JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Subscriber access provided by American Chemical Society

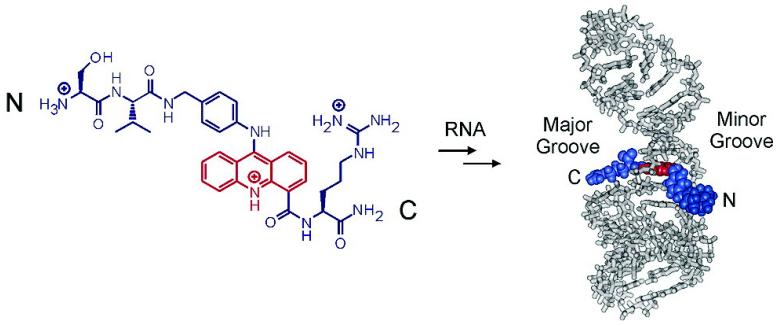
Article

Recognition of Duplex RNA by Helix-Threading Peptides

Barry D. Gooch, and Peter A. Beal

J. Am. Chem. Soc., **2004**, 126 (34), 10603-10610• DOI: 10.1021/ja047818v • Publication Date (Web): 10 August 2004

Downloaded from http://pubs.acs.org on April 1, 2009



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 2 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

ACS Publications
High quality. High impact.

Journal of the American Chemical Society is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

J|A|C|S
ARTICLES
Published on Web 08/10/2004

Recognition of Duplex RNA by Helix-Threading Peptides

Barry D. Gooch and Peter A. Beal*

Contribution from the Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112-0850

Received April 15, 2004; E-mail: beal@chem.utah.edu

Abstract: Many important biological processes, from the interferon antiviral response to the generation of microRNA regulators of translation, involve duplex RNA. Small molecules capable of binding duplex RNA structures with high affinity and selectivity will be useful in regulating these processes and, as such, are valuable research tools and potentially therapeutic. In this paper, the synthesis and duplex RNA-binding properties of EDTA·Fe-modified peptide-intercalator conjugates (PICs) are described. Peptide appendages at the 4- and 9-positions of the planar acridine ring system render these PICs threading intercalators, directing the substituents into both grooves of double helical RNA simultaneously. Directed hydroxyl radical cleavage experiments conducted with varying RNA stem-loop structures indicate a preferred binding polarity with the N- and C-termini of the PIC in the minor and major grooves, respectively. However, this binding polarity is shown to be dependent on both the structure of the PIC and the RNA secondary structure adjacent to the intercalation site. Definition of the minimal RNA structure required for binding to one of these PICs led to the identification of an intercalation site in a pre-microRNA from *Caenorhabditis elegans*. Results presented will guide both rational design and combinatorial approaches for the generation of new RNA binding PICs and will continue to facilitate the identification of naturally occurring RNA targets for these small molecules.

Introduction

Duplex RNA is involved in numerous biological phenomena, including RNA interference, the interferon antiviral response, the trafficking and editing of messenger RNA, and the generation of microRNA regulators of translation.[1−5] A large fraction of the duplex RNA involved in these processes is composed of canonical Watson−Crick base pairs; however, deviations from complementarity giving rise to mismatched, bulged, or looped nucleotides are common and often essential for RNA function. Low-molecular weight, cell-permeable molecules that bind with high affinity and selectivity to double-stranded RNA would be valuable research tools for studying these processes and have the potential to be developed into new therapeutics. Even so, rational development of RNA-binding compounds has proceeded slowly given our poor understanding of RNA recognition principles.[6] Designing compounds to target the unique arrangement of functional groups located in the grooves of a duplex RNA structure is an attractive approach to developing such reagents. However, there is a paucity of synthetically accessible structural scaffolds that can be rapidly optimized that place functionality into the grooves of an RNA duplex.[7−11]

Threading intercalation is a potentially useful paradigm for selective recognition of duplex RNA.[12] Binding in this way involves insertion of an intercalator between base pairs of a nucleic acid duplex and localization of distinguishable substituents on the planar heterocycle in opposite grooves of the helix.[13−15] One of the substituents is required to pass through the duplex (or thread) during the binding event, thus giving this binding mode its name. Given that a transient opening of the double helix at the binding site is necessary, these molecules bind preferentially to dynamic base-paired locations in RNA, such as exist adjacent to bulges and/or internal loops.[16,17] Furthermore, double-stranded RNA contains functional groups that provide unique recognition surfaces that may selectively

(1) Bass, B. L. *Annu. Rev. Biochem.* **2002**, *71*, 817−846.
(2) Hannon, G. J. *Nature* **2002**, *418*, 244−251.
(3) Nelson, P.; Kiriakidou, M.; Sharma, A.; Maniataki, E.; Mourelatos, Z. *TIBS* **2003**, *28*, 534−540.
(4) Ramos, A.; Grünert, S.; Adams, J.; Micklem, D. R.; Proctor, M. R.; Freund, S.; Bycroft, M.; Johnston, D. S.; Varani, G. *EMBO J.* **2000**, *19*, 997−1009.
(5) Stark, G. R.; Kerr, I. M.; Williams, B. R. G.; Silverman, R. H.; Schreiber, R. D. *Annu. Rev. Biochem.* **1998**, *67*, 227−264.
(6) Gallego, J.; Varani, G. *Acc. Chem. Res.* **2001**, *34*, 836−843.
(7) Lee, J.; Kwon, M.; Lee, K. H.; Jeong, S.; Hyun, S.; Shin, K. J.; Yu, J. *J. Am. Chem. Soc.* **2004**, *126*, 1956−1957.
(8) Wong, C.-H.; Hendrix, M.; Manning, D. D.; Rosenbohm, C.; Greenberg, W. A. *J. Am. Chem. Soc.* **1998**, *120*, 8319−8327.
(9) Tamilarasu, N.; Huq, I.; Rana, T. M. *J. Am. Chem. Soc.* **1999**, *121*, 1597−1598.
(10) Mayer, M.; James, T. L. *J. Am. Chem. Soc.* **2004**, *126*, 4453−4460.
(11) Luedtke, N. W.; Carmichael, P.; Tor, Y. *J. Am. Chem. Soc.* **2003**, *125*, 12374−12375.
(12) Wilson, W. D.; Ratmeyer, L.; Zhao, M.; Strekowski, L.; Boykin, D. *Biochemistry* **1993**, *32*, 4098−4104.
(13) Liaw, Y.-C.; Gao, Y.-G.; Robinson, H.; van der Marel, G. A.; van Boom, J. H.; Wang, A. H.-J. *Biochemistry* **1989**, *28*, 9913−9918.
(14) Jourdan, M.; Garcia, J.; Lhomme, J. *Biochemistry* **1999**, *38*, 14205−14213.
(15) Guelev, V.; Lee, J.; Ward, J.; Sorey, S.; Hoffman, D. W.; Iverson, B. L. *Chem. Biol.* **2001**, *8*, 415−425.
(16) Carlson, C. B.; Vuyisich, M.; Gooch, B. D.; Beal, P. A. *Chem. Biol.* **2003**, *10*, 663−672.
(17) Nonin, S.; Jiang, F.; Patel, D. J. *J. Mol. Biol.* **1997**, *268*, 359−374.

10.1021/ja047818v CCC: $27.50 © 2004 American Chemical Society

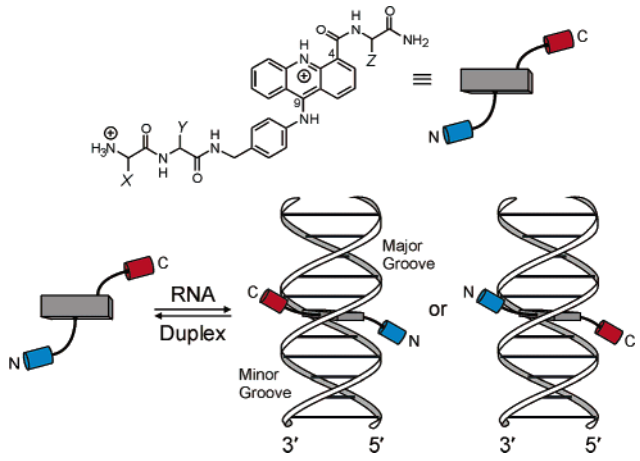


Figure 1. Schematic representation of a peptide-intercalator conjugate (PIC) bound to duplex RNA in a threading intercalation complex with two different binding polarities.

bind substituents of a threading intercalator, such as the Hoogsteen face of base pairs in the major groove or guanosine 2-amino groups and ribose 2′-hydroxyls found in the minor groove.

Although the potential for selective RNA recognition by threading intercalation is evident, no such structural motif has been described for RNA where the groove location of the substituents (i.e., the binding polarity) has been defined. Defining the orientation of threading intercalation complexes is critically important if one wishes to rationally modify the ligand to target recognition sites found in the grooves adjacent to an intercalation site.

We have previously described the synthesis and RNA-binding properties of various peptide-intercalator conjugates (PICs).[16,18,19] PICs contain an intercalating domain (acridine or quinoline) integrated into the backbone of a peptide such that the N- and C-termini project from opposite edges of the heterocycle (Figure 1). This design imparts the potential for threading intercalation with the peptide termini in opposite grooves in the complex. Using in vitro evolution of the RNA (SELEX), we demonstrated that the PIC N-Ser-Val-Acr-Arg-C, where Acr is our acridine-containing amino acid, binds preferentially to 5′-CpG-3′ sites in duplex RNA flanked by an internal loop and a single bulged uridine.[16] The groove location, however, of the peptide substituents in such complexes was not defined.

Presently, high-resolution structural information for RNA−acridine complexes is limited to X-ray crystal structures of simple, nonthreading acridines intercalated into dinucleotides.[20−22] Consequently, the binding orientation of an acridine-based threading intercalator has not been thoroughly examined for duplex RNA. Additionally, the effect nucleic acid structure has on the polarity with which such a threading intercalator binds has not been investigated.

Covalent modification at either the N- or C-terminus of a PIC with EDTA·Fe has allowed us to examine the RNA binding

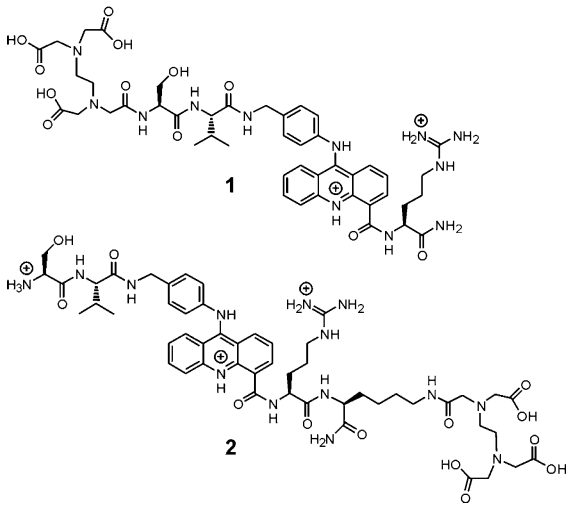


Figure 2. Structures of EDTA-modified PICs used to study the binding polarity of these threading intercalators.

properties of threading intercalators of this type (Figure 2). Using these reagents, we obtained directed hydroxyl radical cleavage data supporting a model with the PIC N-terminus in the minor groove and the C-terminus in the major groove of an RNA stem-loop structure. Furthermore, systematic deletions of nucleotides within an internal loop of the RNA alters PIC binding affinity and polarity, demonstrating a dependence on non-Watson−Crick structure immediately flanking the intercalation site. Finally, these studies led to the definition of the minimal RNA structure required for binding this PIC and the identification of a naturally occurring PIC binding site in a pre-microRNA from *Caenorhabditis elegans*.

Materials and Methods

General. All reagents for solid-phase peptide synthesis (SPPS) were purchased from NovaBiochem, except for ethylenediaminetetraacetic acid (EDTA) monoanhydride, which was synthesized according to Ebright et al.[23] Glassware for SPPS and synthesis of EDTA monoanhydride was oven-dried at 125 °C overnight and cooled in a desiccator prior to use. Reagents for DNA amplification, RNA synthesis and radioactive labeling, hydroxyl radical cleavage, and ribonuclease footprinting were purchased from Amersham Pharmacia Biotech (high-purity solution NTP set (ATP/CTP/GTP/UTP, 100 mM), RNase-Free deoxyribonuclease I (DNase I), RNAguard ribonuclease inhibitor (porcine)), New England Biolabs (T4 polynucleotide kinase (PNK)), PerkinElmer Life Sciences ([γ-^{32}P]ATP (6000 Ci/mmol)), Stratagene (Pfu Turbo DNA polymerase), and USB (PCR nucleotide mix (dATP/dCTP/dGTP/dTTP, 10 mM), shrimp alkaline phosphatase (SAP)). All commercial reagents were used as purchased without further purification. Distilled, deionized water was used for all aqueous reactions and dilutions. Electrospray ionization (ESI) mass spectra were recorded on a Finnigan LCQ ion-trap mass spectrometer. Storage phosphor autoradiography was carried out using imaging screens purchased from Kodak. A Molecular Dynamics Typhoon 9400 and ImageQuant 5.2 software (Molecular Dynamics) were used to obtain and analyze all data from storage phosphor imaging screens.

Preparation of N-EDTA-Ser-Val-Acr-Arg-C (1). PIC **1** was prepared according to established methods.[16]

Preparation of N-Ser-Val-Acr-Arg-Lys(EDTA)-C (2). PIC **2** was synthesized using 9-fluoronylmethoxycarbonyl (Fmoc)-protected amino acids according to standard SPPS protocols. Rink Amide MBHA resin

(18) Carlson, C. B.; Beal, P. A. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1979−1982.
(19) Krishnamurthy, M.; Gooch, B. D.; Beal, P. A. *Org. Lett.* **2004**, *6*, 63−66.
(20) Sakore, T. D.; Jain, S. C.; Tsai, C.-C.; Sobell, H. M. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 188−192.
(21) Sakore, T. D.; Reddy, B. S.; Sobell, H. M. *J. Mol. Biol.* **1979**, *135*, 763−785.
(22) Westhof, E.; Sundaralingam, M. *Proc. Natl. Acad. Sci. U.S.A* **1980**, *77*, 1852−1856.
(23) Ebright, Y. W.; Chen, Y. P.; Pendergrast, S.; Ebright, R. H. *Biochemistry* **1992**, *31*, 10664−10670.

(0.64 mmol/g loading, 50 mg, 0.032 mmol) was added to a 15 mL fritted-glass filter flask reactor equipped with a screw cap and Teflon stopcock. After preswelling in DMF, the resin was shaken for 30 min on a Burrell Model 75 Wrist Action shaker in a solution of 20% piperidine in DMF to remove the Fmoc protecting group from the primary amine of the resin, followed by consecutive washes with DMF, MeOH, and DMF again. The initial lysine residue, protected at the ϵ-amino group with methyltrityl (Mtt), was prepared by dissolving the Fmoc amino acid (Fmoc-Lys(Mtt)-OH, 100 mg, 0.16 mmol, 5 equiv), *N*-hydroxybenzotriazole (HOBt) (24 mg, 0.16 mmol, 5 equiv), and 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) (61 mg, 0.16 mmol, 5 equiv) in DMF. After the solution was cooled to 4 °C, diisopropylethylamine (56 μL, 0.32 mmol, 10 equiv) was added and the solution stirred at 4 °C for 30 min. The activated amino acid solution was subsequently added to the resin. The coupling occurred over ~5 h and was monitored by ninhydrin (Kaiser test). When a negative Kaiser test was obtained, the resin was rinsed with DMF, MeOH, and again with DMF. After the Fmoc protecting group was removed, the resin-bound peptide was extended by coupling Fmoc-Arg(Pbf)-OH in an identical manner.

Following Fmoc deprotection of the arginine residue, the *N*-hydroxysuccinimide ester of a previously described acridine-containing amino acid was added as a THF solution and shaken overnight.[18] The resin was then washed and dried under high vacuum in preparation for removal of the allyloxycarbonyl (Alloc) group.

Freshly distilled DCM was added to the dry resin under an argon atmosphere. Solutions of $Pd(PPh_3)_4$ (37 mg, 0.032 mmol, 1 equiv) and $PhSiH_3$ (95 μL, 0.768 mmol, 24 equiv) in DCM were delivered via an oven-dried glass syringe, and the reactor was shaken for 30 min at room temperature.[24] The procedure was repeated once, revealing the primary amine of the acridine amino acid residue. Following Alloc removal, the resin was washed with 0.5% diisopropylethylamine in DMF, 0.5% (w/v) sodium diethyldithiocarbamate in DMF, straight DMF, MeOH, and again DMF. The remaining amino acids, Fmoc-Val-OH and Fmoc-Ser(Trt)-OH, were coupled according to the procedures described above.

Following coupling of the N-terminal serine residue, the solid support was treated with 1% trifluoroacetic acid (TFA) and 5% triisopropylsilane (TIS) in DCM to remove the Mtt protecting group of lysine, leaving other acid-labile protecting groups intact.[25] The resulting compound was washed with DCM, DMF, MeOH, and again DMF. The washed resin was allowed to react with EDTA monoanhydride (44 mg, 0.16 mmol, 5 equiv) for 15 h in DMF. Upon removal of the N-terminal Fmoc group, the resin was prepared for peptide release by washing with DMF, MeOH, AcOH, and again MeOH and dried under high vacuum overnight. The dry resin was suspended in a cleavage cocktail of TFA/TIS/PhOH/water (88:5:5:2) and shaken for 5 h to effect cleavage from the resin and amino acid side chain deprotection. The resulting solution containing **2** was collected and concentrated under reduced pressure. Ether precipitation of the residue, followed by extraction with water, and subsequent neutralization of the acidic aqueous layer with triethylamine gave the crude product.

The lyophilized crude product was HPLC-purified on a reverse-phase C-18 column (4.6 × 250 mm, Vydac) following absorbance of the compound at $\lambda_{max} = 442$ nm. The PIC was eluted with a gradient of H_2O (0.1% TFA) (mobile phase A) and 60% ACN (0.1% TFA) (mobile phase B) (0−4 min, linear increase to 40% B; 4−16 min, linear increase to 80% B; 16−18 min, linear increase to 100% B) (1.5 mL/min).

PIC **2** was analyzed by ESI-MS by directly infusing into the instrument a solution of **2** dissolved in a 1:1 mixture of ACN/H_2O with 1% formic acid at a flow rate of 10 μL/min. Typical ESI-MS conditions utilized a capillary voltage of 6 V at 175 °C. Ultrapure nitrogen was used as a sheath gas at a flow rate of 60 (arbitrary units). Xcaliber software was used to run the instrument and analyze the data. ESI-MS: 1087.4 ($[M + H]^+$), 544.3 ($[M + 2H]^{2+}$), 363.2 ($[M + 3H]^{3+}$).

RNA Synthesis and 5′-End-Labeling. The 75 nucleotide in vitro selected RNA and the nucleotide deletion RNAs were generated by runoff transcription with T7 RNA polymerase. First, a dsDNA PCR product was amplified from a pUC-19 derived plasmid (*4nt*)[16] or chemically synthesized templates (*3nt*, *2nt*, *1nt*, *del*, and pre-miR39) using 25mer and 45mer DNA oligonucleotide primers. Sequences are as follows: *3nt* template, 5′-CGGAAGCTTCTGCTACATGCAATGGGATTCGCGGTTGAAGCCGACGCCCACTGTCACGTGTAGTATCCTCTCCC-3′; *2nt* template, 5′-CGGAAGCTTCTGCTACATGCAATGGGTTCGCGGTTGAAGCCGACGCCCACTGTCACGTGTAGTATCCTCTCCC-3′; *1nt* template, 5′-CGGAAGCTTCTGCTACATGCAATGGGTCGCGGTTGAAGCCGACGCCCACTGTCACGTGTAGTATCCTCTCCC-3′; *del* template, 5′-CGGAAGCTTCTGCTACATGCAATGGGCGCGGTTGAAGCCGACGCCCACTGTCACGTGTAGTATCCTCTCCC-3′; *pre-miR39* template, 5′-CGGAAGCTTCTGCTACATGCAATGGTACACGGTTGAAGCCGACGACCACTGTCACGTGTAGTATCCTCTCCC-3′; 25mer primer, 5′-CGGAAGCTTCTGCTACATGCAATGG-3′; 45mer primer, 5′-GCGAATTC*TAATACGACTCACTCTC*GGGAGAGGATACTACACGTG-3′, the T7 promoter is italicized. The PCR product was extracted with PhOH/$CHCl_3$, precipitated with EtOH, and dissolved in 100 μL reaction volumes consisting of transcription buffer (80 mM HEPES, 25 mM $MgCl_2$, 2 mM spermidine, 30 mM DTT, pH 7.5), NTPs (8 mM), and ribonuclease inhibitor (1 U/μL). Transcription was initiated with T7 RNA polymerase (0.03 mg/mL) and continued overnight at 37 °C. The reaction mixture was treated with RNase-free DNase I (0.5 U/μL) and $CaCl_2$ (1 mM) for 1 h at 37 °C and purified on a 10.5% denaturing polyacrylamide gel. The transcribed RNA was visualized by UV shadowing, excised from the gel, and eluted overnight via the crush and soak method. After filtration, the resulting solution was extracted with PhOH/$CHCl_3$ and precipitated with EtOH, and the RNA concentration was determined by measuring the absorbance at 260 nm. For the preparation of 5′-end-labeled RNA, 30 pmol of the transcript was treated with SAP (0.1 U/μL) for 45 min at 37 °C, followed by heat inactivation of the enzyme for 15 min at 65 °C. The solution of dephosphorylated RNA was immediately treated with T4 PNK (0.5 U/μL), [γ-^{32}P]-ATP (2 mCi/mL), and DTT (1 mM) and incubated for 45 min at 37 °C. Labeled RNA was gel purified, visualized by storage phosphor autoradiography, and isolated as previously described.

RNA Stem-Loop Secondary Structure Prediction. Secondary structure prediction of all RNA stem-loops was performed using the web-based program *mfold* (version 3.0) developed by Dr. Michael Zuker.[26,27] All RNA stem-loops were predicted to fold into single predominant structures.

PIC-EDTA·Fe Cleavage Experiments. PICs **1** and **2** were dissolved in a 2-fold excess of $Fe(NH_4)_2(SO_4)_2$ (aq) to give the desired concentration of Fe^{2+}-containing EDTA-modified PIC immediately preceding hydroxyl radical cleavage of the RNA. PIC-EDTA·Fe−RNA complexes were formed by incubating varying concentrations of EDTA·Fe **1** and **2** with ~5 nM end-labeled RNA for 15 min in buffer (50 mM Bis-Tris·HCl, 100 mM NaCl, and 10 mM $MgCl_2$, pH 7.0) and 10 μg/mL of yeast $tRNA^{phe}$. The resulting complexes were probed by initiating hydroxyl radical formation with the addition of 0.01% H_2O_2 and 5 mM DTT. The reaction was allowed to proceed for 45 min at room temperature. Affinity cleaving reactions were quenched by the addition of distilled, deionized water, followed by PhOH/$CHCl_3$ extraction and EtOH precipitation. Cleaved RNA was resuspended in formamide loading dye, heat denatured, and analyzed by denaturing 10.5% polyacrylamide gel electrophoresis.

(24) Thieriet, N.; Alsina, J.; Giralt, E.; Guibé, F.; Albericio, F. *Tetrahedron Lett.* **1997**, *38*, 7275−7278.
(25) Park, C.; Burgess, K. *J. Comb. Chem.* **2001**, *3*, 257−266.
(26) Mathews, D. H.; Sabrina, J.; Zuker, M.; Turner, D. H. *J. Mol. Biol.* **1999**, *288*, 911−940.
(27) Zuker, M.; Mathews, D. H.; Turner, D. H. *Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide in RNA Biochemistry and Biotechnology*; Kluwer Academic Publishers: Dordrecht, 1999.

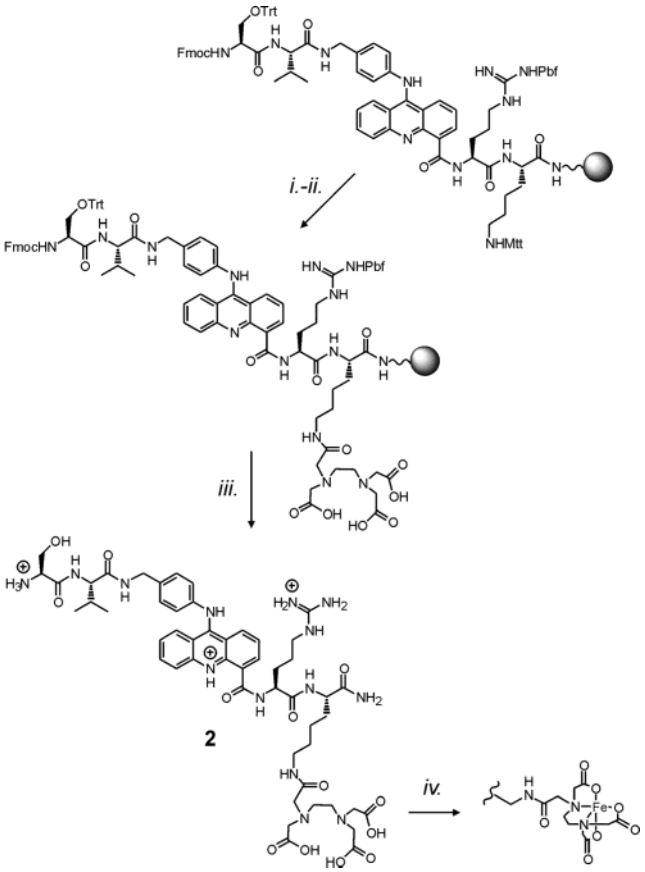
Quantitative Ribonuclease Footprinting. Footprints for **1** and **2** on the RNA stem-loops *4nt*, *1nt*, and pre-miRNA 39 were obtained using RNase V1 under native conditions. PIC−RNA complexes were formed by incubating increasing concentrations of **1** and **2** with ~5 nM end-labeled RNA for 15 min in buffer (50 mM Bis-Tris·HCl, 100 mM NaCl, and 10 mM $MgCl_2$, pH 7.0) and 10 μg/mL of yeast $tRNA^{phe}$. Enzymatic digestions with RNase V1 (0.0001 U/μL) were carried out for 30 min and quenched with hot formamide loading dye. Cleaved RNA was heat denatured and analyzed by 10.5% denaturing polyacrylamide gel electrophoresis. The cleavage efficiency at nucleotide(s) near the binding site was calculated by normalizing for the differential loading for each concentration of the peptide tested. For *4nt*, the footprint at A46 cleaved by RNase V1 was monitored with respect to the V1-dependent constant band U54. For *1nt* and pre-miRNA 39, the footprint at G30-U31 was followed with respect to the V1-dependent constant band G18. The cleavage data for the RNA was converted to binding data for the ligand, assuming that the maximum cleavage efficiency corresponds to 0% occupancy and the minimum cleavage efficiency corresponds to 100% occupancy by the ligand. The fraction of RNA bound by the ligand was plotted as a function of concentration, and the data were fitted to the equation: fraction bound = [ligand]/([ligand] + K_d). The results are reported as the average and standard deviation for three different experiments.

Results

To determine the groove location of the N- and C-termini of a PIC, we developed solid-phase synthesis approaches to variants with EDTA covalently attached to either the N- or the C-terminal amino acid residue. Freely diffusing, nonspecific hydroxyl radicals, generated by EDTA·Fe, abstract hydrogen atoms from the proximal ribose or deoxyribose residues producing carbon-based sugar radicals that can react further with molecular oxygen, ultimately leading to strand scission.[28] The cleavage patterns generated are diagnostic for minor or major groove localization of the EDTA·Fe.[29−32] We have previously published the synthesis of the N-terminally modified PIC **1**.[16] For modification with EDTA at the PIC C-terminus (**2**), we introduced a C-terminal lysine residue protected at the ϵ-amino group with methyltrityl (Mtt) (Scheme 1). Following standard solid-phase synthesis and coupling of the N-terminal serine residue, the solid support was treated with 1% TFA to remove Mtt, while leaving other acid-labile protecting groups intact.[25] The resulting resin-bound compound was allowed to react with a solution of EDTA monoanhydride in DMF overnight. Following piperidine treatment to remove the N-terminal Fmoc group, the C-terminal EDTA-modified PIC was cleaved from the support with simultaneous side-chain deprotection using 88% TFA. The final product was purified to homogeneity by HPLC and the predicted $[M + H]^+$ ion was observed by electrospray ionization mass spectrometry.

EDTA·Fe **2** was used to cleave a 75 nucleotide in vitro selected RNA with a high-affinity intercalation site for N-Ser-Val-Acr-Arg-C.[16] The resulting cleavage pattern was compared to that of the same RNA allowed to react with EDTA·Fe **1**. With EDTA·Fe at the N-terminus, PIC **1** efficiently cleaves nucleotides 28−33 and 43−51 (Figure 3A). As expected, these nucleotides surround the base-paired 5′-CpG-3′ step in the

Scheme 1. Solid-Phase Synthesis of SVAcrRK−EDTA[a]



[a] Reagents and conditions: (i) 1% TFA and 5% TIS in DCM, rt, 30 min (×2); (ii) EDTA monoanhydride in DMF, rt, 15 h; (iii) TFA/TIS/PhOH/water (88:5:5:2), rt, 5 h; (iv) $Fe(NH_4)_2(SO_4)_2$ (provided as an aqueous solution immediately preceding hydroxyl radical cleavage).

structure, an established intercalation site for acridine. Furthermore, there is a subtle 3′-asymmetrical shift to the pattern when mapped onto the predicted secondary structure (highlighted blue in Figure 3B), consistent with a minor groove-localized EDTA·Fe.[29,32] The C-terminally modified compound (**2**) produces a significantly different cleavage pattern in which nucleotides 25−27 and 38−42 are most efficiently cleaved (Figure 3A). These nucleotides flank the proposed intercalation site and are 5′-asymmetrically shifted (highlighted red in Figure 3B), characteristic of a major groove-localized EDTA·Fe.[30,31]

When presented in three-dimensions, the locations of nucleotides cleaved by the modified PICs on this duplex RNA structure become more apparent. A molecular model of the PIC−RNA complex (Figure 4) was generated using A-form geometry for the base-paired nucleotides and arbitrarily chosen conformations for the loop and bulged nucleotides. The intercalation domain of the PIC was docked into the 5′-CpG-3′ step as observed in a crystal structure of the dinucleotide bound to 9-aminoacridine.[21] The nucleotides cleaved by the N- and C-terminally EDTA·Fe-modified PICs are highlighted in blue and red, respectively. According to the model, nucleotides cleaved with the two different reagents reside on opposite sides of the duplex. This supports a threading intercalation binding mode with the PIC N-terminus in the minor groove and C-terminus in the major groove of this in vitro-selected RNA.

To better understand how the non-Watson−Crick secondary structure flanking the intercalation site contributes to PIC

(28) Pogozelski, W. K.; Tullius, T. D. *Chem. Rev.* **1998**, *98*, 1089−1107.
(29) Taylor, J. S.; Schultz, P. G.; Dervan, P. B. *Tetrahedron* **1984**, *40*, 457−465.
(30) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645−650.
(31) Han, H.; Dervan, P. B. *Nucleic Acids Res.* **1994**, *22*, 2837−2844.
(32) Dervan, P. B. *Science* **1986**, *232*, 464−471.

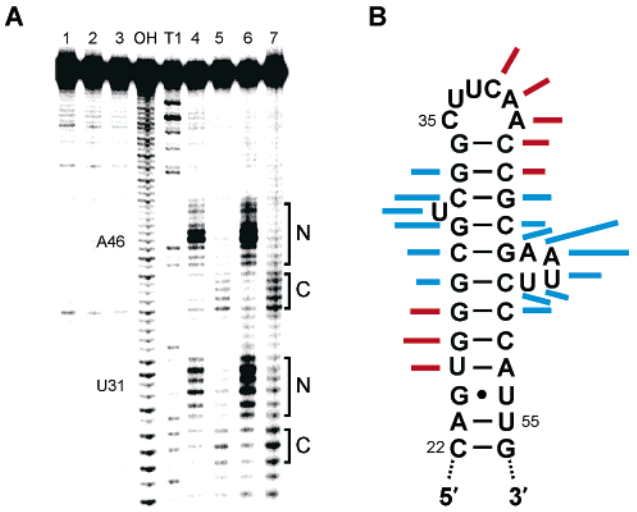


Figure 3. (A) Storage phosphor autoradiogram of a 10.5% denaturing polyacrylamide gel separating 5′-^{32}P-labeled RNA cleavage products. Lane 1 contains EDTA (50 μM) and $Fe(NH_4)_2(SO_4)_2$ (100 μM), while lane 2 contains SVAcrR (50 μM), EDTA (50 μM), and $Fe(NH_4)_2(SO_4)_2$ (100 μM). Lanes 3, OH, and T1 are RNA in buffer alone, alkaline hydrolysis, and denaturing RNase T1, respectively. Lanes 4 and 6 contain EDTA·Fe−SVAcrR (**1**) at 10 and 50 μM, respectively. Lanes 5 and 7 contain SVAcrRK−EDTA·Fe (**2**) at 10 and 50 μM, respectively. Bracketed nucleotides are labeled according to the PIC terminus that was modified with EDTA·Fe. (B) Partial secondary structure of the in vitro-selected RNA showing location and efficiency of hydroxyl radical cleavage by EDTA·Fe **1** (blue) and **2** (red).

binding, additional RNA stem-loops were evaluated using directed hydroxyl radical cleavage. Nucleotides in the asymmetric internal loop (46−49) were systematically deleted from the 3′ end, resulting in four additional RNA stem-loops containing three (*3nt*), two (*2nt*), or one (*1nt*) nucleotide(s) in the internal loop, as well as one with the entire loop removed (*del*) (Figure 5). These loop-altered RNA structures were allowed to react with EDTA·Fe **1**, and their respective cleavage patterns were compared to that of the original RNA stem-loop (*4nt*) cleaved with this N-terminal EDTA·Fe modified compound (Figure 5). Cleaved regions concentrated about U31 and A46 are maintained throughout the series of RNA stem-loops, except for the absence of any significant cleavage on *del*. The persistence of the 3′-asymmetric cleavage pattern suggests that reducing the size of the internal loop from four nucleotides to a single nucleotide does not affect the location of the N-terminus of **1** when bound to these RNA stem-loops; the PIC remains intercalated with its N-terminus located in the minor groove. Importantly, the affinity with which this compound binds the RNA is affected by the loop nucleotide deletions. The dissociation constants of **1** for *4nt* and *1nt* were measured using quantitative V1 ribonuclease footprinting, and an 18-fold loss of affinity is observed when three nucleotides are removed from the loop (Table 1). It should be noted that a loop consisting of a single nucleotide is a minimum requirement for **1** to bind this RNA, given that no binding is observed when the loop is removed entirely.

When the altered RNA stem-loops react with the C-terminally modified EDTA·Fe **2**, efficiently cleaved nucleotides are observed for all RNA structures except *del* (data not shown), suggesting that a single nucleotide in the loop is also minimally required for **2** to bind. However, unlike the constant pattern observed in the RNA series cleaved with EDTA·Fe **1**, an

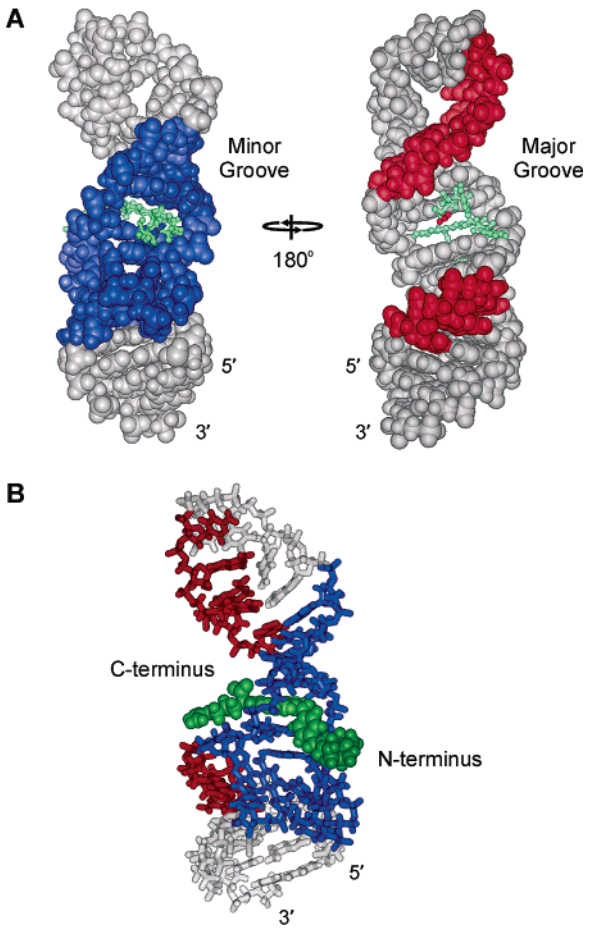


Figure 4. Molecular model of a PIC−RNA complex displaying the location of nucleotides cleaved by **1** (blue) and **2** (red). (**A**) Two orientations of PIC SVAcrRK bound to the in vitro-selected RNA viewed head-on to the N- or C-terminus of the PIC highlighted in blue and red, respectively. (**B**) A side-on view of a model of SVAcrR bound to the RNA combining the hydroxyl radical cleavage data from the two orientations above.

interesting phenomenon is apparent when the *3nt* and *2nt* loops are compared (Figure 6). The most efficiently cleaved nucleotides in *3nt* are identical to those cleaved in *4nt* and display the same characteristic 5′ shift. But *2nt* is bound by **2** with noticeably less selectivity as demonstrated by the much broader range of nucleotides cleaved. The less selective pattern is also seen on stem-loop *1nt* (data not shown), which binds **2** with a 13-fold decrease in affinity when compared to *4nt* (Table 1). A more careful examination reveals that the expanded cleavage pattern is a combination of the two patterns already established: nucleotides 28−33 and 43−51 cleaved by minor groove-localized EDTA·Fe and nucleotides 25−27 and 38−42 cleaved by major groove-localized EDTA·Fe. Thus, the broad cleavage pattern observed on *2nt* is most likely a consequence of a mixture of PIC−RNA threading intercalation complexes with two different polarities. Binding is partitioned between the previously observed mode in which the C-terminus of **2** is situated in the major groove of the RNA and the newly observed mode that places the C-terminus in the minor groove.

Results of the cleavage experiments with the altered RNA stem-loops have defined the minimum structural requirements of a binding site for this PIC. By deleting the non-Watson−Crick secondary structure on either side of the intercalation site, either the single uridine bulge (U31, data not shown) or the asymmetric internal loop (nt 46−49), binding was significantly

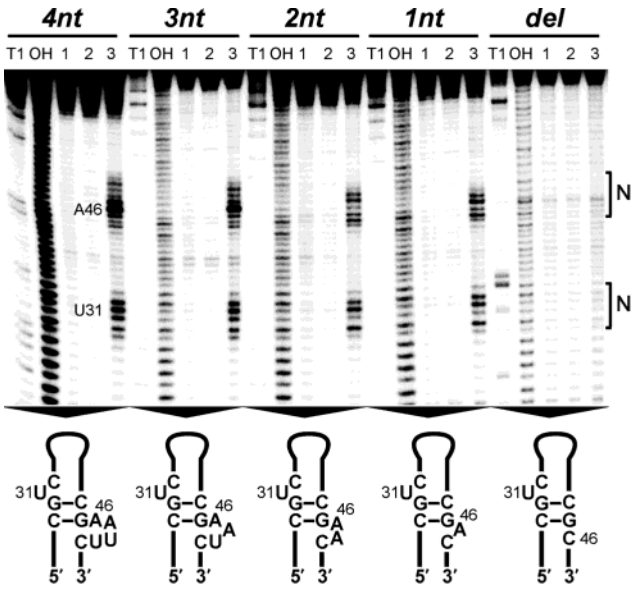


Figure 5. Storage phosphor autoradiogram of a 10.5% denaturing polyacrylamide gel separating 5′-^{32}P-labeled RNA cleavage products of the RNA stem-loops *4nt*, *3nt*, *2nt*, *1nt*, and *del* for EDTA·Fe **1**. For each RNA, lanes T1 and OH are denaturing RNase T1 digestion and alkaline hydrolysis of the RNA, respectively. Lanes 1 contain RNA incubated with 50 μM EDTA, 100 μM $Fe(NH_4)_2(SO_4)_2$, and 50 μM SVAcrR. Lanes 2 contain RNA in reaction buffer alone. Lanes 3 contain RNA incubated with 10 μM EDTA·Fe−SVAcrR (**1**). Bracketed nucleotides are labeled according to the PIC terminus that was modified with EDTA·Fe.

Table 1. Effect of C-Terminal Modification and RNA Internal Loop Structure on PIC Binding

PIC	K_d (μM) for *4nt*[a]	K_d (μM) for *1nt*[a]
EDTA−SVAcrR (**1**)	0.5 ± 0.2	8.8 ± 2.5
SVAcrRK−EDTA (**2**)	1.4 ± 0.5	18.8 ± 3.0

[a] Determined from fit of fraction bound = [ligand]/([ligand] + K_d) and reported as the average and standard deviation for three different experiments (see Materials and Methods for details).

reduced or completely abolished, respectively. Thus, binding to the RNA and the formation of a stable complex leading to hydroxyl radical cleavage requires at least a single-base bulge 3′ to the intercalation site on both strands. This observation led us to consider possible binding sites that fit these criteria in naturally occurring RNAs. Receiving much attention as of late, microRNAs (miRNAs) are naturally occurring single-stranded oligonucleotides (21−24 nucleotides in length) that regulate translation of target messenger RNAs by binding into the 3′-untranslated regions.[33] These miRNAs are generated by the processing of pre-miRNA hairpin stem structures, formed by pairing of the miRNA sequence with nearby sequences in the initial transcript. This pairing is imperfect and the duplex formed is invariably interrupted with bulges, mismatches, and loops. Such duplex RNA structures appear to be ideal candidates for binding by PICs. Indeed, the precursor for miRNA 39 from *C. elegans* has a structure remarkably similar to the binding site for **1** found on the *1nt* RNA, with a base-paired 5′-CpG-3′ step flanked by uridine bulges on the 3′-side on each strand (Figure 7A).[34] To determine if this structure is indeed a binding site for the PIC, we transcribed and evaluated a new stem-loop

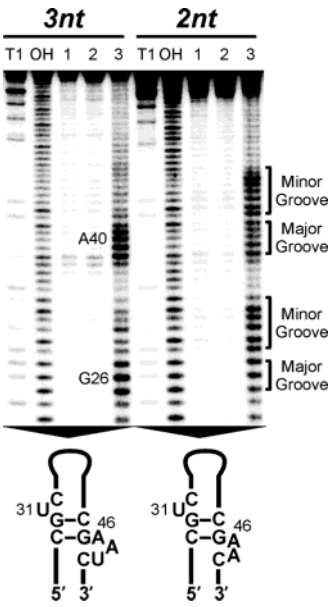


Figure 6. Storage phosphor autoradiogram of a 10.5% denaturing polyacrylamide gel separating 5′-^{32}P-labeled RNA cleavage products of RNA stem-loops *3nt* and *2nt* for EDTA·Fe **2**. For each RNA, lanes T1 and OH are denaturing RNase T1 digestion and alkaline hydrolysis of the RNA, respectively. Lanes 1 contain RNA incubated with 50 μM EDTA, 100 μM $Fe(NH_4)_2(SO_4)_2$, and 50 μM SVAcrR. Lanes 2 contain RNA in reaction buffer alone. Lanes 3 contain RNA incubated with 50 μM SVAcrRK−EDTA·Fe (**2**). The labels "minor groove" and "major groove" describe the RNA groove from which the bracketed nucleotides were cleaved.

containing the putative binding site from *C. elegans* pre-miRNA 39. As expected, EDTA·Fe **1** cleaved the pre-miRNA structure, yielding a pattern nearly identical to that observed with the *1nt* RNA stem-loop, consistent with a minor groove-localized N-terminus (Figure 7B). Analysis by quantitative V1 ribonuclease footprinting revealed a dissociation constant of 16.5 ± 6.9 μM for **1** binding to the pre-miRNA 39 construct (Supporting Information).

Discussion

Directed hydroxyl radical cleavage using covalently attached EDTA·Fe has allowed us to examine RNA recognition properties for threading intercalators of the peptide-intercalator conjugate type. Prior to this work, the groove location of the substituents of an acridine-based threading intercalator had not been clearly defined for any RNA target site. Early crystallographic studies of 9-aminoacridine bound to the RNA self-complementary dinucleotide 5′-CpG-3′ showed the acridine in two distinct intercalation complexes in the crystal.[21] These two binding modes differed in the groove location of the 9-amino group, suggesting that an acridine 9-substituent may occupy either the major or minor groove in duplex RNA. Studies of DNA complexes also revealed uncertainties in the orientation of 9-anilino-4-carboxamide acridines. Early kinetic experiments were interpreted to suggest a binding mode with the 9-anilino substituent positioned in the major groove and the 4-carboxamide side chain residing in the minor groove.[35] More recently, DNA cleavage patterns observed when footprinting with an iron-binding chelidamic acid derivative of 9-anilinoacridine-4-carboxamide suggested minor groove localization of the aniline

(33) Ke, X.-S.; Liu, C.-M.; Liu, D.-P.; Liang, C.-C. *Curr. Opin. Chem. Biol.* **2003**, *7*, 516−523.
(34) Lau, N. C.; Lim, L. P.; Weinstein, E. G.; Bartel, D. P. *Science* **2001**, *294*, 858−862.
(35) Wakelin, L. P. G.; Chetcuti, P.; Denny, W. A. *J. Med. Chem.* **1990**, *33*, 2039−2044.

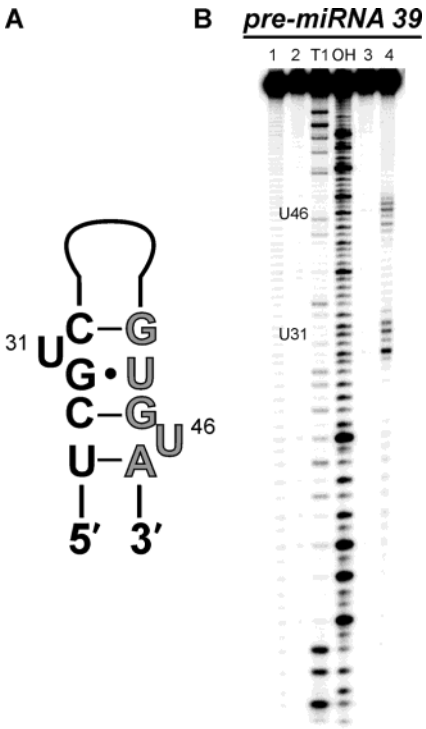


Figure 7. (A) Predicted partial secondary structure of an RNA containing a PIC binding site found in the pre-miRNA 39 from *C. elegans*.[30] The sequence colored in gray is found in the mature miRNA. (B) Storage phosphor autoradiogram of a 10.5% denaturing polyacrylamide gel separating 5′-^{32}P-labeled RNA cleavage products of the pre-miRNA 39 for EDTA·Fe **1**. Lanes T1 and OH are denaturing RNase T1 digestion and alkaline hydrolysis of the RNA, respectively. Lanes 1 and 2 contain RNA in reaction buffer alone and RNA incubated with 50 μM EDTA, 100 μM $Fe(NH_4)_2(SO_4)_2$, and 50 μM SVAcrR, respectively. Lanes 3 and 4 contain RNA incubated with 1 and 10 μM EDTA·Fe−SVAcrR (**1**), respectively.

ring of this compound.[36] Additionally, nitrogen mustard derivatives of the topoisomerase II poison and the clinical anticancer drug amsacrine were shown to have a slight preference for this same binding polarity.[37] In support of this DNA binding mode, X-ray crystal structures, available only for monosubstituted acridines, show the intercalator situated with 4-carboxamide substituents in the major groove.[38−42] However, this recently defined polarity is not absolute, given that a substituent at the 4-position has proven to be capable of dictating the reverse polarity.[43]

It appears the DNA groove locations observed for the acridine substituents on the chelidamic acid and the nitrogen mustard derivatives of amsacrine mentioned above are maintained for the disubstituted threading intercalator N-EDTA·Fe-Ser-Val-Acr-Arg-C (**1**) on the duplex RNA targets described here. This information is valuable for further optimization of the RNA binding affinity and selectivity of this class of compounds. For instance, with the knowledge that the PIC N-terminus lies in the minor groove, additional RNA-selective affinity may be realized by introducing functional groups designed to bind 2′-hydroxyls into this portion of the molecule.

The PICs examined herein bind with a defined polarity, dictated by the structure of both the small molecule and the RNA. When bound to an RNA stem-loop containing a binding site such as the one located in *4nt*, modification with EDTA·Fe at either the N- or C-terminus of a PIC does not prevent its preferred binding polarity of the N- and C-termini in the minor and major grooves, respectively. However, in an RNA stem-loop that contains only two nucleotides in the internal loop, such as *2nt*, an EDTA·Fe group attached to the C-terminus causes the PIC to reverse its polarity and place its C-terminus in the minor groove (Figure 8). From the dissociation constants enumerated in Table 1, it is clear that modifying the C-terminus of the PIC with EDTA·Fe is destabilizing to the PIC−RNA complex. In A-form RNA, the major groove is narrower than the minor groove, creating an environment that may not tolerate an EDTA·Fe moiety as well as the minor groove. Removing nucleotides from the internal loop also leads to a destabilized complex. This may result from a loss of specific RNA contacts to the deleted nucleotides or from increased strain of a conformationally restricted binding site. When both destabilizing characteristics are present in the complex of **2** bound to *2nt*, the combination of a major groove-localized EDTA·Fe and deletion of nucleotides from the internal loop may be so destabilizing that a complex of equal stability can form with the modified C-terminus in the minor groove. It is also possible that the polarity reversal observed for **2** binding to *2nt* is related to the threading mechanism. That is, a bulky EDTA·Fe substituent slows threading of the C-terminus through a conformationally restricted binding site. However, this seems unlikely since incubating EDTA·Fe **2** with *2nt* overnight at room temperature or for 15 min at 55 °C immediately preceding hydroxyl radical cleavage did not change the cleavage pattern (data not shown). Future experiments will be required to delineate the exact origin of the polarity reversal phenomenon.

A significant result of the directed hydroxyl radical cleavage experiments was the determination of minimum structure requirements of a binding site for the evaluated PICs. It became apparent from these studies that a productive complex leading to hydroxyl radical cleavage of the RNA requires at least a single-base bulge 3′ to the intercalation site on both strands. This differs from earlier studies conducted by White and Draper, in which it was noted that the classical intercalators methidium propyl EDTA·Fe and ethidium bromide preferentially bound duplex RNA adjacent to a single nucleotide bulge on one strand.[44] These differing results may be ascribed to the nature of the intercalators studied. While a single-base bulge is adequate for a classical intercalator to effectively insert between base pairs in double-stranded RNA, a threading intercalator apparently requires an additional bulge adjacent to the intercalation site to allow a substituent to pass through the base pairs en route to the opposite side of the duplex. It is important to note, however, that even among binding sites that have a CpG step flanked by a non-Watson−Crick structure, preferential binding is observed. The *4nt* site is bound by PIC **1** with an approximately 30-fold higher affinity than the pre-miRNA 39 site. Thus, the RNA

(36) Searcey, M.; McClean, S.; Madden, B.; McGown, A. T.; Wakelin, L. P. G. *Anti-Cancer Drug Des.* **1998**, *13*, 837−855.
(37) Fan, J.-Y.; Ohms, S. J.; Boyd, M.; Denny, W. A. *Chem. Res. Toxicol.* **1999**, *12*, 1166−1172.
(38) Adams, A.; Guss, J. M.; Collyer, C. A.; Denny, W. A.; Wakelin, L. P. G. *Biochemistry* **1999**, *38*, 9221−9233.
(39) Adams, A.; Guss, J. M.; Collyer, C. A.; Denny, W. A.; Prakash, A. S.; Wakelin, L. P. G. *Mol. Pharmacol.* **2000**, *58*, 649−658.
(40) Adams, A. *Curr. Med. Chem.* **2002**, *9*, 1667−1675.
(41) Adams, A.; Guss, J. M.; Denny, W. A.; Wakelin, L. P. G. *Nucleic Acids Res.* **2002**, *30*, 719−725.
(42) Todd, A. K.; Adams, A.; Thorpe, J. H.; Denny, W. A.; Wakelin, L. P. G.; Cardin, C. J. *J. Med. Chem.* **1999**, *42*, 536−540.
(43) Bourdouxhe-Housiaux, C.; Colson, P.; Houssier, C.; Waring, M. J.; Bailly, C. *Biochemistry* **1996**, *35*, 4251−4264.
(44) White, S. A.; Draper, D. E. *Biochemistry* **1989**, *28*, 1892−1897.

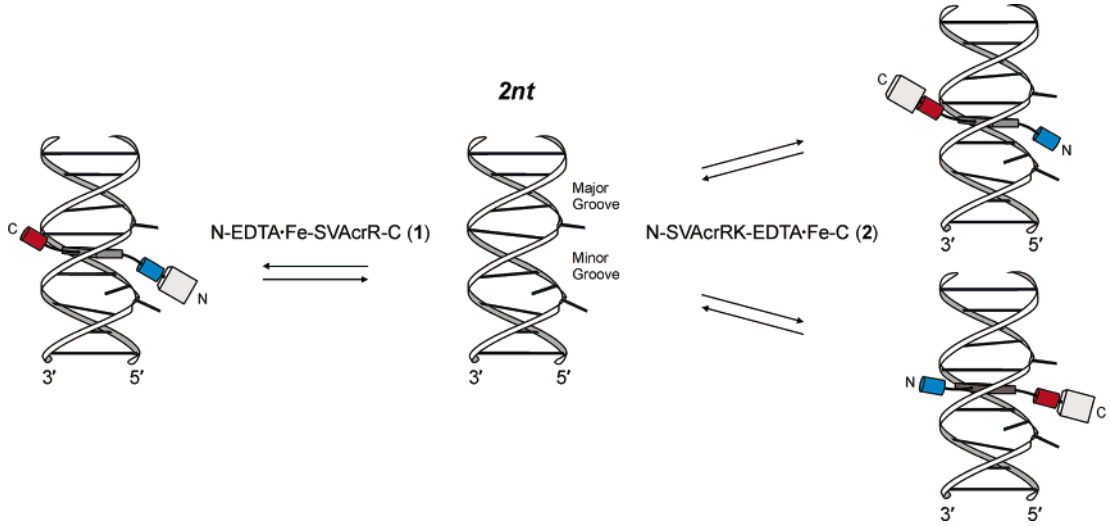


Figure 8. Schematic representation of the binding polarities of PICs when modified at either the N- or C-terminus with EDTA·Fe and bound to *2nt*.

structure flanking the intercalation site plays a currently undefined role in selective binding by these compounds.

Information on minimal binding site requirements is valuable when choosing other RNA targets for PICs. By minimizing the structural complexity of the binding site in the *4nt* RNA to that found in *1nt* while maintaining reasonable affinity for a PIC, the probability of identifying biologically relevant PIC-binding RNAs becomes greater. Indeed, we have shown that PIC **1** binds a duplex RNA structure found in a pre-miRNA from *C. elegans* with identical polarity and similar affinity as when bound to the *1nt* RNA stem-loop. Once binding is observed on such an RNA, one can envision optimizing affinity and specificity through the synthesis of new PICs with structurally altered groove-localized substituents[45] or intercalation domains.[19] Specifically, by reducing the stacking potential of the intercalator while increasing the binding contribution of the peptide substituents, we may obtain greater specificity. As we have previously demonstrated, changing acridine to a 2-phenylquinoline within the context of the PIC N-Ser-Val-PheQ-Arg-C results in a 10-fold loss of affinity for the same in vitro-selected RNA studied in this publication (*4nt*).[19] However, specificity for the high-affinity binding site is maintained. The loss of affinity could be recovered, or perhaps surpassed, by optimizing or creating RNA contacts to the peptide substituents. A PIC−RNA complex more reliant on specific contacts and less on intercalation would show greater discrimination among duplex RNAs with similar structures as pre-miRNA 39. In the case of pre-miRNAs, a high-affinity ligand discovered this way could be used to control the processing steps necessary to generate the mature miRNA, thus controlling translational regulation of the target message.

In summary, we have shown that a peptide intercalator conjugate can bind duplex RNA structures via threading intercalation with a preferred polarity placing its N-terminus in the minor groove and its C-terminus in the major groove. However, the binding polarity preference is apparently lost with a sterically demanding C-terminal EDTA·Fe modification and a conformationally restricted RNA binding site. Deletion of nucleotides in an internal loop found in a high-affinity binding site for a PIC helped define the minimal structural requirements for binding and led to the identification of a PIC-binding site in a pre-miRNA. These results will guide both the rational design and library screening approaches to the discovery of high-affinity and highly selective ligands for naturally occurring duplex RNAs.

Acknowledgment. P.A.B. acknowledges support from the NIH (AI-49062) and a Camille Dreyfus Teacher Scholar Award. We also acknowledge Cabe Clark for assistance with the generation of loop altered RNAs.

Supporting Information Available: Average plots of fraction bound versus concentration for **1** and **2** with RNA stem-loops *4nt* and *1nt* and pre-miRNA 39; storage phosphor autoradiograms of polyacrylamide gels showing footprints of **1** and **2** on *4nt* and *1nt* and pre-miRNA 39 (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA047818V

(45) Carlson, C. B.; Spanggord, R. J.; Beal, P. A. *ChemBioChem* **2002**, *3*, 859−865.