microcentrifugation for direct use in subsequent experiments.

Dual luciferase reporter assays

To generate reporter plasmids psi-hnRNPH-S (sense reporter) and psi-hnRNPH-AS (antisense reporter), a 343-bp PCR fragment of hnRNPH cDNA (Acc.: NM_005520) was cloned in the 3’-UTR of the humanized Renilla luciferase gene in the psiCHECK™-2 vector (Promega) in either the sense or antisense orientation (40). HCT116 cells at 60–80% confluence in 24-well plates were transfected in duplicate or triplicate with 100ng of reporter DNA, 25–200 pM siRNA and 0.5μl Lipofectamine2000 per well, in a total transfection volume of 500μl. Cells transfected with an irrelevant siRNA were used as mock controls, and an average was calculated from the replicates to set Renilla/luciferase expression to 100%.

For EGFP repression, psi-EGFPS1-S (sense reporter) or psi-EGFPS1-AS was generated by cloning a 245-bp PCR fragment of EGFP in either the sense or antisense orientation in the 3’-UTR of the psiCHECK™-2 vector (Promega) (40). Forty nanograms of reporter DNA, 20 or 200 pM siRNA and 0.5μl Lipofectamine2000 per well were used to transfect HEK293 cells in duplicate at
60–80% confluency in 48-well plates, in a total transfection volume of 200 μl. HEK293 cells were collected 24-h post-transfection and subjected to Dual Luciferase assays (Promega). Transfection was done in duplicate and repeated at least three times. *Renilla* luciferase expression was normalized to internal control Firefly luciferase expression. Cells transfected with an irrelevant siRNA were used as mock controls, and an average was calculated from the replicates to set *Renilla*/Firefly luciferase expression to 100%.

For siDicer, siAgo2 or siTRBP treatment, HEK293 cells were seeded on 10-cm dishes and transfected with 40-nM siRNA-targeting Dicer, Ago2 or TRBP. The sequences of the siRNAs used in this study are as follows: Dicer, 5'-UUUGUGUUCCGAGGCUCGUU; Ago2, 5'-GCAAGGAAGUCAUCUGA; TRBP, 5'-UCUACGAAUUUCAGUAGG. Forty-eight hours later, the siRNA treated cells were seeded on either 48-well plates or 6-well plates for Dual Luciferase assays or RT-qPCR, respectively. In the second-round transfection, the same procedure described above was used for Dual Luciferase assays. Either with or without 20-nM siDicer, siAgo2 or siTRBP in the second round of transfection produced the same results.

**IC50 measurements**

To generate 50% inhibitory concentration (IC50) curves, HEK293 or HCT116 cells at 60–80% confluency in 48-well plates were transfected in duplicate with 40 ng of reporter DNA (either psiEGFP-S or –AS with HEK293 cells and psi-H1-S or –AS with HCT116 cells), siRNAs ranging from 0.2 pM to 15 nM in serial dilutions (a minimal of 10 points measurements) and 0.5 μl Lipofectamine2000 per well in a total transfection volume of 200 μl. Cell lysates were collected 24-h post-transfection and subjected to the Dual-Luciferase Reporter Assay System (Promega). For each siRNA, transfection was carried out in duplicate and repeated at least three times. Normalization of *Renilla* luciferase expression to internal control Firefly luciferase expression was as described above. IC50 curves were generated with GraphPad Prism version 5.01 using the dose response curve equation: \( Y = \text{Bottom} + (\text{Top} - \text{Bottom})/ (1 + 10^{[(\log IC50 - X) \times \text{HillSlope}]}) \) with variable slope. The 95% confidence intervals of IC50 values for siRNAs in this study are listed in Figures 3 and 4.

**Recombinant proteins**

N-terminally His6-tagged TRBP2 was expressed from the pET11a (+) vector (Novagen) in the *Escherichia coli* BL21DE3 strain. Following 4-h induction with 0.5 mM IPTG, purification of His6-TRBP2 was performed with Nickel-Nitrilotriacetic acid (Ni2+-NTA) (the QIA expression system, Qiagen) under denaturing conditions. Efficiency of elution steps was determined with 12% SDS–PAGE/comassie blue staining. The protein was further purified by dialyzing with a 3500 MWCO dialysis cassette (#66110, PIERCE, Rockford, IL), concentrated with an Amicon Ultra centrifugal filter 30KMNWL and stored in buffer D containing 10 mM TCEP. Purity of His-TRBP was also determined by Coomassie blue staining and western-blot analyses with an α-His tag antibody (ab18184, Abcam Inc., Cambridge, MA). Recombinant human Dicer was purchased from Ambion, and its concentration was determined with A260 measurements. Titration-binding assays were performed to evaluate optimal concentrations of the recombinant proteins (data not presented), ranging from 0.1 μg to 5 μg for Dicer and 4 ng to 300 ng for TRBP, 2 μg of rDicer enzyme and 80 ng of rTRBP were used in subsequent gel shift assays (Figure 5).

**Immunodepletion assays**

Immunodepletion analyses were carried out as previously described (49). To immunodeplete Dicer, TRBP or Ago 2 from HEK293 whole cell extracts, 2 mg of pre-cleared extract were incubated with each primary antibody—α-Dicer (15 μg of sc-30226, Santa Cruz Biotechnology, Inc., Santa Cruz, CA), α-TRBP (25 μl of TRBP-JBX, (50)), α-Ago2 (15 μg of ab57113, Abcam Inc., Cambridge, MA). In the case of Dicer/TRBP or Ago2/ TRBP double immunodepletion, TRBP precipitation was performed following Dicer or Ago2 precipitations. Pre-immune serum (2.5 μl of sc-2007 or sc-2025, Santa Cruz Biotechnology, Inc.) was used as a control. Following 2-h incubation, protein A/G plus agarose beads (Santa Cruz Biotechnology, Inc.) were added and incubated overnight at 4°C with gentle agitation. The following day, supernatants were collected for additional incubation with the beads for 6 h. After removing the beads by a brief microcentrifugation, buffer D was added to bring the cell extracts to a final concentration of 4 mg/ml. Of total protein extracts, 15 μg were used in the Gel Shift assays. Of the cell extracts, 150 μg were also used in western-blot analyses to evaluate the efficiency of the immunodepletions. Equal amounts of total protein were electrophoresed in SDS–PAGE gels (16 cm × 16 cm), and each protein was detected sequentially. After confirming the reproducibility of the procedure, the membrane was cut in three pieces for three independent Western blot analyses. The first trough was used to detect Dicer (~215 kDa), Ago2 (~100 kDa), TRBP (~40 kDa) and β-actin (~37 kDa), and the whole membrane was scanned at once for consistency using two different wave lengths. Antibodies α-Dicer (ab14601, Abcam Inc.), α-Ago2 (ab57113, Abcam Inc.), α-TRBP (ab42018, Abcam Inc.) or α-β-Actin (A5441, Sigma-Aldrich, St. Louis, MO) were used to detect the corresponding proteins. Additionally, the molecular weights of TRBP and Beta-actin are very close. Thus in order to avoid stripping the membrane between each probing, the secondary antibodies—anti-mouse-800 nm (610-132-121, Rockland, Gilbertsville, PA) or anti-rabbit-680 nm (A21109, Molecular Probes/Invitrogen)—were detected at two different wavelengths on an Odyssey Infrared imaging system (LI-COR biosciences, NE). Scanned results in green, red or yellow were visualized as black and white. The reproducibility of the immunodepletions was verified four times.
Gel-shift assay
For incubations with radio-labeled siRNAs, single-stranded siRNAs were 5′-end labeled with T4 polynucleotide kinase (New England Biolabs) and γ-32P-ATP (MP Biomedical) for 1 h and purified with a G25 column (GE Healthcare). For siRNA duplexes, independently radiolabeled sense and antisense strands were mixed and annealed at 95°C in a heat block for 2 min and allowed to cool slowly at room temperature by removing the heat block from the heat source until they reached ambient temperature. SiRNA duplexes that were radiolabeled after annealing yielded identical results. Unlabeled EGFP-1A, –B, hnRNPH-1 and -3 siRNA duplexes were used as cold competitors. For siRNAs with only one 5′-end-labeled strand, the sense or antisense strand was annealed with the unlabeled complementary strand to generate S/A or S/A* EGFP-1A or unlabeled sense and antisense strands were annealed and used as a cold competitor. Of total protein in buffer D, 15.0 µg were incubated with 1 nM (4 × 10^6 c.p.m.) of labeled siRNA duplexes for 30 min at room temperature. In competition assays, cold siRNA competitors of concentrations at 2, 5, 10, 25, 50 or 100 nM were incubated in addition to 1 nM of labeled EGFP-1A siRNA duplex. The samples were mixed with 4X native gel loading dye and resolved in a 4.5% non-denaturing polyacrylamide gel (29:1) (20 × 30 cm glass plates, 1.5 mm spacers) with 1X TBE for 3.0 h at 200V and 4°C. The gel was dried at 80°C for 20 min and fixed with a 10% Acetic Acid: 10% methanol mix for 1 h for autoradiography. Quantifications of the RNA–protein complexes were performed with densitometric scanning of the gel (Typhoon scanner). The percent complex binding was calculated as percent bound [bound/ (bound + free) × 100] of siRNAs relative to input siRNA incubated without the cell extract. Percentage bound of the HEK293 WCE sample was set to 100%, and relative intensities of complexes formed in immunodepleted cell extracts were calculated accordingly.

For the gel supershift assays, HEK293 cell extract (15.0 µg of total protein) was incubated with 1 µg of antibody specific for Dicer (ab14601), Ago1 (07-599, Upstate, Lake Placid, NY), Ago2 (07-590, Upstate) or 1 µL TRBP antibody Ab672 (51) for 15 min prior to the incubation with siRNA duplexes. The samples were resolved in a composite polyacrylamide (3%)-agarose (0.5%) gel and run for 7 h at 100V at 4°C. The gel was fixed with a 10% Acetic Acid: 10% methanol mix for 20 min and dried at 80°C for 1 h for autoradiography. The antibodies used in the supershifts were also used in western blot analyses (49) to detect the proteins in the cell extract. The assays described above were repeated multiple times and representative gel figures are shown in this study.

RT-qPCR
Twenty-four hours after the second round of transfection with 20-nM siRNA-targeting Dicer, Ago2 or TRBP, and 200-pM siEGFP-1A variants, RNA was extracted with RNA STAT-60 (TEL-TEST), treated with Turbo DNA-Free Kit (Ambion, Austin, Texas) and reverse-transcribed into complementary DNA (cDNA) using random hexamer primers and Maloney Murine Leukemia Virus (MMLV) reverse transcriptase (Invitrogen, Carlsbad, California). One RNA sample of each preparation was processed without MMLV RT as a negative control in subsequent real time PCR reactions. Quantitative analysis of Dicer, Ago2 and TRBP expressions was performed by real time PCR SYBR Green I (Bio Rad) analysis (C1000 Thermal Cycler, Bio Rad, Hercules, California). Dicer, Ago2 and TRBP expression was detected using 50 ng of cDNA, amplified with corresponding primer sets Dicer-A (5′-CATGGATAGTGGAATGTGACAC-3′), Dicer-B (5′-CTACTTCCCCAGTGCATCCTG-3′); Ago2-A (5′-CGCGTTCCGAAGGCTGCTCTA-3′), Ago2-B (5′-TGGGCTGTGCTTAAACGCTT-3′) and 5′TRBP (5′-GGGCCGCTGTATGAGGACCC-3′), 3′TRBP-2 (5′-GACCCCGAAAGTGAATAG-3′). GAPDH expression was detected as internal control using 50 ng of cDNA, with primers GAPDH-A (5′-CGCTTCTCGTCTTCTGTT-3′) and GAPDH-B (5′-CGCTTTGTTCTGAGCGAAT-3′). PCR conditions are as follows: (i) Dicer, Ago2 and GAPDH PCR: 95°C for 5 min, followed by 40 cycles of 95°C for 30 s, 60°C for 30 s and 72°C for 1 min, and (ii) for TRBP and GAPDH: 95°C for 5 min, followed by 40 cycles of 95°C for 1 min, 60°C for 1 min and 72°C for 1 min.

RESULTS
Differential whole-cell-extract complex formation and target knockdown correlate with siRNA end structure
To test the structural requirements for binding of siRNAs to components of whole cell extracts and target knockdown, various 3′-end modified versions of a single EGFP-targeting siRNA (EGFP-1A) were incubated in whole-cell extracts for evaluation of complex formation and in cell culture for target down-regulation assays. The sequence of the antisense (guide) strand of the siRNA duplexes in an experimental setting could result in masking the target knockdown. When the siRNA duplexes in Figure 1a were incubated with a HEK239 whole-cell extract, the strongest binding was observed with the 21-nt 3′ overhang (EGFP-1A or 19 + 2 in Figure 1a and b) and was reduced with subsequent alterations in the siRNA end structure. Binding reduced <40% with the 21+0 (blunt 3′-end) siRNA and complex formation and was further reduced with 5′ overhangs (20-1 or 19-2 in Figure 1a and b), indicating that the complex is formed preferentially with siRNAs via the 3′ overhangs, consistent with binding being initiated with siRNA 3′ overhang-PAZ domain interactions.

To determine whether the 3′-end modifications introduced in the EGFP-1A variants affect RNAi efficacy, the above siRNA variants were tested in the Dual Luciferase reporter assays (Figure 1c). Silencing efficiency for many siRNAs is known to be concentration dependent. Since a high concentration of a given siRNA in an experimental setting could result in masking the
effects of various structural features, we chose to use concentrations of 20 and 200 pM, or a concentration at the EGFP-S1A IC50 value (Figure 3) versus one 10-fold above this. The variant siRNAs were co-transfected with the psiEGFP-S1 reporter plasmid in HEK293 cells and the effects on Renilla luciferase activity were monitored. The 19+2 siRNA was most potent, consistent with it being the best binder in the extracts. The other siRNAs followed in graded fashion with the +1, blunt, −1 and −2 siRNAs being progressively worse in target knockdown. These

![Figure 1. The effect of 3’ overhang on complex formation and RISC function. (a and b) Four variants of EGFP-S1A siRNA differing in the number of nucleotides at 3’-end (a) were evaluated for complex formation competence (b) +2, +1, +0, −1, −2 indicate numbers of nucleotides at the 3’-end as overhangs [19+2, 20+1, 21+0 (blunt ends), 20 – 1 (with 1-nt 5’ overhangs) and 19 – 2 (with 2-nt 5’ overhangs)]. The bottom strands are the antisense sequences and the upper strands the sense sequences. Without HEK293 whole-cell extract, 19+2 was included as a negative control. The band shifts were quantified by densitometric scanning of the gel (Typhoon scanner) and percent bound siRNAs were calculated by [bound/(free+bound)] x 100] of siRNAs relative to input siRNA without cell-extract incubation. The assays were conducted multiple times with similar results, and a representative gel is shown here. (c) Dual Luciferase assays. To determine the efficiency of the 3’-end modified siRNAs in intracellular target knockdown efficiency, silencing by the antisense strand of the end modified EGFP-S1A siRNA duplexes was assayed in HEK293 cells by co-transfecting the psi-EGFP-S1 sense reporter and 20- or 200-pM siRNAs. Target-specific Renilla luciferase expression was normalized to the control Firefly luciferase expression for all replicates (determined from multiple co-transfections).
cellular assays were consistent with their binding activities in the cell extracts.

We next tested the specificity of complex formation with single- or double-stranded siRNAs to determine if the high-molecular weight complex observed is an RNA duplex specific (Figure 2). When either the EGFP1A labeled sense (S*) or antisense (AS*) strands were tested individually in the binding assays, no binding to the high-molecular weight complex was observed (Figure 2) whereas a faster migrating complex was observed with both of the single-strand RNAs. We further tested the duplex requirement of the complex using EGFP1A containing either one (S*/AS or S/AS*) or both (S*/AS*) 5'-radiolabeled strands. Both S*/AS and S/AS* were incorporated into the complex, and cold EGFP1A (S/ AS) efficiently competed with the S*/AS* siRNA–protein complex, confirming that the complex in this study is duplex specific (Figure 2a). These observations demonstrate that the complex formation in this study is preferentially with the duplex form of siRNAs. Our observations are also consistent with reported observations that the human RLC forms exclusively with the duplex form of siRNAs (27–29).

To test the complex formation with several different canonical siRNAs harboring 19-base duplexes and 2-nt overhangs, we created a set of ten siRNAs targeting EGFP (EGFP1-A–J, Figure 3a) in which each sequence is shifted by 1 nt towards the 5'-end of the target RNA. Several, but not all of the siRNAs incorporated published siRNA features that contribute to the overall efficiency of RNAi (such as an A at position 3 or 19 and 30–50% overall GC content in the passenger strand) (37,39,52–54). Duplex end free energies—from 2 to 5 bp on both the 5'- and 3'-ends—of the siRNAs were also calculated and are listed in Table 1, panel a. When the 10 32P-labeled anti-EGFP siRNA duplexes were incubated in HEK293 cell extracts and subjected to EMSA analyses, the duplexes formed complexes that migrated similarly. However, while the majority of siRNAs formed strong complexes, EGFP1B, I and J formed weak complexes (Figure 3b). We next examined the relative levels of siRNA-directed mRNA knockdown using Dual Luciferase assays. IC50 values for both sense (S) and antisense (AS) targets of the siRNAs were determined using dose-dependent target knockdown measurements (Figure 3c and Supplementary Figure S1). When the EGFP1 siRNAs were co-transfected with Renilla luciferase target reporter plasmids—psiEGFP1-S or psiEGFP1-AS (40)—into HEK293 cells, differential knockdown efficiencies of the siRNAs were observed (Figure 3c). The general trend observed was that the siRNAs with the stronger complex formation percentages (Figure 3b) had better overall combined IC50 values (Figure 3c). In contrast, the three siRNAs with the poorest complex formation percentages (EGFP1B, I and J, Figure 4b) had high-combined IC50 values (Figure 3c), and in the case of S1J there was no measureable target knockdown for the psi-EGFP1-AS target (Figure 3c).

As representatives of good and poor binders in the HEK293 cell extract assay, EGFP1A and B were subsequently incubated with HCT116 cell extracts. Consistent with the observation from the HEK293 cell extract, EGFP1A and B formed strong and weak complexes, respectively, in the HCT116 cell extract (data not shown), validating that the siRNA-binding differentials are not extract dependent.

To further explore the relationship between complex formation in the whole-cell extracts and RNAi efficacy, we created an additional set of siRNAs targeting a human gene encoding the heterogeneous nuclear ribonucleoprotein (hnRNP) H and tested these for
complex formation in HCT116 whole cell extracts (Figure 4). Seven 21-nt-long siRNAs (hnRNP H1–7) were designed with their sequences shifted 1 nt toward the 3’-end of the target RNA hnRNP H (Figure 5a) (40) and these were evaluated in the extracts by EMSA. As with the first set of siRNAs, duplex end-free energies—from 2 to 5 bp on both the 5’- and 3’-ends—of the siRNAs were also calculated and are listed in Table 1, panel b. When the seven 32P-labeled anti-hnRNP H siRNA duplexes were incubated in HCT116 cell extracts and subjected to EMSA analyses, the duplexes formed complexes that migrated similarly (Figure 4b). However, among the seven anti-hnRNP H siRNAs, which were shifted by a single base from each other, we observed a range of binding efficiencies: hnRNP H1, H2, H4 and H7 were effectively incorporated into the complex, whereas...
hnRNPH3, H5 and H6 were poorly incorporated (Figure 4b).

To determine the target knockdown efficacies of the siRNAs in intact cells the relative levels of target inhibition in HCT116 cells was evaluated using Dual Luciferase assays. The reporter system contained an hnRNPH PCR fragment in either a sense (psi-H1-S) or antisense (psi-H1-AS) orientation in the 3′-UTR of a Renilla luciferase reporter gene (40). The individual siRNAs were co-transfected with the sense or antisense oriented target plasmids and IC50 values were determined for both siRNA strands (Figure 4c and Supplementary Figure S2). The three siRNAs that showed poor complex formation (H3, H5 and H6) had one or both strands with high-IC50 values, whereas those that had stronger complex interactions (H1, H2, H4 and H7) had overall lower combined S/AS IC50 values (Figure 4c).

The differential strength of complex formation between EGFPs1A versus EGFPs1B and hnRNPH1 versus hnRNPH3 were examined in greater detail using cold siRNAs as competitors for complex formation (Figure 5). Using EGFPs1A as a radio-labeled ligand, unlabeled EGFPs1A, EGFPs1B, hnRNPH1 or hnRNPH3 at 1:2 to 1:100 molar excess were incubated in the HEK293 cell extracts. When hnRNPH1 (IC50s and AS = 9.03 and 5.04 pM, Figure 4c) was incubated with EGFPs1A (IC50s and AS = 18.5 and 17.6 pM, Figure 3c), it competed effectively throughout the range of concentrations tested, while EGFPs1B (IC50s and AS = 50.4 and 80.4 pM, Figure 3c) was a weak competitor. HnRNPH3 (IC50s/AS = 29.7 and 7.66 pM, Figure 4c) was moderately competitive with EGFPs1A for complex formation. These results demonstrate that siRNAs with good potency in cells are good competitors for complex formation in cell extracts, and vice versa, implicating selective siRNA incorporation takes place in the RISC-loading complex (RLC) as well as in RISC itself.

### Examination of the roles Dicer, TRBP and Ago2 in complex formation

The results described above were carried out using either the HEK293 or HCT116 cell lines. In order to further explore the complex composition, we carried out additional biochemical experiments in the HEK293 cell extracts. The human RLC consists of Ago2, Dicer and TRBP (26–29). To determine whether any of these components are involved in the siRNA-complex binding observed in this study, we carried out antibody-mediated super-shift assays with HEK293 cell extracts using antibodies against Dicer, TRBP, Ago2 (and Ago1) to monitor complex formation with the EGFPs1A and hnRNPH1 siRNAs (Figure 6). Addition of the anti-Dicer antibody resulted in a strong super-shift of the complex whereas the anti-TRBP antibody yielded a weak but reproducible supershift. These results indicate that the complex observed in Figures 1–3 contain at least Dicer and TRBP, consistent with the previous observation by Pellino et al. (55). Although no supershifts were observed with the anti-Ago2 or anti-Ago1 antibodies, these antibodies did recognize the cognate proteins in the HEK293 whole cell extract (Figure 6).

To further analyze the complex components, Dicer, TRBP, Dicer/TRBP, Ago2 or Ago2/TRBP, imuno-depletions were also evaluated by western-blot analyses (Figure 7). The efficiencies of the immunodepletions were also evaluated by western-blot analyses (Figure 7). HEK293 cell extracts immunodepleted of Dicer and/or TRBP (Id Dicer, Id TRBP, Id Dicer/TRBP), and Ago2 and TRBP (Id Ago2/TRBP) resulted in a loss of siRNA binding to the major
complex. We were unable to immunodeplete Ago2 in the cell extract (Id Ago2, Figure 7), suggesting poor recognition of Ago2 by this antibody, perhaps due to interactions of Ago2 with other proteins in the extracts. Nevertheless, residual levels of Ago2 in Id Dicer, Id TRBP and Id Dicer/TRBP cell extracts had no effect on complex formation, consistent with the previous report that Ago2 does not associate with siRNA duplexes by itself (27). It is of interest to note that the combined use of anti-Ago2 and TRBP antibodies resulted in depletion of Ago2 in addition to Dicer and TRBP and this also abolished complex formation. This result is most likely a consequence of tight interactions among TRBP, Dicer and Ago2 (27). Most importantly we observed loss of siRNA gel shifts when the cell extracts were immunodepleted of Dicer, or the Dicer/TRBP heterodimer (Figure 7, lanes 4–6 and 8 in the left panel, and lanes 3–5 and 7 in the right panel). In total, these observations support Dicer as the siRNA-binding factor in the cell extracts.

To further explore the role of Dicer in the complex formation we replenished the Dicer immunodepleted extracts using purified recombinant Dicer and/or TRBP and tested for EGFPS1A complex formation (Figure 8). While rTRBP alone had no effect on complex restoration in both Id Dicer (lane 5) or Id Dicer/TRBP cell extracts (lane 13), addition of rDicer or rDicer in combination with rTRBP (lanes 6, 7, 12 and 14) resulted in restoration of the siRNA/complex formation (Figures 8a–c).
Given the differential complex formation of several of the siRNAs described above, we next asked if Dicer alone can discriminate binding of a strong versus weak complex forming siRNA. To do this rDicer was incubated with either the EGFPsi-A or -B siRNA duplexes. As we observed in the high-molecular weight complex formation, rDicer alone showed differential binding affinities between these two siRNAs, forming a much stronger interaction with the EGFPsiA than with the EGFPsiB duplex (Supplementary Figure S3). Although rDicer aggregated and was poorly resolved in the composite gel, the differential binding was reproducible.

To dissect the relative roles of Dicer, Ago2 and TRBP in selective incorporation of siRNAs in vivo, we used RNAi to knockdown each of these proteins and assayed for EGFPsiA-mediated target knockdown using the panel of EGFPsiA siRNAs that differ in their 3′-end structures (depicted and tested in Figure 1). If Dicer binding initially determines the 3′-end structure and guide strand selection, the differential efficiencies of the EGFPsiA variants should be absent in Dicer knockdown cells. Of course knockdown of Ago2 would also reduce RNAi overall since it is required for the RNAi effector step of cleaving the target transcripts. We compared the knockdowns of EGFP via the EGFPsiA panel in HEK293 in cells pre-treated with siRNAs targeting Dicer, Ago2 or TRBP (Figure 9). Consistent with a key role for Dicer in siRNA selection, the previously observed differences in knockdown efficiencies of the EGFPsiA variants were substantially reduced in cells pre-treated with the anti-Dicer siRNA. In these cells, we determined that levels of Dicer and Ago2 expression were repressed ~85% and 65%, respectively (Supplementary Figure S4). Although TRBP mRNA levels were reduced by ~60%, the effects of this reduction on the differential knockdown efficiencies by the EGFPsiA panel were minimal, which is

Figure 6. Analysis of proteins comprising the complex. Super shift assays. α-Dicer, α-TRBP, α-Ago2 or α-Ago1 antibodies were added to the HEK293 cell extract and incubated for 15 min prior to addition of either EGFPsiA or hnRNPH1 siRNA. Western analyses of proteins detected with the antibodies were also shown on the right. Of total protein, 50 μg for detection of Dicer and Ago1 or 150 μg for Ago2, TRBP and β-actin was loaded on 8% SDS–PAGE gels.

Figure 7. Analysis of RISC components required in high-molecular-complex formation. Gel-shift assays. EGFPsiA duplex was incubated with HEK293 cell extracts immunodepleted of Dicer (Id Dicer), TRBP (Id TRBP), Dicer and TRBP (Id Dicer/ TRBP), Ago2 (Id Ago2) or Ago2 and TRBP (Id Ago2/ TRBP). Evaluation of immunodepletion of RISC proteins with western blot analyses was shown on the right. Of Id cell extracts, 150 μg were loaded on a 8% SDS–PAGE gel. Pre-immune serum was used as an immunodepletion control.
consistent with our EMSA results (Figures 7 and 8). To further validate these results, we also tested the effect of Dicer knockdown on siRNA-triggered gene silencing by comparing knockdown efficiencies of hnRNPH5 and H6 in siDicer-treated HEK293 cells (Supplementary Figure S5). When the siRNAs were co-transfected with the psi-H1-S or psi-H1-AS reporter plasmid at 200 pM, asymmetric knockdowns of the sense or antisense targets by the siRNA duplexes were observed; however knockdown efficiencies of hnRNPH5 and H6 siRNAs were substantially reduced in siDicer treated cells (Supplementary Figure S5). Our observations indicate that the overall efficacies of the siRNAs were diminished in the Dicer knockdown cells, consistent with previous observations (27,56), which is most likely a consequence of disruption of the RLC.

To discriminate the selective binding aspect of Dicer from direct incorporation of the siRNAs into Ago2, we modified EGFPs1A to generate an asymmetric 27/25-nt-long Dicer substrate form (DsiRNA)—a blunt end with two deoxyribonucleotide bases at the 3'-end of the passenger strand—forcing the duplex to be bound and undergo Dicer cleavage activity. EGFPs1B was also modified to a DsiRNA form for direct comparison, and the IC50 values were determined for both DsiRNAs (Figure 10a and b). For both EGFPs1A and B DsiRNAs, the strand harboring the 2-nt 3' overhang was predominantly selected as the guide strand (IC50, Dsi-S1A-AS = 31.9 pM, IC50, Dsi-S1A-Sense = 101 pM; IC50, Dsi-S1B-AS = 27.5 pM, IC50, Dsi-S1B-Sense = 433 pM), consistent with previous observations (40). In contrast, the dominant selection of the antisense strand (relative to the sense EGFP target) was absent or less pronounced with the use of 19+2 siRNAs. Both the dsiRNAs and siRNAs form dicer-dependent complexes in whole-cell extracts (57) (Supplementary Figure S6). These results support a model in which the orientation of Dicer binding to its substrates sets the preference for which strand will be chosen to act as the guide strand. Taken together, our observations suggest that the initial Dicer–DsiRNA interaction is an important determinant for establishing siRNA-strand selection in RISC and ultimately

Figure 8. Analysis of siRNA–RISC complex formation. (a) Reconstitution of the siRNA complex with recombinant proteins. Id Dicer (left) or Id TRBP (right) cell extracts were mixed with 2 μg of recombinant Dicer, 80 ng of recombinant TRBP or both proteins to rescue complex formation. (b and c) Percent-bound-reconstituted siRNA complexes were calculated relative to the HEK293 WCE which was set as 100%.
the potency of the siRNAs. The Dicer substrates force Dicer to enter the RNA via its PAZ domain and the 2-base 3' overhang of the Dicer substrate whereas 19 + 2 siRNAs allow Dicer to bind from either end using the 2-base 3' overhang for PAZ domain interactions.

DISCUSSION

We have used siRNA binding to a high-molecular-weight complex in whole-cell extracts to better understand what cellular factors are involved in the initial binding and discriminatory selection of highly functional versus poorly functional siRNAs. In particular, we have been interested in what happens to siRNAs following transfection into cultured cells. We find that the extent of siRNA complex formation reflects siRNA-directed overall target knockdown efficiency in cultured cells. Using antibodies to core RNAi proteins to determine the key proteins of the complex, we find that it minimally consists of Dicer/TRBP, presumably as the pre-RLC. Our observations suggest that human RISC silencing efficiency with exogenously supplied siRNAs reflects the extent of siRNA-pre-RLC formation.

Previously, mammalian Dicer was shown to be dispensable for RISC assembly and siRNA-mediated cleavage of target RNA transcripts (41–43). In contrast, the absence of a functional mouse Dicer resulted in a lack of shRNA-mediated RNAi (43), showing that Dicer-cleavage activity was necessary to generate the siRNAs from this precursor. The lack of a Dicer requirement for pre-formed siRNA function led to the conclusion that only Ago2 was required for siRNA selection. Conversely, two groups demonstrated that the absence of Dicer abolished gene silencing effects (27,56). Our results suggest that when present, Dicer itself binds siRNAs that mimic the products of Dicer cleavage (19 bp + 2-nt 3' overhangs) and Dicer can serve as a ‘gate-keeper’ discriminating between potent and non-potent siRNAs by selective binding of the siRNAs in the pre-RLC.

Schwarz et al. reported that 2–4 bp on both ends of a siRNA are important for their relative loading efficiency into RISC in Drosophila (35). Our calculated duplex end free energies for siRNAs used in this study (Table 1) are consistent with these observations. Moreover, we observed that in general, siRNAs with thermodynamically unstable base pairs at both ends effectively form high-molecular-weight complexes in cell extracts and have better overall silencing efficiencies. These observations suggest that the thermodynamic properties of 2–4 bp on both ends of effective siRNAs are used as functional determinants by the pre-RLC and determine overall silencing efficiency in humans as well.

It is widely known that siRNAs targeting the same region of a target mRNA can have a wide range of
potencies. To circumvent this obstacle, reported siRNA features such as an A at position 3 or 19 and 30–50% overall GC content in the passenger strand (37,39,52–54) are widely used to predict potent siRNAs in addition to thermodynamic end properties. However, the importance of the AU base pair at these positions remained elusive. Using two sets of siRNAs targeting EGFP site I and hnRNP H site I (40), we observed that siRNAs containing an A:U base pair at positions 1 and/or 19 result in more effective binding to the high-molecular-weight complex than those with C:G pairs at these positions. This finding is consistent with the previous report of Katoh and Suzuki (58) which suggested that an AU base pair at these positions makes the siRNA a good substrate for the pre-RLC interaction and selection as an effective siRNA for target downregulation.

The manipulations of the 2-nt 3' overhangs, which resulted in reductions in target downregulation (Figure 1) verify that the presence of a 1- or 2-nt 3' overhang is required not only for efficient binding to the high-molecular-weight complex, but is also required for efficient target downregulation. Previously, Pellino et al. (55) reported that the complex observed with HEK293 cell extracts contained Dicer. Our siRNA gel shift and binding assays along with the immunodepletion assays in HCT116 and HEK293 cell extracts demonstrate that Dicer is required for siRNA binding in these extracts, a result which is consistent with reports that the initial siRNA–Dicer interaction is via the 2-nt 3' overhang (59,60).

The marked differences in Dicer/TRBP binding to siRNAs and the correlations with efficacy of target
knockdown in cell culture, even for siRNAs differing by only a single nucleotide in sequence, point to the importance of understanding how siRNAs are selected for entry into RISC. Understanding the rules for effective siRNA design is further complicated by the fact that siRNAs which do not bind to Dicer can also be incorporated into RISC (41–43). We propose that when Dicer is present it will preferentially bind and select siRNAs harboring structurally favorable features, and this interaction leads to the subsequent handoff step to Ago2.

Selective binding of Dicer to siRNAs could potentially determine the strand to be incorporated as a guide strand. The binding of the PAZ domain to a Dicer substrate allows the catalytic domain to interact with the opposite end of the dsRNA. Studies of the interaction of the human Dicer catalytic site with the PIWI domain of Ago2 (61) and the binding of the \textit{Archaeoglobus fulgidus} PIWI protein with the 5'-end of the guide strand in the crystal structure (62) suggest the coupling of the initiation and effector steps of RNAi. The handoff of the guide strand from Dicer to Ago2 via the Dicer-catalytic domain could explain not only why Dicer processing of longer precursors often increases the efficacy of processed siRNAs (40,63), but also why the 5'-end of the siRNA within RISC leads to and determines the stable association of RISC and the target RNA (64). In this regard, Dicer substrates in the studies of Rose et al. (40) have a blunt end with two deoxybases at the 3'-end of the passenger strand, which precludes Dicer PAZ domain interactions from that end of the duplex. The polarity of Dicer entry from the end harboring the 2-base 3' overhang results in substantially reduced function of the passenger strand relative to that observed with the corresponding 19+2 siRNA (Figure 10). Moreover, a single-nucleotide shift in the siRNA relative to the target site also resulted in a strong bias in guide strand selection for the dsiRNAs as demonstrated by EGFPS1A-25/27 (IC_{50}, Dsi-S1A-Sense = 101 pM) and EGFPS1B-25/27 DsiRNAs (IC_{50}, Dsi-S1B-Sense = 433 pM) (Figure 10). Thus, Dicer interaction with the 2-nt 3' overhang is an important step in Dicer binding, and for 19+2 siRNAs Dicer can bind in two orientations. Nevertheless, the thermodynamic end properties of the siRNAs also influence Dicer binding as we have demonstrated.

Recent single-particle EM analyses of the RLC-containing Dicer/TRBP/Ago2, show that Ago2 is associated with the C-terminal domain of Dicer, while TRBP is associated with the N-terminal or helicase domain of Dicer (65,66). This structural configuration would facilitate Dicer handoff of the 5'-end of the guide strand to Ago2. Our observations that Dicer binds selectively to 2-base 3' overhangs, most likely via the PAZ domain, would orient the enzyme such that the C-terminal portion of the enzyme is associated with the 5'-end of the guide strand to facilitate handoff of the guide strand to Ago2. An important observation is that the stability of Dicer binding to an siRNA is not solely mediated by the 3' 2-base overhang, but obviously involves interactions with the duplex itself. This is made evident by the fact that many of the siRNAs used in our binding analyses contain the 2-base overhang but are not bound strongly to the Dicer-containing complex.

Thus far, the binding efficiency of Dicer/TRBP to pre-miRNAs appears to be maintained regardless of whether or not the pre-miRNA can be cleaved by Dicer (67). Our work and that of Pellino et al. (55) demonstrate that Dicer/TRBP can bind to siRNAs as well. However, these observations are in contrast to a previous study that human Dicer is incapable of siRNA binding (68). One possible explanation for this discrepancy is that the siRNA used in the previous study may fall into the group of siRNAs that are weak binders to the pre-RLC—similar to the poor binding siRNAs tested in our study.

While immunodepletion of Dicer (Id Dicer) left-residual TRBP, Id of TRBP resulted in nearly complete depletion of Dicer in the HEK 293 cell extracts, as did co-Id of Dicer/TRBP. These results indicate that the majority of Dicer is bound to TRBP and that distinctive roles of Dicer and TRBP augment each other to function as the pre-RLC. Id of Ago2/TRBP resulted in immunodepletion of Dicer, TRBP and Ago2 (Figure 7), suggesting that TRBP functions as a bridge between Dicer and Ago2.

We have demonstrated that for many siRNAs what we deemed as the passenger strand could be more effective at target knockdown than the desired antisense guide strand. Since Dicer/TRBP binding to a siRNA duplex determines the strand to be selected as the guide strand, off-targeting will often be a consequence of both strands being incorporated into RISC. Thus the two strands of a siRNA can also compete with one another for incorporation into RISC following Dicer selection, thereby potentially reducing the potencies of both.

In conclusion our results demonstrate that human Dicer plays an important role in selection of efficacious siRNAs via direct binding to siRNAs that harbor certain structural features, which include the 2-nt 3' overhangs and the thermodynamically unstable base pairing at the ends of the duplex. These results are similar yet distinct from \textit{Drosophila} in which R2D2 and Dcr-2 bind to the thermodynamically stable and unstable ends of siRNA duplexes, respectively. Our results demonstrate that the human Dicer enzyme can recognize and selectively bind siRNAs with thermodynamically favorable termini. Our results demonstrate that an often overlooked feature of 19+2 siRNAs is the ability of both strands to be chosen for RISC entry. This is sometimes observed for miRNAs as well, but often only one strand serves as the guide. In this regard, Dicer substrate siRNAs mimic the miRNA pathway, providing strong polarity for guide strand selection, perhaps a useful attribute for minimizing off target effects and competition between the two strands for entry into RISC.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**ACKNOWLEDGEMENTS**

The authors acknowledge members of the Rossi laboratory for helpful discussions and support, especially Lisa
Scherer for the critical reading of this article. They are grateful for the suggestions of Y. Chen, R.J. Lin, J Shively and J. Termini.

**FUNDING**

Norwegian Research Council (a postdoctoral fellowship to M.A.); Arnold and Mabel Beckman Foundation (to D.K.); National Institutes of Health (grants AI29329, AI42552 and HL074704 to J.J.R.). Funding for open access charges: National Institutes of Health National Heart Lung and Blood Institute and National Institute of Allergy and Infectious Diseases.

**Conflict of interest statement.** M.A.B. is employed by IDT DNA, which manufactures siRNAs and dicer substrate siRNAs. J.J.R. is chairman of the SAB at Dicerna Pharmaceuticals which uses Dicer substrate siRNAs.

**REFERENCES**


