# Identification of RNA Pseudoknot-Binding Ligand That Inhibits the −1 Ribosomal Frameshifting of SARS-Coronavirus by Structure-Based Virtual Screening

So-Jung Park,[†] Yang-Gyun Kim,[‡] and Hyun-Ju Park*[,†]

[†]School of Pharmacy and [‡]Department of Chemistry, Sungkyunkwan University, Suwon 440-746, Korea

ⓢ Supporting Information

ABSTRACT: Programmed −1 ribosomal frameshifting (−1 RF) is an essential regulating mechanism of translation used by SARS-CoV (severe acute respiratory syndrome coronavirus) to synthesize the key replicative proteins encoded by two overlapping open reading frames. The integrity of RNA pseudoknot stability and structure in the −1 RF site is important for efficient −1 RF. Thus, small molecules interacting with high affinity and selectivity with the RNA pseudoknot in the −1 RF site of SARS-CoV (SARS-pseudoknot) would disrupt −1 RF and be fatal to viral infectivity and production. To discover ligands for the SARS-pseudoknot by virtual screening, we constructed a 3D structural model of the SARS-pseudoknot and conducted a computational screening of the chemical database. After virtual screening of about 80 000 compounds against the SARS-pseudoknot structure, high-ranked compounds were selected and their activities were examined by in vitro and *cell-based* −1 RF assay. We successfully identified a novel ligand **43** that dramatically inhibits the −1 RF of SARS-CoV. This antiframeshift agent is an interesting lead for the design of novel antiviral agents against SARS-CoV.
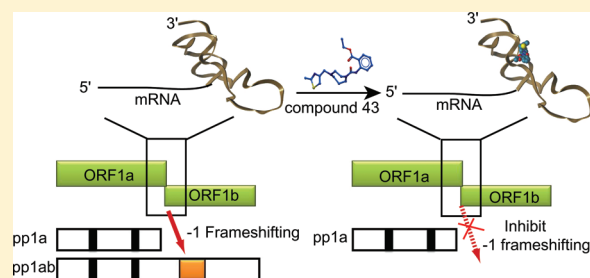
INTRODUCTION

Programmed −1 ribosomal frameshifting (−1 RF) is an essential common strategy utilized by RNA viruses that regulate the relative expressions of proteins encoded in two overlapping translational read frames.[1] −1 RF is a change in the translational reading frame such that the codon starts one nucleotide upstream from the original position in the mRNA sequence. Many pathogenic RNA viruses such as IBV (infectious bronchitis virus),[2] SARS (severe acute respiratory syndrome) coronavirus,[3] and HIV-1 (human immunodeficiency virus-1)[4−6] have an absolute requirement for a −1 RF event during translation in order to synthesize structural and enzymatic proteins. As small changes in −1 RF efficiency have been found to significantly decrease the production of infectious virions,[7] the discovery of antiframeshift agents which alter −1 RF efficiencies will greatly benefit the development of antiviral therapies against critical pathogenic viruses.[8] If such an agent acting against −1 RF is identified, it should compliment the current portfolio of drugs available for combination therapy treatment of infected individuals.

A −1 RF requires two cis-acting mRNA signals: the slippery site,[9] at which the ribosome changes reading frame during a ribosomal pause, and a downstream region of a secondary mRNA structure, most often an RNA pseudoknot.[10] An mRNA pseudoknot induces −1 RF by hindering the elongating ribosome by positioning itself just outside the narrow mRNA entry tunnel of the ribosome.[11] In addition, the integrities of mRNA pseudoknot stability and structure are the important features for maintaining the efficiency of −1 RF.[12] Small molecules that bind tightly to the

RNA pseudoknot may interfere with the ribosome's ability to engage with the stable upper RNA pseudoknot during −1 RF. Thus, the RNA pseudoknot inducing −1 RF is an attractive target for drug development in RNA viruses using −1 RF.

The −1 RF of SARS coronavirus (SARS-CoV) is essential for the synthesis of viral RNA-dependent RNA polymerase (RdRp) and other replication components which are essential genes for viral replication.[13] The stability of RNA pseudoknot in the −1 RF site of SARS-CoV also produces a dramatic effect on −1 RF efficiency.[14−16] Mutants of the slippery site that lower SARS-CoV −1 RF efficiency eventually reduced the infectivities of infected clones of SARS-CoV by more than 4 orders of magnitude.[16] An example of targeting the SARS-CoV −1 RF is antisense oligonucleotides designed to disrupt the RNA pseudoknot in the −1 RF site of SARS-CoV, which have been shown to be very effective in inhibiting −1 RF.[17] Therefore, the RNA pseudoknot of the SARS-CoV −1 RF site (SARS-pseudoknot) could provide a promising target for new antiviral agents which interfere with translational regulation. There are no previous publications on the small molecule ligands targeting SARS-pseudoknot which alter the −1RF efficiency.

We have published a successful study that identified compounds which regulate −1 RF by interacting with the biotin-binding RNA pseudoknot using computational structure-based virtual screening methods.[18] In the present study, we selected the

ligands for the SARS-pseudoknot using virtual screening, and their −1 RF efficiencies were analyzed via in vitro and cell-based −1 RF assays. We identified a novel compound, named **43**, which drastically inhibits the SARS-CoV −1 RF, an interesting lead for the design of novel antiviral agents which may regulate the protein synthesis induced by −1 RF in SARS-CoV.

■ EXPERIMENTAL SECTION

Modeling of SARS-Pseudoknot. In the first step, the structure of the SARS-pseudoknot was predicted using the PSEUDOVIEWER program (available at http://pseudoviewer.inha.ac.kr),[19] which is executable within a Web browser on any PC with Windows as its operating system. The SARS-pseudoknot sequence (GeneBank accession number AY291315) was saved in .txt format. PSEUDOVIEWER took as an input an RNA sequence with its structure data in bracket view. The outline view displays the structure in the form of a backbone in which loops were replaced by polygons and helices by line segments. On the next step, the 3D structure of the SARS-pseudoknot was built on the Sybyl 6.9 system using tools of build/edit based on the 2D structure obtained by PSEUDOVIEWER. The SARS-pseudoknot 3D structure was prepared in .pdb format using Sybyl 6.9. After building the 3D structure, this structure was optimized and all procedures were carried out using AMBER8.0 software package[20] (http://ambermd.org). The initial model was optimized in the explicit water system using the SANDER module. To run the SANDER program, the coordinates of constructed 3D structure of the SARS-pseudoknot was generated by XLEaP program. This structure was loaded into XLEaP, built with AMBER ff99 force field, and solvated with 10 Å of TIP3P water using a truncated octahedron periodic box. This structure was neutralized by 67 sodium ions. The minimization procedure for solvated RNA consisted of two approaches. In the first step, minimization was performed with fixing RNA and just minimized the positions of water and ions. The 1000 steps of minimization including 500 steps of steepest descent minimization and 500 steps of conjugate gradient minimization were run with restraining RNA position. A force constant of 500 kcal mol^{-1} Å^{-2} and 10 Å cutoff was used. In the second step, the whole system was minimized with running 25 000 steps without fixing RNA. Periodic boundaries and the SHAKE algorithm were used during molecular dynamics (MD) two-step simulation. The condition of MD involved a constant temperature of 300 K and was controlled using the Langevin temperature equilibration scheme under 1 atm of pressure. In the first step of MD, 15ps of MD simulation was carried out with 10 Å cutoff and time step of 2 fs per steps, and every 100 steps the coordinates were written to trajectory file. In the second step, 1ns of MD simulation was run using SHAKE. SHAKE algorithm removed the center of hydrogen mass motion every 500 steps. SHAKE was run by default tolerances that were used during the simulation steps. The production of MD trajectories with measuring RNA backbone root-mean-square deviation (rmsd) from the starting structure and averaging of th ecoordinate set were conducted by using the ptraj module. RNA backbone rmsd along the simulation was 1.71 ± 0.72 Å. This averaged structure was briefly minimized using SANDER. The MD simulation with 33 Mg^{2+} counterions was also carried out under the same condition as those of the Na^{+} simulation. The production of MD trajectories with measuring RNA backbone rmsd from the starting structure and averaging of the coordinate set was conducted by using the ptraj module. RNA backbone rmsd was 1.91 ± 0.86 Å. The difference between two optimized structures was evaluated by the backbone rmsd value. (Supporting Information Figure S1).

Virtual Screening of the Chemical Database Using the DOCK4.0 Program. Flexible docking screening was performed with the "Flexible_ligand" option using the program DOCK4.0, which requires the following receptor files (SARS-pseudoknot): a mol2 format file containing coordinates for hydrogen and the Kollman charge; a PDB

file coordinate without hydrogen or water atoms; and a PDB file excluding the active site. The active site of the SARS-pseudoknot was defined as all residues within 8.5 Å centered by U21, one of the key residues for maintaining −1 RF efficiency. U21 is positioned at the junction of stem1/stem2 whose active site includes 11 residues. The LeadQuest database (Tripos Inc.), the commercially available compound collections, was used in this study. The database was docked into the active site of the built SARS-pseudoknot structure. The ligand coordinates were provided in mol2 format. From the first docking output, 500 highly ranked compounds were selected, and saved as the first filtered database. All selected compounds were analyzed with the standard Tripos force field, and Gasteiger−Hückel charges were assigned to the ligand atoms. The minimization was run until convergence reached a maximum derivative of 0.001 kcal mol^{-1} Å^{-1}. The second docking was run again for the minimized previously filtered compounds. Each top-scored pose per ligand was analyzed by energy score and visual inspection. Finally, 58 high-scored compounds were selected and purchased to examine their effects on −1 RF efficiency. To validate the activity of the best hit compound **43**, it was resynthesized by LeadGenex Inc. (Daejeon, Korea). The chemical structure and purity (>98%) of **43** were confirmed by HPLC, proton NMR, and high-resolution fast atom bombardment (FAB) mass spectrometry (data given in the Supporting Information). Mass spectrometry data were obtained from the Korea Basic Science Institute (Daegu) on a Jeol JMS 700 high-resolution mass spectrometer. All biological activity data presented in this paper are obtained from resynthesized **43**.

Template Constructs for −1 Frameshifting Assay. p2luc-UUUAAC vector was based on the p2luc dual luciferase vector.[21] The slippery sequence (UUUAAC) was inserted into a p2luc vector insert site between the renilla luciferase (rluc) gene and the firefly luciferase (fluc) gene. The wild type of the SARS-pseudoknot sequence was inserted into p2luc-UUUAAC at the XhoI and SpeI sites by using annealed duplex DNA oligomers. This vector containing the slippery site and the SARS-pseudoknot site was named p2luc-SARS-WT. In the present study, to compare the selectivity of candidate compounds for the SARS-pseudoknot, we used two different RNA pseudoknots. First, the biotin-binding pseudoknot sequence[22] was inserted into the p2luc-UUUAAC vector instead of the SARS-pseudoknot. The aptamer RNA pseudoknot sequences are named p2luc-biotin-pseudoknot which regulates −1 RF. Second, the PEMV-pseudoknot (RNA pseudoknot from pea enation mosaic virus RNA1) sequence,[23] which is a well-known RNA pseudoknot, was inserted into the p2luc-UUUAAAC vector instead of SARS-pseudoknot. This vector named p2luc-PEMV and used to identify specificity of candidate compounds for SARS-pseudoknot. All used RNA pseudoknots for this study have their own specific sequence and regulate −1 RF. Mutation in the SARS-pseudoknot sequence was introduced by inserting duplex oligomers containing mutant sequences at XhoI and SpeI for the pseudoknot. We deleted nine nucleotides (13 457−13 465) which were a part of the **43** binding site, predicted by the docking program. This plasmid is named p2luc-SARS-9D (Figure 7d).

In Vitro Transcription/Translation Coupled Assay. The transcription/translation coupled (TNT) assay system (Promega) was used following the manufacturer's instructions. Template DNA (500 ng) and 0.5 μL of 100 mM candidate compounds were used in 20 μL reactions containing 16 μL of reticulocyte lysate and 0.8 μL of 10 μCi/μL [^{35}S]-labeled methionine (NEN). The nonframeshifting protein product (NRF) is 40 kDa with nine methionines, whereas the frameshifting fusion protein (RF) yields a 100 kDa protein containing 22 methionine residues. The results of the TNT assay were examined by 12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and exposed to PhosphorImager screens, and signals were quantified. The reported −1 RF efficiency was calculated with the formula $(I[\mathrm{RF}]/22)/[(I[\mathrm{RF}]/22) + (I[\mathrm{NRF}]/9)] \times 100$, where $I[\mathrm{RF}]$ is the signal intensity of the frameshifting product and $I[\mathrm{NRF}]$ the signal intensity of the nonframeshifting product.

10095

dx.doi.org/10.1021/ja1098325 |J. Am. Chem. Soc. 2011, 133, 10094–10100

Dual Luciferase Assay. −1 RF efficiencies were confirmed by dual luciferase assay (Promega) which provides more accurate −1 RF efficiency values. Activities of reporter genes such as firefly and renilla luciferases are measured by TD-20/20 luminometer (Turner Designs). The activity of firefly luciferase was calculated by 50 μL of luciferase assay reagent II. After measuring the firefly luciferase luminescence, this reaction is stopped, and then the renilla luciferase reaction was started by simultaneously adding Stop & Glo reagent (50 μL). For 10 s, light emission was recorded respectively for both luciferases and a ratio between the two measurementscalculated was calculated. All individual assays to determine the average −1 RF efficiency were repeated three times or more. −1 RF efficiency was calculated with the formula % = [(firefly luciferase of sample/renilla luciferase of sample)/(firefly luciferase of p2luci/renilla luciferase of p2luci)] × 100. Plasmid p2luci that expresses in-frame renilla-firefly fusion protein was used as a positive control for −1 RF (defined as 100% efficiency).

In Vitro Transcription Assay. To prepare plasmid for the production of "run-off" transcripts, the vector was linearized with NotI which restricted after firefly luciferase gene. After the restriction digestion, the linearized plasmid was extracted with phenol:chloroform:iso-amyalcohol, ethanol precipitate, and dissolved water before using an in vitro transcription assay. The Riboprobe RNA system-T7 (Promega) was used according to the manufacturer's protocol. Linearized template DNA p2luc-SARS-WT (1 μg) was used in 20 μL reactions containing 4 μL of transcription optimized 5× buffer, 2 μL of 100 mM dithiothreitol, 40 units of recombinant RNasin ribonuclease inhibitor, 1 μL of 2.5 mM each of rATP, rGTP, and rUTP, 2.4 μL of 100 μM rCTP, 5 μL of 10 μCi/μL [^{32}P]-rCTP (NEN), 20 units of T7 polymerase, and 0.5 μL of candidate compounds. These reactions were incubated at 37 °C for 1 h. The DNA template was removed by adding 25 units of RNase-free DNase I with further incubation for 30 min at 37 °C. The reaction mixture was phenol/chloroform extracted twice, ethanol-precipitated, vacuum-dried, and resuspended in water. The amplification reaction products were treated with 20 μL of 10% formaldehyde, 80% formamide, and 10% RNA buffer containing ethidium bromide at 65 °C for 5 min, and electrophoresed through 10% poyacrlamide gel containing 7 M urea in TAE (Tris acetate-EDTA; EDTA = ethylenediaminetetraacetic acid) buffer. The gel was analyzed by autoradiography.

Synthesis of RNA for In Vitro Translation Assay. For in vitro translation assay, RNA was synthesized using a Ribomax large-scale RNA production-T7 system (Promrga) by standard methods. A linear DNA template (10 μg) in 100 μL reaction included 25 mM rNTPs (ATP, CTP, GTP, and UTP) and a T7 enzyme mix. Reaction was mixed gently by pipetting and incubated at 37 °C for 4 h. The DNA template was removed by adding 25 unit RNase-free DNaseI. This reaction was purified using phenol/chloroform extraction and ethanol-precipitated. The RNA concentration was quantified by ultraviolet light absorbance.

In Vitro Translation Assay. The rabbit reticulocyte lysate system (Promega) was used according to the manufacturer's guideline. Synthesized RNA (1 μg) was used in 50 μL reactions containing 35 μL of rabbit reticulocyte lysate, 1 μL of 1 mM amino acid mixture minus methionine, 2 μL of 10 μCi/μL [^{35}S]-labeled methionine (NEN), 40 units of RNasin ribonuclease inhibitor, and 0.5 μL of 100 mM candidate compounds. The reaction mixtures were incubated at 30 °C for 2 h. The result of translation was analyzed the same as in vitro TNT assay.

Cell-Based **−1 Frameshifting Assay.** For cell-based −1 RF assay, the human embryonic kidney cells (HEK 293) were grown at 37 °C in Dulbecco's modified eagle's medium (DMEM) (Sigma) containing 1% penicillin−streptomycin (Hyclone) and supplemented with 10% fetal bovine serum (Hyclone). Reporter plasmids (1 μg) were transfected onto 40% confluence in 24-well plates by cationic polymer-mediated transfection using JetPEI (Qbiogen). The reporter DNA (1 μg) was transfected and incubated for 18 h. In the next step, a compound was treated and incubated for another 18 h. The cells were
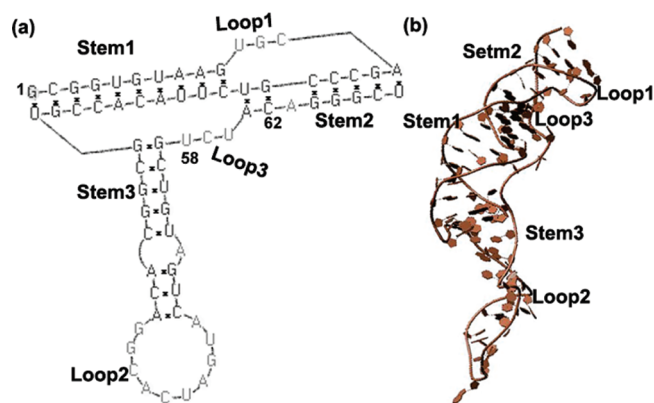

Figure 1. (a) Two-dimensional model of SARS-pseudoknot generated by the PSEUDOVIEWER[19] program. (b) Three-dimensional structural model of the SARS-pseudoknot used in this study. It was optimized by molecular dynamics simulation using the Amber 8.0 program. Brown ribbon renders the phosphate backbone of the RNA pseudoknot.

assayed for transient expression of reporter genes at 36 h after the first transfection. To dual luciferase assay transfected cells was dispensed 150 μL of passive lysis buffer (Promega) to each culture well after washing with phosphate-buffered saline (PBS) twice. The cells were then lysed by rocking the culture plates at room temperature for 15 min and spun to pellet cell debris. A 1 μL aliquot of supernatant was used instead of translation product in vitro assay. Luciferaes activities were measured by the same way of in vitro experiment.

■ RESULTS AND DISCUSSION

Modeling of SARS-Pseudoknot Structure and Virtual Screening. To perform a virtual screening, we constructed a three-dimensional structure model of the SARS-pseudoknot. Since a 3D structure of the RNA pseudoknot involved in −1 RF of SARS-CoV has not been determined, we constructed a 3D structure based on the two-dimensional predicted model using the software PSEUDOVIEWER.[19] As an output, a typical H-type two-stemmed and a unique three-stemmed (Figure 1a) RNA pseudoknot structure models were generated. Generally, RNA pseudoknot consists of two stems and two loops, but SARS-pseudoknot has an unusually long 29-nucleotide loop2. Based on mapping and 2D-NMR analysis, the possibility of the third loop and stem formation in the SARS-pseudoknot has already been reported.[14,15] Then, we selected a three-stemmed SARS-pseudoknot model for our study and optimized it by molecular dynamics (MD) simulation using the program Amber 8.0.[20] The RNA structure was neutralized by Na$^+$ and Mg^{2+} counterions, respectively, and MD simulations were performed. To examine the structural difference given by two simulations, backbone root mean square deviation (rmsd) for averaged structures was measured. The rmsd was 1.65 Å, suggesting that Na$^+$ and Mg^{2+} ions do not give significant difference in overall SARS-pseudoknot structure model (Supporting Information Figure S1). Therefore, we used the model obtained from Na$^+$ simulation for virtual screening as a target receptor structure. The optimized 3D architecture of the SARS-pseudoknot is shown in Figure 1b. Stem1 and stem2 are not coaxial, but are bent relative to each other. Loop1 crosses the deep major groove of stem2, and loop3 crosses the shallow minor groove of stem1.
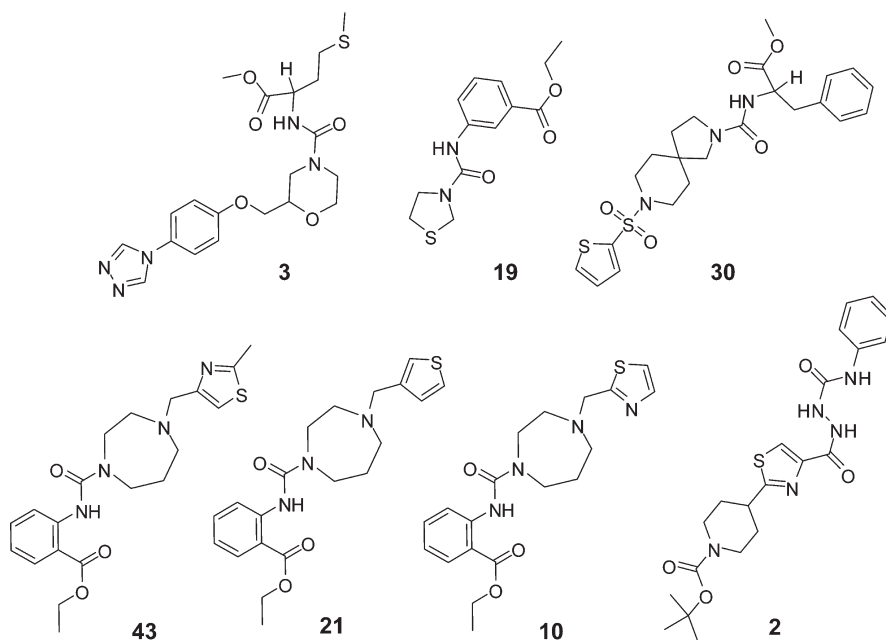
Figure 2. Chemical structures of compounds mentioned in this work.

To search ligands selectively interacting with the SARS-pseudoknot, we performed a docking screening against the SARS-pseudoknot using the DOCK4.0 program. To define the active site of the SARS-pseudoknot for virtual screening, we considered the published mutational studies of −1 RF promoting pseudoknots.[15] Mutational studies demonstrated key sequences required for −1 RF in the SARS-pseudoknot. The active site of the SARS-pseudoknot was defined as all nucleotide residues within 8.5 Å of U21, one of the key residues for maintaining −1 RF efficiency. U21 is positioned at the junction of stem1 and stem2, and the active site includes 11 residues (G11, U12, G13, U21, C22, U23, U58, C59, U60, A61, and C62). About 80 000 compounds from LeadQuest (Tripos Inc.), a commercially available chemical database, were docked into the active site of the SARS-pseudoknot. After the first DOCK run, 500 compounds which bound into the active site were selected. The top scoring pose for each ligand was energy-minimized and put into a new database. After the second docking filtering of the focused database, all compounds were ranked according to their DOCK energy scores. For any given molecule, there may be many orientations and conformations that fit into the SARS-pseudoknot active site, and each orientation is evaluated by visual inspection and the DOCK scoring function. Only the best scored orientation for each compound is stored to the output file, and compounds are ranked in order of their energy scores. A set of 58 highly ranked compounds was selected and purchased to examine their effects on −1 RF via biological assays.

In Vitro −1 Frameshifting Efficiency Assay. To test −1 RF efficiency, we designed a dual reporter construct containing the −1 RF system induced by the SARS-pseudoknot between two reporter genes, the renilla luciferase (rluc) gene and the firefly luciferase (fluc) gene (Supporting Information Figure S2). Because the termination codon UAG of the rluc gene is located immediately after the slippery sequence, if a −1 RF occurs, the termination codon of the rluc gene is not read, and renilla luciferase-firefly luciferase fusion protein is produced. If −1 RF does not occur, the termination codon UAG of the rluc gene is
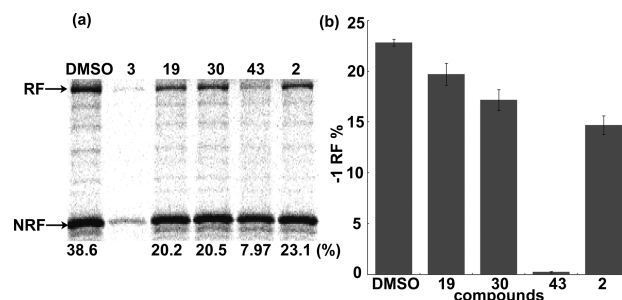

Figure 3. Measurements of −1 RF efficiencies by in vitro TNT assay. (a) SDS-PAGE analysis of [^{35}S]-methionine-labeled translation products from −1 frameshifting assay of the SARS-pseudoknot in the presence of candidate compounds (250 μM). The nonframeshifting product (NRF) is renilla luciferase protein, and the frameshifting product (RF) is firefly luciferase-renilla luciferase fusion protein. (b) The −1 RF efficiencies obtained from dual luciferase assays. The −1 RF efficiencies (%) were calculated by the formula described in the Experimental Section.

read, and only renilla luciferase is produced. We first measured −1 RF efficiency values as determined by SDS-PAGE after in vitro TNT assay. Among tested 58 compounds, 35 compounds slightly or dramatically decreased −1 RF efficiency and the rest of them have little effects (Supporting Information Figure S3). As an example, the result of the frameshifting assay in the presence of several compounds (3, 19, 30, 43, and 2 in Figure 2) that reduce the −1RF was shown in Figure 3a. In particular, with the addition of 43, −1 RF efficiency was decreased by 80% (Figure 3a: DMSO (dimethyl sulfoxide), 38.6% lane 1; 43, 7.97% lane 5). Compound 3 also extremely reduced the −1 RF efficiency value. However, we do not consider 3 a −1 RF inhibitor because it decreased both RF and NRF band intensity, suggesting that 3 may affect the normal protein synthesis.

Next, the dual luciferase assay was conducted to calculate −1 RF efficiency because this method quantifies the amount of
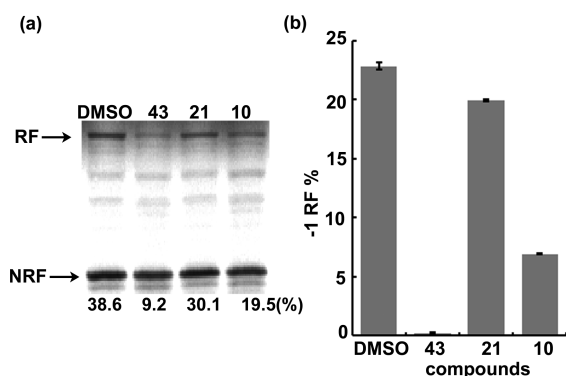
Figure 4. Measurements of −1 RF efficiencies by in vitro TNT assay. (a) The −1 RF efficiencies (%) in the presence of **43**, **21**, and **10** obtained from SDS-PAGE analysis. The nonframeshifting product (NRF) is the renilla luciferase protein, and the frameshifting product (RF) is a firefly luciferase-renilla luciferase fusion protein. (b) The −1 RF efficiencies obtained from dual luciferase assays.

reporter proteins produced and provides more accurate −1 RF efficiency values. As shown in Figure 3b, all compounds produced the similar pattern of decreasing −1 RF efficiency in the dual luciferase assay. Interestingly, **43** more dramatically reduced −1 RF to 1%, showing much stronger inhibition than the results from SDS-PAGE analysis. A −1 RF efficiency was calculated using the equation described in the Experimental Section. Both −1 RF assays revealed that **43** dramatically decreased the −1 RF induced by the SARS-pseudoknot. In case of compound **19**, **30**, and **2**, they did not decrease −1 RF as much as **43**. On the basis of the result of the TNT assay, we identified **43** as the best antiframeshift hit compound which showed the strongest inhibition of −1 RF in SARS CoV. Among the compounds that inhibit the −1 RF (see the result shown in Supporting Information Figures S3 and S4), we recognized two compounds, **21** and **10**, structurally very similar to **43**. As shown in Figure 2, their structures are distinguished only by a five-membered heterocyclic moiety in one end. Their effects on −1 RF have been compared in Figure 4 and −1 RF inhibitory activity revealed in the order of **43** > **10** > **21**. Compound **10**, having a thiazole ring instead of 2-methylthiazole of **43**, showed antiframeshift activity similar to that of **43**, while thiophene-containing compound **21** exhibited weak inhibition of −1 RF. The results suggested that the five-membered heterocyclic moiety of these compounds play a critical role in the interaction with SARS pseudoknot and regulation of −1 RF.

In Vitro Transcription and Translation Assay. Since the −1 RF assays were conducted via the transcription/translation coupled reactions, it should be verified that hit compounds have an effect on only the translation step. Thus, in vitro transcription and in vitro translation assays were conducted separately. In vitro transcription reactions were carried out at 37 °C, as described in the Experimental Section, and used the same template as the TNT assay. When their effects on transcription were examined (Supporting Information Figure S5a), addition of compounds (**43**, **21**, and **10**) gave no effect on the synthesis of RNA compared with the control DMSO, resulting in formation of the same amount of transcripts as the control no matter what their different effects on −1 RF.

To confirm that the inhibition of −1 RF by these compounds has occurred at the translation level, the next in vitro translation
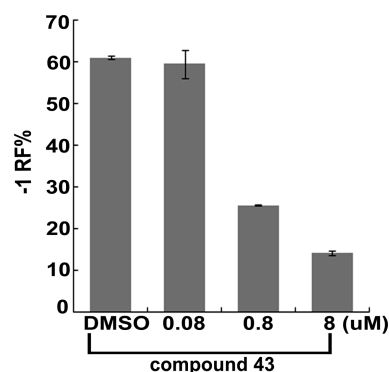

Figure 5. Concentration-dependent inhibition of SARS-CoV −1 RF by **43** in HEK 293 cells. The −1 RF % was obtained from dual luciferase assay.
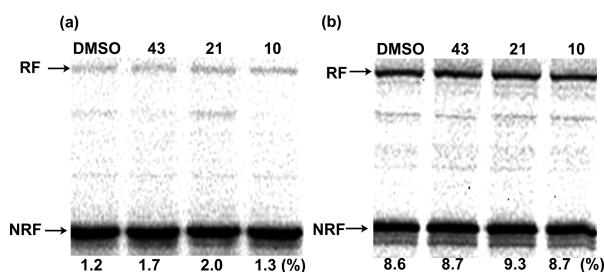

Figure 6. In vitro TNT assays using the constructs containing biotin-pseudoknot (a) and PEMV-pseudoknot (b) in the −1 RF site: (a) In vitro TNT assay using p2luc-biotin-pseudoknot; (b) result using p2luc-PEMV-pseudoknot. The −1 RF efficiencies were obtained by SDS-PAGE analysis.

assay was performed. For this assay, RNA template was synthesized using RiboMax large-scale RNA production systems (T7). An SDS-PAGE analysis on the in vitro translation reaction products was demonstrated in Supporting Information Figure S5b. The calculated −1 RF efficiency values revealed a −1 RF inhibition pattern similar to that obtained from in vitro coupled TNT assays (Figure 4a). The results supported that the compound interrupts only the translation step through its interaction with the RNA, not transcription.

Cell-Based −1 Frameshifting Assay. In addition to the in vitro −1 RF assay system, we tested efficiency of the most active compound **43** by cell-based −1 RF assay. Compound **43** was treated into the cell culture from 0.08 through 8 μM concentrations in order to monitor dose responses. As shown in Figure 5, 0.08 μM **43** slightly decreased −1 RF efficiency, but 0.8 μM **43** decreased it by more than 2-fold, and 8 μM **43** decreased −1 RF efficiency by about 5-fold relative to the control. The IC_{50} value for **43** is approximately 0.45 μM, and the results confirmed that **43** reduced −1 RF both in vitro and in a cell-based −1 RF system.

Selectivity of Hit Compounds for SARS-Pseudoknot. To examine whether hit compounds have selectivity toward only the SARS-pseudoknot, we performed an in vitro −1 RF assay using template constructs containing different RNA pseudoknot sequences because the virtual screening was conducted to target the SARS-pseudoknot structure. We compared effects of **43**, **21**, and **10** on the −1 RF induced by a biotin-binding RNA pseudoknot (biotin-pseudoknot) (Figure 6a) versus an RNA pseudoknot
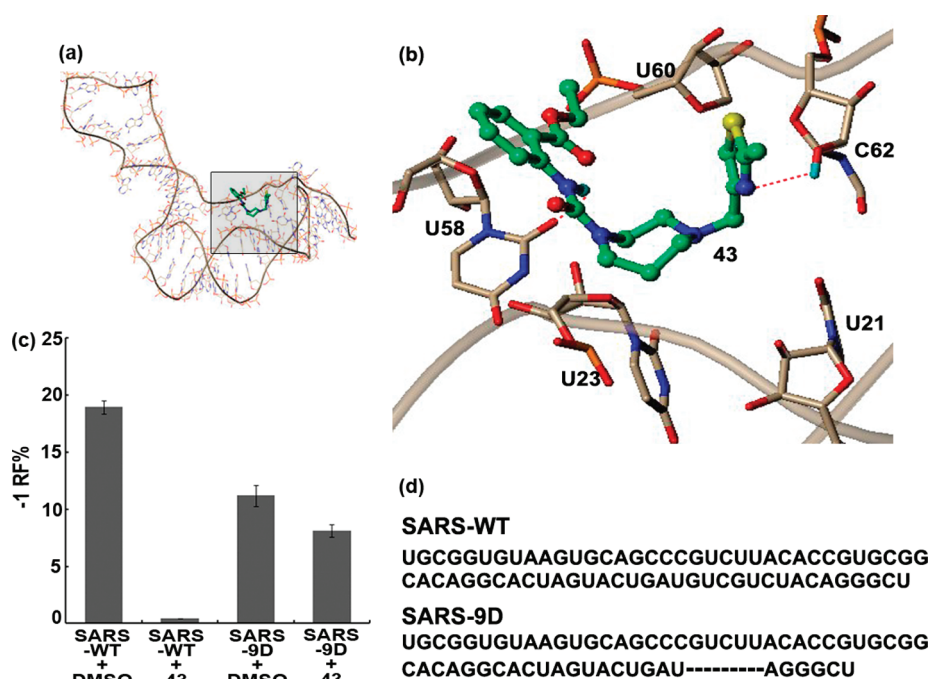
Figure 7. (a) Proposed binding pose of **43** (ball and stick) in the active site of SARS-pseudoknot, generated by DOCK 4.0. (b) Enlarged inset box in a. Docked model of **43**:SARS-psudoknot complex generated by DOCK 4.0. Several residues in the binding site are rendered in capped stick (brown carbon), **43** in ball and stick (green crbon), and red dashed lines indicate hydrogen bonds. (c) Comparison of the effect of **43** on the wild type SARS-pseudoknot (SAR-WT) versus the deletion mutant of the SARS-pseudoknot (SAR-9D). The −1 RF efficiencies were obtained in the presence of 250 μM **43** using a dual luciferase assay. (d) Sequences of mutant pseudoknot and dashed line indicates deleted residues.

from pea enation mosaic virus RNA1 (PEMV-pseudoknot) (Figure 6b) using an in vitro TNT assay. The biotin-pseudoknot and PEMV-pseudoknot have sequences and structures absolutely different from that of the SARS-pseudoknot (Supporting Information Figure S6). In the −1 RF site of template constructs, PEMV-pseudoknot and biotin-pseudoknot sequences were inserted respectively instead of the SARS-pseudoknot. All three compounds including **43** did not alter the −1 RF efficiencies induced by either pseudoknot, suggesting that they interact selectively with the SARS-pseudoknot to regulate the −1 RF. To map the correct binding site for **43** in SARS-pseudoknot structure, an enzymatic RNA structure-probing experiment is also underway.

Docking Model and Mutational Study. A docked model of **43** and the SARS-pseudoknot is illustrated in Figure 7a,b, and the model demonstrated that **43** interacts with various residues of loop3, a part of the key sequence altering −1 RF. The nitrogen atom in the thiazole ring of **43** forms a hydrogen bond (2.8 Å) with the 2′-OH group of ribose of C62, and the amide NH of **43** forms a hydrogen bond (2.6 Å) with the carbonyl oxygen atom (O2) of the U58 uracil base. The hydrogen bond between the thiazole moiety and the receptor pseudoknot was identified as one of key intermolecular interactions, which is in good agreement with biological assay results showing that antiframeshift activities fluctuated by heterocycle moiety in compounds **43**, **21**, and **10**. For this study, only neutral forms of 80 000 ligands were screened due to limited computational cost. Our hit compounds contain tertiary amine, which can be protonated in our biological experimental condition. Therefore, we examined whether protonated forms of **43** bind to SARS-pseudoknot in the same way as neutral species. Protonation of tertiary amine gave two stereoisomers, and only the (S)-form of protonated **43** fit well into the

binding pocket of SARS-pseudoknot. In comparing the docking pose between protonated **43** and the neutral form of **43**, the thiazole ring nitrogen commonly forms a hydrogen bond with the 2′-OH group of ribose of C62 (see Supporting Information Figure S7). The detailed determination of the mechanical stability of the RNA pseudoknot and −1 RF efficiency is not fully understood; however, the ligand-induced alterations in the stability and structure of the SARS-pseudoknot may contribute to the inhibition of −1 RF. On the basis of the docking model of **43** bound to the SARS-pseudoknot, we designed a mutant pseudoknot structure. We deleted sequences which are important for **43** binding (Figure 7c,d). In −1 RF assay using the construct containing a deleted pseudoknot (SARS-9D), **43** decreased the −1 RF efficiency by approximately 28%, while it reduced by 98% for wild type SARS-pseudoknot (Figure 7c). A dramatic change of antiframshift activity of **43**, depending upon the integrity of SARS-pseudoknot structure, supports that **43** may selectively interact with the SARS-pseudoknot and regulate the −1 RF.

■ CONCLUSIONS

The −1 RF mechanism is exploited by SARS-CoV for the synthesis of the enzymes required for replication, and therefore −1 RF signal is highly conserved. On the basis of this phenomenon, we have proposed that this mRNA structure is a potential target for the design of an antiviral drug. In this study, we discovered a compound that may interact with the SARS-pseudoknot and which effectively inhibits the −1 RF using computational screening of a chemical database. For some corona viruses, RNA pseudoknot structures involved in −1 RF have been characterized, but their precise contributions to −1 RF are not completely understood. Our approach can be applied to the discovery of

ligands with antiframeshift activity not only in SARS-CoV but also in many other RNA viruses in which −1 RF is crucial for replication.

■ ASSOCIATED CONTENT

ⓢ Supporting Information. Information about the template constructs for −1 RF assays (Figures S2 and S6), supporting results (Figures S1, S3, S4, S5, and S7), and structure analysis data for **43** (Figures S8−S12). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
hyunju85@skku.edu

■ ACKNOWLEDGMENT

■ REFERENCES

(1) (a) Kang, H.; Tinoco, l., Jr. *Nucleic Acids Res.* **1997**, *25*, 1943–1949. (b) Brierley, I.; Rolley, N. J.; Jenner, A. J.; Inglis, S. C. *J. Mol. Biol.* **1991**, *220*, 889–902. (c) Farabaugh, P. J. *Prog. Nucleic Acid Res. Mol. Biol.* **2000**, *64*, 131–170.

(2) (a) Brierley, I.; Digard, P.; Inglis, S. C. *Cell* **1989**, *57*, 537–547. (b) Somogyi, P.; Jenner, A. J.; Brierley, I.; Inglis, S. C. *Mol. Cell. Biol.* **1993**, *13*, 6931–6940.

(3) (a) Ziebuhr, J.; Heusipp, G.; Siddell, S. G. *J. Virol.* **1997**, *71*, 3992–3997. (b) Plant, E. P.; Dinman, J. D. *Front Biosci.* **2008**, *13*, 4873–4881.

(4) Hung, M.; Patel, P.; Davis, S.; Green, S. R. *J. Virol.* **1998**, *72*, 4819–4824.

(5) Paulus, C.; Hellebrand, S.; Tessmer, U.; Wolf, H.; Kräusslich, H. G.; Wagner, R. *J. Biol. Chem.* **1999**, *274*, 21539–21543.

(6) Baril, M.; Dulude, D .; Gendron, K.; Lemay, G.; Brakier-Gingras, L. *RNA* **2003**, *9*, 1246–1253.

(7) (a) Dulude, D.; Berchiche, Y. A.; Gendron, K.; Brakier-Gingras, L.; Heveker, N. *Virology* **2006**, *345*, 127–136. (b) Shehu-Xhilaga, M.; Crowe, S. M.; Mak, J. *J. Virol.* **2001**, *75*, 1834–1841.

(8) Dinman, J. D.; Ruiz-Echevarria, M. J.; Peltz, S. W. *Trends Biotechnol.* **1998**, *16*, 190–196.

(9) Brierley, I.; Jenner, A. J.; Inglis, S. C. *J. Mol. Biol.* **1992**, *227*, 463–479.

(10) (a) Liphardt, J.; Napthine, S.; Kontos, H.; Brierley, I. *J. Mol. Biol.* **1999**, *288*, 321–335. (b) ten Dam, E. B.; Pleij, C. W. A.; Bosch, L. *Virus Genes* **1990**, *4*, 121–136. (c) Paul, C. P.; Barry, J. K.; Dinesh-Kumar, S. P.; Brault, V.; Miller, W. A. *J. Mol. Biol.* **2001**, *310*, 987–999.

(11) (a) Namy, O.; Moran, S. J.; Stuart, D. I.; Gilbert, R. J.; Brierley, I. *Nature* **2006**, *441*, 244–247. (b) Giedroc, D. P.; Cornish, P. V. *Virus Res.* **2009**, *139*, 193–208.

(12) (a) Jacks, T.; Townsley, K.; Varmus, H. E.; Majors, J. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 4298–4302. (b) Kim, Y.-G.; Su, L.; Maas, S.; O'Neill, A.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *25*, 14234–14239. (c) Chen, G.; Chang, K. Y.; Chou, M.-Y.; Bustamante, C.; Tinoco, I., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 12706–12711.

(13) (a) Dos Ramos, F.; Carrasco, M.; Doyle, T.; Brierley, I. *Biochem. Soc. Trans.* **2004**, *32*, 1081–1083. (b) Brierley, I; Dos Ramos, F. J. *Virus Res.* **2006**, *119*, 29–42.

(14) Plant, E. P.; Perez-Alvarado, G. C.; Jacobs, J. L.; Mukhopadhyay, B.; Hennig, M.; Dinman, J. D. *PLoS Biol.* **2005**, *3*, No. e172.

(15) Su, M. C.; Chang, C. T.; Chu, C. H.; Tsai, C. H.; Chang, K. Y. *Nucleic Acids Res.* **2005**, *33*, 4265–4275.

(16) Plant, E. P.; Rakauskaite, R.; Taylor, D. R.; Dinman, J. D. *J. Virol.* **2010**, *84*, 4330–4340.

(17) Plant, E. P.; Dinman, J. D. *Nucleic Acids Res.* **2005**, *33*, 1825–1833.

(18) Park, S.-J.; Jung, Y. H.; Kim, Y.-G.; Park, H.-J. *Bioorg. Med. Chem.* **2008**, *16*, 4676–4684.

(19) (a) Han, K.; Byun, Y. *Nucleic Acids Res.* **2003**, *31*, 3432–3440. (b) Byun, Y.; Han, K. *Nucleic Acids Res.* **2006**, *34*, W416–W422.

(20) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

(21) Grentzmann, G; Ingram, J. A.; Kelly, P. J.; Gesteland, R .F.; Atkins, J. F. *RNA* **1998**, *4*, 479–486.

(22) Nix, J.; Sussman, D.; Wilson, C. *J. Mol. Biol.* **2000**, *296*, 1235–1244.

(23) Nixon, P. L.; Rangan, A.; Kim, Y.-G.; Rich, A.; Hoffman, D. W.; Hennig, M.; Giedroc, D. P. *J. Mol. Biol.* **2002**, *322*, 621–633.