Published in final edited form as:

J Inorg Biochem. 2017 May ; 170: 202–208. doi:10.1016/j.jinorgbio.2017.02.024.

Cobalt-based paramagnetic probe to study RNA-protein interactions by NMR

Leah M. Seebald, Christopher M. DeMott, Srivathsan Ranganathan, Papa Nii Asare-Okai, Anastasia Glazunova, Alan Chen, Alexander Shekhtman, and Maksim Royzen*

Department of Chemistry, University at Albany, SUNY, 1400 Washington Ave. Albany, NY 12222, United States

Abstract

Paramagnetic resonance enhancement (PRE) is an NMR technique that allows studying three-dimensional structures of RNA-protein complexes in solution. RNA strands are typically spin labeled using nitroxide reagents, which provide minimal perturbation to the native structure. The current work describes an alternative approach, which is based on a Co^{2+} -based probe that can be covalently attached to RNA in the vicinity of the protein's binding site using 'click' chemistry. Similar to nitroxide spin labels, the transition metal based probe is capable of attenuating NMR signal intensities from protein residues localized $<40 \text{ \AA}$ away. The extent of attenuation is related to the probe's distance, thus allowing for construction of the protein's contact surface map. This new paradigm has been applied to study binding of HIV-1 nucleocapsid protein 7, NCp7, to a model RNA pentanucleotide.

Keywords

HSQC; RNA-protein interactions; Paramagnetic NMR; NCp7; HIV-1; Cobalt

1. Introduction

There is a strong interest in the biochemical community to develop a better fundamental understanding of how proteins interact with RNA [1]. RNA-protein interactions are ubiquitous in biology and critical at many regulatory steps of gene expression and stages of cell development [2,3]. X-ray crystallography has been the main tool in structural biology having paved way towards understanding a number of key ribonucleoprotein assemblies, such as ribosomes [4–6]. Nuclear magnetic resonance (NMR) spectroscopy is a complementary technique that, in addition to structure determination, allows studying dynamics interactions between biomolecules in solution and inside the live cells [7,8]. NMR is especially useful when the structural complexity precludes crystallization of a ribonucleoprotein complex.

The classical two-dimensional ^1H and ^{15}N correlation experiments (i.e. heteronuclear single quantum coherence (^{15}N -HSQC)) correlate the ^1H and ^{15}N nuclei, giving rise to cross peaks

*Corresponding author. mroyzen@albany.edu (M. Royzen).

reflective of labeled protein's environment [8]. The resonance frequencies of individual residues are determined by the amino acid specific chemical environments and reflect the overall 3-D conformation of an isotope labeled protein [9]. These two-dimensional NMR techniques have been used to study interactions between proteins and RNA aptamers [10–12]. Short distance atomic coupling, $<6 \text{ \AA}$, can be quantified to determine close range interactions in these compact structures. ^{15}N -HSQC however does not provide long range structural information, which is important for larger or more flexible RNA-protein complexes.

Paramagnetic relaxation enhancement (PRE) is an NMR technique which can provide valuable structural information about RNA-protein complexes relying on distance constraints up to $<40 \text{ \AA}$ away [13,14]. Extrinsic paramagnetic centers, incorporated onto RNA oligonucleotides via a linker, are capable of attenuating HSQC cross-peak signal intensities of the protein's backbone once the latter is attached to RNA. The extent of attenuation is related to the proximity of the protein's domains to the paramagnetic center, thus allowing construction of molecular maps describing protein's interactive surface. RNA has typically been modified with nitroxide spin labels due to their small size which minimally perturbs the native RNA structure [15]. The technique termed site-directed spin labeling (SDSL) has also allowed studying RNA's dynamic behavior in solution [16,17]. Nitroxide labels provide strong PRE effects due to slow electron relaxation, and NMR experiments which require these stronger PRE effects benefit from these types of paramagnetic tags. Clore and coworkers used a nitroxide spin-label to detect a rare conformational state in apo maltose-binding protein undergoing chemical exchange [18]. Qin and coworkers use SDSL to study dynamics of the substrate recognition duplex within the Tetrahymena ribozyme [19]. Additional reported nitroxide labels have provided important structural data related to RNA-protein binding in solution [20,21].

Paramagnetic metal ions can also be used to study long range molecular interactions with the benefits of additional versatility both through the choice of metal ion and ligand design. The latter can be engineered for site-specific incorporation into protein's metal binding domain [22]. Two-armed ligands can be employed for stronger and more rigid protein binding [23]. Analogous to nitroxide spin labels, paramagnetic centers with isotropic electron density, such as Mn^{2+} , and Gd^{3+} , can produce PRE effects which arise from slow electron relaxation that occurs during NMR experiments [24]. On the other hand, paramagnetic metal centers, such as high spin Co^{2+} , Yb^{3+} , and Pr^{3+} , have anisotropic electron density which, in addition to PRE, can give structural information through pseudocontact shifts (PCS), and residual dipolar coupling (RDC).

The metal ion-based paramagnetic probes have been explored to study protein structure and dynamics. For example, a paramagnetic Cu^{2+} cyclen tag has been used to differentiate between two conformational states of human Ras protein [25]. Meanwhile, Ni^{2+} was utilized by Led and co-workers to obtain the structural information of thioredoxin [26,27]. Otting and co-workers genetically encoded Co^{2+} -binding amino acid, bipyridylalanine, into West Nile virus NS2B-NS3 protease [22]. Work described herein is the first example where a paramagnetic transition metal-based probe is applied towards understanding binding interactions between RNA and proteins by NMR.

Inspired by the work of Graham [28], the Co^{2+} -based paramagnetic NMR probe, shown in Scheme 2, has been designed for facile attachment to RNA strands using ‘click’ chemistry. The placement of the probe on the RNA has been strategically chosen to be in proximity to the protein’s binding site and yet to be minimally disruptive of the RNA-protein interaction. Short linker reduces the probe’s rotational flexibility thus simplifying interpretation of the NMR data [29]. Herein, we demonstrate that the Co^{2+} -based probe is suitable for studying RNA-protein interactions by PRE NMR.

2. Experimental

2.1. Materials and instrumentation

All chemicals were received from commercial sources and used without further purification. Chromatographic purifications were conducted using SiliaSphere™ spherical silica gel 5 μm , 60 Å silica gel (Silicycle). Thin layer chromatography (TLC) was performed on SiliaPlate™ silica gel TLC plates (250 μm thickness) purchased from Silicycle. Preparative TLC was performed on SiliaPlate™ silica gel TLC plates (1000 μm thickness). HPLC purification was performed using Phenomenex Luna 5 μ C18(2) semi-preparative column (250 $\times$ 10 mm). ^1H and ^{13}C NMR spectroscopy was performed on a Bruker NMR at 400 (^1H), and 100 (^{13}C) MHz.

The RNA oligonucleotide synthesis was carried out on a 1.0 μmol scale using MerMade 8 DNA synthesizer. All the natural nucleoside phosphoramidites (TBDMS as the 2′-OH protecting group) and the accessory reagents were purchased from ChemGenes. After synthesis, the unmodified RNA oligo was cleaved from the beads and deprotected by the treatment with AMA solution (a 1:1 aqueous solution of methyl-amine and concentrated ammonium hydroxide) at 65 °C for 2 h. After evaporating the resulting solution to dryness, the 2′-TBDMS deprotection was performed using NEt_3 -HF solution at 65 °C for 2.5 h. The RNA was finally precipitated by using ammonium acetate and butanol, and re-dissolved in water.

The samples were analyzed on a Thermo Fisher Scientific (West Palm Beach, CA) LTQ Orbitrap Velos Mass spectrometer, using quartz capillary emitters. To facilitate spray optimization, 10% isopropyl alcohol was added to each sample prior to MS analysis. Cyclam and cyclen derivative were dissolved in MeOH and analyzed in the positive mode and the RNA was analyzed in a solution containing 150 mM ammonium acetate in the negative mode.

All 2-D NMR experiments were performed at RT on a 500 MHz Bruker Avance III NMR spectrometer equipped with the TCI cryoprobe. To prepare the NMR sample, 0.15 mM of [U - ^{15}N] HIV- nucleocapsid protein 7, NCp7, was dissolved in the NMR buffer 10 mM potassium phosphate, pH 6.5, and 10% D_2O . **RNA1** or **RNA2** were added to the molar ratio of 1:1 between the NCp7 and RNA. ^1H - ^{15}N HSQC with Watergate water suppression was used to monitor protein chemical shifts changes due to RNA binding [30]. 1024 $\times$ 256 points in proton and nitrogen dimensions, respectively, were acquired with 128 transients. The spectra were processed by using Topspin 2.1 (Bruker Inc) and analyzed by using Cara software [31]. The peaks were assigned based on published results [32]. The peak volumes

were normalized by dividing the peak volume of each residue by the peak volume of E42 for NCp7-RNA1 complex.

The peaks intensity changes were quantified by using the ratio between the peak volumes:

$$\frac{\text{normalized peak volume of residue from RNA1}}{\text{normalized peak volume of residue from RNA2}} \\ = \text{Ratio of Peak Volume Change}$$

2.2. Synthesis of probe

2.2.1. Synthesis of 5—A solution of Boc anhydride (2.95 g, 13.50 mmol) in anhydrous CH₂Cl₂ (60 mL) was added dropwise over 1 h to the solution of cyclam (1.50 g, 7.48 mmol) and DIPEA (6.50 mL, 37.3 mmol), in CH₂Cl₂ (250 mL) under nitrogen atmosphere. After addition, the mixture was cooled to −15 °C and a second solution of Boc anhydride (1.96 g, 8.98 mmol) in CH₂Cl₂ (60 mL) was added dropwise over 1 h. The solution was slowly warmed up to rt and stirred for 18 h. The solution was then washed with aqueous 0.5 M K₂CO₃ (2 × 150 mL), dried with Na₂SO₄ and concentrated under reduced pressure. The title product was obtained as a white foam by flash chromatography using a gradient of 0–100% EtOAc in hexanes, followed by 1–10% MeOH in EtOAc. R_F = 0.32 in 1:9 MeOH:EtOAc with 0.1% Et₃N. Yield = 1.77 g (47.32%); ¹H NMR (CDCl₃, 400 MHz) δ 3.39–3.28 (m, 12H), 2.79 (t, *J* = 5.5 Hz, 2H), 2.62 (t, *J* = 5.5 Hz, 2H), 1.92 (br s, 2H), 1.70 (br s, 2H), 1.45 (s, 27H); HRMS (DART) *m/z*: calcd. for C₂₅H₄₉N₄O₆ [M + H]⁺ 501.3652; found 501.3661. The observed spectra were consistent with the ones reported previously [33].

2.2.2. Synthesis of 6—Sodium carbonate (1.50 g, 14.16 mmol) and propargyl bromide (~80% toluene, 268 μL, 3.54 mmol) were added to a solution of **5** (1.77 g, 3.54 mmol) in anhydrous acetonitrile (55 mL). The solution was stirred at reflux under nitrogen for 18 h. The insoluble salts were filtered and the solution was concentrated under reduced pressure. The title product was obtained as a white foam by flash chromatography using a gradient of 5% - 80% EtOAc in Hexanes. R_F = 0.58 in 7:3 EtOAc:Hexanes. Yield = 0.98 g (51.31%); ¹H NMR (CDCl₃, 400 MHz) δ 3.40–3.29 (m, 14H), 2.67 (br s, 2H), 2.51 (t, *J* = 5.5 Hz, 2H), 2.16 (s, 1H), 1.89 (br s, 2H), 1.69 (br s, 2H), 1.46 (s, 27H); HRMS (ESI) *m/z*: calcd. for C₂₈H₅₁N₄O₆ [M + H]⁺ 539.3809; found 539.3778. The observed spectra were consistent with the ones reported previously [33].

2.2.3. Synthesis of 7—Compound **6** (0.98 g, 1.82 mmol) was dissolved in a 1:1 mixture of TFA and CH₂Cl₂ (10 mL) and stirred for 1.5 h at room temperature. The solution was concentrated under reduced pressure and co-evaporated with MeOH (5 × 10 mL). The TFA salt was dissolved in a 1:9 mixture of MeOH:CH₂Cl₂ (50 mL) and washed with aqueous 1 M NaOH (150 mL), and brine (150 mL). The organic layer was dried with Na₂SO₄ and concentrated under reduced pressure. The title product was obtained as a light yellow solid.

Yield = 0.43 g (60.78%); ¹H NMR (CDCl₃, 400 MHz) δ 3.50–3.25 (m, 14H), 2.91 (t, *J* = 5.4 Hz, 2H), 2.84 (t, *J* = 5.5 Hz, 2H), 2.62 (s, 1H), 2.18 (t, *J* = 5.5 Hz, 2H), 1.93 (t, *J* = 5.5 Hz, 2H); ¹³C NMR (CD₃OH, 100 MHz) δ 76.67, 76.46, 51.15, 50.00, 46.02, 45.73, 45.63,

45.23, 44.70, 42.49, 39.07, 23.59, 23.41; HRMS (ESI) m/z : calcd. for $C_{11}H_{23}N_4$ $[M + H]^+$ 211.1923; found 211.1903.

2.2.4. Synthesis of 1— Co^{2+} chloride hexahydrate (59.7 mg, 0.252 mmol) was added to a solution of **7** (50.0 mg, 0.210 mmol) in EtOH (5 mL) and refluxed for 1.5 h. The metal complex was precipitated with ether and filtered. The crude product was dissolved in MeOH (1 mL). X-ray quality crystals were obtained by slow diffusion of ether. HRMS (ESI) m/z : calcd. for $C_{13}H_{26}ClCoN_4$ $[Co(\mathbf{12})]^{3+}$ 166.0584; found 166.0575. IR (neat) 3159, 2982.05, 2937.76, 2515.59, 2030.18, 2105.99, 1977.50, 1610.60, 1454.87, 1422.51, 1383.28, 1253.78, 1240.95, 1136.13, 1107.25, 1092.25, 1065.06. 1H NMR (D_2O , 400 MHz) δ 3.73 (t, $J = 6.8$ Hz, 2H), 3.58 (t, $J = 6.9$ Hz, 4H), 2.98–2.89 (m, 4H), 2.77 (t, $J = 6.8$ Hz, 3H), 2.53 (t, $J = 6.9$ Hz, 4H) 2.44–2.35 (m, 1H), 2.03–1.66 (m, 2H). *Anal. Calcd.* for $C_{13}H_{28}Cl_5CoN_4$ (**M** + 2HCl) C, 32.76; H, 5.92; N, 11.76. Found: C, 32.43; H, 6.06; N, 11.33.

2.3. X-ray crystallography

2.3.1. Crystals of 1 (CCDC deposition number: 1438951)—Data collection was performed on a Bruker D8 VENTURE X-ray diffractometer with PHOTON 100 CMOS detector and mirror-monochromated Cu-K α radiation ($\lambda = 1.54178$ Å) at $T = 100(2)$ K. Data was corrected for absorption effects using the empirical method SADABS. The structures were solved with the SHELXTL-97 software package by direct methods and refined against F^2 [34–36]. In general, all non-hydrogen atoms were located on the difference-Fourier map and refined anisotropically. Hydrogen atoms were placed in idealized locations and given isotropic thermal parameters equivalent to either 1.5 (terminal CH_3 or NH_3 hydrogen atoms) or 1.2 times the thermal parameter of the atom to which they were attached. At the final stages of refinement, the structures were checked for higher symmetry using PLATON [37].

All non-hydrogen atoms were refined anisotropically. Hydrogen atoms of two sumanenyl anions without disorder were located in the difference-Fourier map and refined with the C—H distances constrained to 1.01 ± 0.02 Å for CH and 0.95 ± 0.02 Å for CH_2 . All hydrogen atoms of the disordered sumanenyl dianion were included in idealized positions for structure factor calculations with $U_{iso}(H) = 1.2 U_{eq}(C)$. This anion and K7 are disordered over two orientations that were resolved and treated with an occupancy ratio of 0.66:0.34. Both orientations were refined with anisotropic thermal parameters but all corresponding atoms' anisotropic thermal parameters were constrained to be the same.

2.4. Synthesis and preparation of modified RNA

Compound **2** was synthesized and immobilized on GE Healthcare Custom Primer Support™ Amino resin (GE Healthcare cat#17-5214-98) by following the previously reported procedure [38]. The functionalized resin was packed into empty Bioautomation MerMade columns (MM-1000-1) (4.00 mg resin per column). The RNA oligonucleotide synthesis was performed as described above. After synthesis, the RNA was cleaved from the resin as previously described [38]. After evaporating the resulting solution to dryness, the RNA was resuspended in DMSO (anhydrous, 100 μ L). The 2'-TBDMS deprotection was performed using $NEt_3 \cdot HF$ (1:1, 125 μ L) solution at 65 °C for 2.5 h. The RNA was cooled briefly and

precipitated with 3 M NaOAc (25 μ L), vortexed for 15 s, followed by butanol (1 mL) vortexed for 30 s. The solution was cooled at -80°C for 30 min, then centrifuged for 15 min at 12500 rpm. The supernatant was removed and the precipitate was dried and resuspended in H_2O . The synthetic RNA strands were purified by RP-HPLC using Phenomenex, Luna 5u C18(2) 100A column. Buffer A (100mM TEAA and 5% acetonitrile); Buffer B (100 mM TEAA 50% acetonitrile). The purification was achieved using an 18 min gradient of 0–85% Buffer B. After purification the solvents were lyophilized and the sample was desalted, re-suspended in water and quantified using nano-drop. **3** - HRMS (ESI) m/z : calcd. for $\text{C}_{49}\text{H}_{61}\text{N}_{21}\text{O}_{33}\text{P}_4$ [M-2] $^{-}$ 797.6346; found 797.6353. **RNA2** - HRMS (ESI) m/z : calcd. for $\text{C}_{47}\text{H}_{58}\text{N}_{18}\text{O}_{33}\text{P}_4$ [M-2] $^{-}$ 763.1182; found 763.1222.

Click was done following the previously reported procedure [39]. Reaction progress was monitored by ESI-MS. Upon completion, the reaction mixture was lyophilized, resuspended in H_2O , and purified by RP-HPLC using Phenomenex, Luna 5u C18(2) 100A column. Buffer A (100 mM TEAA and 5% acetonitrile); Buffer B (100 mM TEAA 50% acetonitrile). The purification was achieved using an 18 min gradient of 0–85% Buffer B. After purification the solvents were lyophilized and the sample was desalted, re-suspended in water and quantified using nano-drop. **RNA1** - HRMS (ESI) m/z : calcd. for $\text{C}_{62}\text{H}_{85}\text{CoN}_{25}\text{O}_{33}\text{P}_4$ [Co(**RNA1**)-4] $^{2-}$ 944.6905; found 944.6933.

The clicked RNA strands were purified by RP-HPLC using Phenomenex, Luna 5u C18(2) 100A column. Buffer A (100 mM TEAA and 5% acetonitrile); Buffer B (100 mM TEAA 50% acetonitrile). The purification was achieved using an 18 min gradient of 0–85% Buffer B. After purification the solvents were lyophilized and the sample was desalted, re-suspended in water and quantified using nano-drop.

Prior to the 2-D NMR experiments, the RNA strands were ion exchanged by dissolving in 300 μ L of aqueous 100 μ M KCl. After 20 min equilibration, the solutions were loaded into SpinOUTTM columns (G biosciences GT-100, cat# 786-866) and the supernatants containing RNA were collected after centrifugation at 1000 G. RNA concentrations were determined by nano-drop.

2.5. Preparation and purification of protein

NCp7 was isotope labeled by overexpression in BL21 pLysE cells grown in $^{15}\text{NH}_4\text{Cl}$ enriched media and purified by FPLC following the procedure from previous literature [32].

3. Results and discussion

3.1. Synthesis and characterization

We designed an air-stable Co(III) complex, **1**, that can be easily attached to an RNA strand of interest using click chemistry under aerobic conditions. Compound **1** is diamagnetic, thus allowing full characterization by NMR, in addition to mass spectrometry and X-ray crystallography. The metal can be reduced to paramagnetic Co(II) using sodium ascorbate after the probe is covalently attached to RNA and just prior to PRE NMR experiments [40]. The tetravalent macrocyclic ligand, cyclam, was chosen for its small size (relative to other known cobalt-binding ligands) and strong affinity to the metal ($\log K = 12.7$ for Co(II)-

cyclam) [41,42]. Macrocyclic ligands are known to provide higher kinetic stability to coordinating metals in comparison to their acyclic counterparts [43].

Compound **1** was synthesized in four steps from commercially available cyclam, as illustrated in Scheme 1. The synthesis commenced with selective protection of three secondary amine groups with Boc [33]. The remaining secondary amine group was alkylated with propargyl bromide. The Boc groups were cleaved using trifluoroacetic acid. Compound **1** was achieved upon refluxing CoCl_2 in the ethanolic solution of **7** under aerobic conditions. Oxidation state of cobalt was evident based on the green color of the crystals which agrees with previous literature [44].

An X-ray crystallographic analysis was undertaken to confirm the identity of **1**. Slow diffusion of ether vapor into 10 mM methanolic solution of **1** yielded green blocks suitable for X-ray analysis. The molecular structure, solved in monoclinic space group $C 2/c$, was formulated as $[\text{Co}(\text{7})\text{Cl}_2]\text{Cl}$. An ORTEP representation of the X-ray structure is shown in Fig. 1. The hexacoordinate Co^{3+} cation resides in a distorted octahedral geometry. The cobalt ion is coordinated in equatorial plane by the four nitrogen atoms of the macrocycle with interatomic distances of 1.965, 1.971, 1.975, and 2.066 Å. The apical positions are coordinated by the chloride ions with the observed interatomic distances of 2.2588 and 2.2678 Å. The observed crystallographic parameters are in agreement with previously reported Co^{3+} structure [44].

Diamagnetic Co^{3+} was reduced back to paramagnetic Co^{2+} using sodium ascorbate prior to NMR experiments. There is considerable amount of literature describing octahedral Co^{2+} cyclam complexes that were obtained by reduction of the corresponding Co^{3+} compounds. The reported magnetic susceptibility data suggests that octahedral Co^{2+} complexes containing modified cyclam ligands are typically in a high spin state with three unpaired electrons [45–47].

3.2. Incorporation of the probe into RNA pentanucleotide

As a proof-of-principle, **1** was utilized to study binding of the RNA pentanucleotide.

5'-ACGCU*-3' (U* represents the labeled nucleotide), and HIV-1 nucleocapsid protein, NCp7. The latter plays a number of key roles in the pathogenic lifecycle of the virus, including packaging of viral RNA into budding virions [48]. NCp7 has higher affinity for single stranded DNA or RNA than for double stranded DNA and has specificity for particular oligonucleotide sequences. The azide group was introduced at the 3'-end of the RNA strand using previously reported nucleoside, **2**, immobilized on polystyrene resin [38]. The RNA strand, **RNA1**, having the sequence 5'-ACGCU*-3' was synthesized by solid phase synthesis, as shown in Scheme 2. The paramagnetic probe was subsequently attached via 'click' chemistry using conditions that minimize Cu(I)-catalyzed hydrolysis of the RNA's phosphodiester backbone [39,49]. The later click reaction was done with excess sodium ascorbate to reduce cobalt back to its paramagnetic state, Co^{2+} [40,50,51].

There is a wealth of data describing oligonucleotide binding by NCp7, including an NMR structure of the protein's zinc finger domain bound to the single-stranded DNA

pentanucleotide, 5'-ACGCC-3' [52, 53]. The key binding interaction in this structure is pi-stacking between the indole ring of W25 inserted between the C2 and G3 bases, and the guanosine residue. Previously reported binding studies indicated that NCp7 binds equally tightly to the DNA and RNA pentanucleotide of the sequence described above [54,55]. This reported structure further predicts that attachment of the paramagnetic probe at the 3'-end of the pentanucleotide would place it in proximity of the protein, while minimally affecting its binding affinity. We replaced the cytidine residue at the 3'-end of the pentanucleotide by uridine to simplify the synthesis.

3.3. Molecular dynamics simulations

Molecular dynamics (MD) simulations, based on the published NMR structure of the NCp7-d(ACGCC) complex, were carried out to predict which protein residues would be most affected by the paramagnetic probe placed at the 3'-uridine. Ten parallel MD simulations initiated with different starting structures of the protein-nucleic acid complex were performed, generating a total of 1 microsecond simulation data. The simulation data was analyzed to assess the proximity of the paramagnetic probe to different amino acid residues of the protein. Fig. 2B shows the predicted interaction surface map of the RNA binding domain of NCp7 color-coded to reflect the proximity to the paramagnetic probe. A distance cut-off of 0.5 nm between the cobalt ion to the protein backbone was used for this purpose. Due to its attachment to the RNA's 3'-end nucleotide, which is not directly involved in protein binding, the paramagnetic probe will have some rotational flexibility bringing it within the contact distance with several sites on the protein surface, as shown in Fig. 2. From the figure it is evident that the amino acid residues R29, G35, and D48 would be in closest proximity to the probe and therefore will experience the greatest attenuation effects.

3.4. PRE NMR experiments

The proposed paradigm was tested by titrating ^{15}N -labeled NCp7 with either paramagnetically tagged **RNA1** or an unlabeled control RNA pentanucleotide, **RNA2**. ^{15}N -HSQC NMR spectra were obtained by titrating substoichiometric solutions of RNA to 500 μM [U - ^{15}N] NCp7 (450 μL) in 10 mM potassium phosphate buffer, 90% H_2O /10% D_2O . HSQC spectra of ^{15}N -labeled NCp7 protein, as well as NCp7 bound to the unlabeled RNA were consistent with the ones reported previously [54,55].

Presence of the Co^{2+} probe resulted in global attenuation of the cross peak intensities. The extent of attenuation was found to be dependent on the predicted proximity of the protein's backbone to the paramagnetic center. Fig. 3A shows an overlay of the HSQC spectrum corresponding to NCp7 bound to **RNA1** and NCp7 bound to the unlabeled RNA, **RNA2**. The peak volumes from the spectra were normalized against the residue most unaffected by the paramagnetic probe: E42. The extent of attenuation, illustrated in Fig. 3B, was quantitated by computing the ratios of each corresponding cross peak area between the two spectra.

The observed attenuation is consistent with the MD predictions. The strongest attenuation of intensity was observed for the G35 residue. Residues R29 and D48 were also attenuated, as predicted by the simulations. In addition, strong attenuation was observed for the following

protein residues: K14, C15, N17, C18, A25, K26, N27, C28, H44, and R52. The graph bars were color coded to indicate the relative extent of attenuation indicative of the spatial proximity of the paramagnetic probe. Based on these results we were able to construct the protein's interactive surface map, shown in Fig. 3C, which is strikingly similar to that obtained by using molecular simulations (Fig. 2B). Residues shown in red are the most attenuated and thus the closest to the paramagnetic probe, while the ones shown in blue are the least attenuated.

The molecular map, shown in Fig. 3C, provides a visual representation of the residues most attenuated by proximity of the paramagnetic probe. Fig. 3D is a reduced representation of Fig. 3C, showing specifically the attenuation experienced by the N-H cross peaks from the protein backbone. Both of these describe a unique spectroscopic signature representing NCp7 bound to the paramagnetically labeled **RNA1**. Their distinguishing feature is the HSQC cross peaks whose intensities are attenuated as a function of proximity to the paramagnetic probe.

Paramagnetic metal-based probes can potentially offer three NMR experimental observables that yield long-range structural information: PREs, PCSs, and RDCs. While the PRE effect can be detected in any paramagnetic system, PCSs and RDCs can only be observed in systems with an anisotropic electron g -factor and there are a number of reports describing high spin Co(II)-based probes having PCS effects. The latter are dependent on the magnetic susceptibility tensor (usually referred to as the χ tensor). In general, if the electron g tensor is anisotropic, the χ tensor is also anisotropic [14]. The relationship between g -tensor and χ -tensor has been mathematically described [56]. The paramagnetic probe, **1**, is attached to the RNA pentanucleotide through a flexible linker that allows the principal axes of the χ tensor to fluctuate within the frame of the macromolecule. In addition to the flexible linker, the model system also experiences a degree of conformational mobility from the oligonucleotide. The MD calculations, shown in Fig. 2, predicted proximity of the probe to several residues of NCp7. Because of the flexibility from both the linker and the oligonucleotide, the magnitude of the PCSs has been significantly reduced and only the PRE effects have been observed.

4. Conclusions

This work describes a new approach for studying RNA-protein interactions using two-dimensional correlation NMR experiments. We have shown that labeling a model RNA pentanucleotide with a Co(II)-based paramagnetic NMR probe allowed for creation of visual maps describing the RNA's binding to HIV-1 nucleocapsid protein, NCp7. The probe resulted in distance-dependent attenuation of HSQC crosspeaks corresponding to amide backbone of specific amino acid residues of NCp7 protein. The signal attenuation observed in the two-dimensional NMR experiments was consistent with the predictions obtained by MD simulation. The structural data obtained from this proof-of-concept study is in agreement with the previously described NOE experiments that elucidated binding of the unlabeled pentanucleotide, 5'-ACGCC-3', and NCp7 [55]. This serves as a confirmation that the paramagnetic probe was attached just outside of the binding site, having minimal impact on the RNA-protein binding. In the future, the Co(II)-based paramagnetic probe

could be used to study more complex systems involving RNA-protein interactions with contact distances further than the NOE distance constraint of 6 Å.

Future work will be focused on redesigning the probe to reduce its rotational flexibility. A Co(II)-based probe with a rigid attachment will be able to provide PCS data, in addition to PRE. Our ultimate goal is to apply the paramagnetic metal-based NMR probes to study how NCp7 binds viral RNA in live human cells [57,58]. To achieve this, the probe will have to be covalently attached at a strategically chosen internal positions of significantly larger RNA strands, such as 20-nt long stem loop SL3 [57].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

M. R. would like to thank the Research Foundation of the State University of New York at Albany for financial support of this project. We also thank the National Science Foundation, CHE-MRI-1337594, for instrumentation support of this work, as well as Dr. Zhang for solving the crystal structure of **1**.

Appendix A Supplementary data

Simulation methods described (PDF)

Table 1. X-ray crystallographic data collection and refinement parameters for compound **1**.

X-ray crystallographic files in CIF format for the structure determination of **1**.

HPLC trace of **RNA1**.

Figure S4. HSQC overlay of **RNA1** and **RNA2**.

¹H and ¹³C NMR of compounds **5–7**.

Mass spectra of compounds **3**, **RNA1**, **RNA2**.

Supplementary data associated with this article can be found in the online version, at doi: 10.1016/j.jinorgbio.2017.02.024.

Abbreviations

NCp7	Nucleocapsid protein 7
HSQC	Heteronuclear single quantum coherence
PRE	Paramagnetic resonance enhancement
TLC	Thin layer chromatography
NMR	Nuclear magnetic resonance
RNA	Ribonucleic acid

A	Adenosine
C	Cytidine
U	Uridine
G	Guanosine
MD	Molecular dynamics
TFA	Trifluoroacetate
FPLC	Fast protein liquid chromatography
BOC	<i>tert</i> -Butoxycarbonyl

References

1. Li H, Sun H. In-cell NMR: an emerging approach for monitoring metal-related events in living cells. *Metallomics*. 2014; 6:1–2. <http://dx.doi.org/10.1039/c3mt00224a>.
2. Castel SE, Martienssen RA. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat Rev Genet*. 2013; 14:1–2. <http://dx.doi.org/10.1038/nrg3355>. [PubMed: 23165185]
3. van Kouwenhove M, Kedde M, Agami R. MicroRNA regulation by RNA-binding proteins and its implications for cancer. *Nat Rev Cancer*. 2011; 11:1–2. <http://dx.doi.org/10.1038/nrc3107>.
4. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. The complete atomic structure of the large ribosomal subunit at 2. *Science*. 2000; 289:1–2. <http://dx.doi.org/10.1126/science.289.5481.905>.
5. Carter AP, Clemons WM, Brodersen DE, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature*. 2000; 407:1–2. <http://dx.doi.org/10.1038/35030019>. [PubMed: 10993036]
6. Schlutzenzen F, Tocilj A, Zarivach R, Harms J, Gluehmann M, Janell D, Bashan A, Bartels H, Agmon I, Franceschi F, Yonath A. Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell*. 2000; 102:1–2. [http://dx.doi.org/10.1016/S0092-8674\(00\)00084-2](http://dx.doi.org/10.1016/S0092-8674(00)00084-2). [PubMed: 10929706]
7. Hänsel R, Luh LM, Corbeski I, Trantirek L, Dötsch V. In-cell NMR and EPR spectroscopy of biomacromolecules. *Angew Chem Int Ed* (. 2014; :n/a–n/a. DOI: 10.1002/anie.201311320
8. Selenko P, Wagner G. Looking into live cells with in-cell NMR spectroscopy. *J Struct Biol*. 2007; 158:1–2. <http://dx.doi.org/10.1016/j.jsb.2007.04.001>. [PubMed: 17251039]
9. Sakakibara D, Sasaki A, Ikeya T, Hamatsu J, Hanashima T, Mishima M, Yoshimasu M, Hayashi N, Mikawa T, Wälchli M, Smith BO, Shirakawa M, Güntert P, Ito Y. Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature*. 2009; 458:1–2. <http://dx.doi.org/10.1038/nature07814>.
10. Amarasinghe GK, De Guzman RN, Turner RB, Chancellor KJ, Wu ZR, Summers MF. NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. *J Mol Biol*. 2000; 301:1–2. <http://dx.doi.org/10.1006/jmbi.2000.3979>. [PubMed: 10926487]
11. Schubert M, Dominguez C, Duss O, Ravindranathan S, Allain FHT. How to study the structure and dynamics of protein-RNA complexes by NMR spectroscopy. *Adv Biomed Spectrosc*. 2011; 3:1–2. <http://dx.doi.org/10.3233/978-1-60750-695-9-249>.
12. Foot JN, Feracci M, Dominguez C. Screening protein - single stranded RNA complexes by NMR spectroscopy for structure determination. *Methods*. 2014; 65:1–2. <http://dx.doi.org/10.1016/j.ymeth.2013.09.018>. [PubMed: 24184353]
13. Hennig, J., Warner, LR., Simon, B., Geerlof, A., Mackereth, CD., Sattler, M. Structural analysis of protein-RNA complexes in solution using NMR paramagnetic relaxation enhancements; *Methods Enzymol*. 2015. p. 1-2.<http://dx.doi.org/10.1016/bs.mie.2015.02.006>

14. Marius Clore G, Iwahara J. Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem Rev.* 2009; 109:1–2. <http://dx.doi.org/10.1021/cr900033p>. [PubMed: 19140772]
15. Fawzi NL, Fleissner MR, Anthis NJ, Kálai T, Hideg K, Hubbell WL, Clore GM. A rigid disulfide-linked nitroxide side chain simplifies the quantitative analysis of PRE data. *J Biomol NMR.* 2011; 51:1–2. <http://dx.doi.org/10.1007/s10858-011-9545-x>.
16. Qin PZ, Dieckmann T. Application of NMR and EPR methods to the study of RNA. *Curr Opin Struct Biol.* 2004; 14:1–2. <http://dx.doi.org/10.1016/j.sbi.2004.04.002>.
17. Nguyen P, Qin PZ. RNA dynamics: perspectives from spin labels, *Weiley Interdiscip. Rev RNA.* 2012; 3:1–2. <http://dx.doi.org/10.1002/wrna.104>. [PubMed: 21898828]
18. Clore GM, Tang C, Iwahara J. Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Curr Opin Struct Biol.* 2007; 17:1–2. <http://dx.doi.org/10.1016/j.sbi.2007.08.013>.
19. Grant GPG, Boyd N, Herschlag D, Qin PZ. Motions of the substrate recognition duplex in a group I intron assessed by site-directed spin labeling. *J Am Chem Soc.* 2009; 131:1–2. <http://dx.doi.org/10.1021/ja808217s>.
20. Anthis NJ, Clore GM. The length of the calmodulin linker determines the extent of transient interdomain association and target affinity. *J Am Chem Soc.* 2013; 135:1–2. <http://dx.doi.org/10.1021/ja4051422>.
21. Anthis NJ, Doucleff M, Marius Clore GM. Transient, sparsely populated compact states of apo and calcium-loaded calmodulin probed by paramagnetic relaxation enhancement: interplay of conformational selection and induced fit. *J Am Chem Soc.* 2011; 133:1–2. <http://dx.doi.org/10.1021/ja2082813>. [PubMed: 21142155]
22. Nguyen THD, Ozawa K, Stanton-Cook M, Barrow R, Huber T, Otting G. Generation of pseudocontact shifts in protein NMR spectra with a genetically encoded cobalt(II)-binding amino acid. *Angew Chem Int Ed.* 2011; 50:1–2. <http://dx.doi.org/10.1002/anie.201005672>.
23. Keizers PHJ, Ubbink M. Paramagnetic tagging for protein structure and dynamics analysis. *Prog Nucl Magn Reson Spectrosc.* 2011; 58:1–2. <http://dx.doi.org/10.1016/j.pnmrs.2010.08.001>. [PubMed: 21241883]
24. Su XC, Otting G. Paramagnetic labelling of proteins and oligonucleotides for NMR. *J Biomol NMR.* 2010; 46:1–2. <http://dx.doi.org/10.1007/s10858-009-9331-1>. [PubMed: 19924385]
25. Rosnizeck IC, Graf T, Spoerner M, Tränkle J, Filchtinski D, Herrmann C, Gremer L, Vetter IR, Wittinghofer A, König B, Kalbitzer HR. Stabilizing a weak binding state for effectors in the human ras protein by cyclen complexes. *Angew Chem Int Ed.* 2010; 49:1–2. <http://dx.doi.org/10.1002/anie.200907002>.
26. Jensen MR, Led JJ. Metal-protein interactions: structure information from Ni²⁺-induced pseudocontact shifts in a native nonmetalloprotein. *Biochemistry.* 2006; 45:8782–8787. [PubMed: 16846221]
27. Petersen G, Lauritzen C, Pedersen J, Led JJ. Metal binding sites in proteins: identification and characterization by paramagnetic NMR relaxation. *Biochemistry.* 2005; 44:11014–11023. [PubMed: 16101285]
28. Swarbrick JD, Ung P, Dennis ML, Lee MD, Chhabra S, Graham B. Installation of a rigid EDTA-like motif into a protein α -helix for paramagnetic NMR spectroscopy with cobalt(II) ions. *Chem - A Eur J.* 2016; 22:1–2. <http://dx.doi.org/10.1002/chem.201503139>.
29. Pintacuda G, Moshref A, Leonchiks A, Sharipo A, Otting G. Site-specific labelling with a metal chelator for protein-structure refinement. *J Biomol NMR.* 2004; 29:1–2. <http://dx.doi.org/10.1023/B:JNMR.0000032610.17058.fe>. [PubMed: 15017135]
30. Martial Piotto VS, Vladimir Saudek, Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR.* 1992; 2:661–665. [PubMed: 1490109]
31. Keller, Rochus. The Computer Aided Resonance Assignment Tutorial. 2004
32. De Guzman RN, Wu ZR, Stalling CC, Pappalardo L, Borer PN, Summers MF. Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. *Science.* 1998; 279:384–388. <http://www.ncbi.nlm.nih.gov/pubmed/9430589>. [PubMed: 9430589]

33. Yu M, Price JR, Jensen P, Lovitt CJ, Shelper T, Duffy S, Windus C, Avery VM, Rutledge PJ, Todd MH. Copper, Nickel and Zinc Cyclam-amino Acid and Cyclam-peptide Complexes may be Synthesized with “Click” Chemistry and are Non-toxic. 2006
34. Sheldrick, GM. A short history of SHELX. 2008. p. 1-2. <http://dx.doi.org/10.1107/S0108767307043930>
35. Sheldrick GM. SHELXL93. Program for the Refinement of Crystal Structures. 1993
36. Sheldrick, GM. SADABS: Area-Detector Absorption Correction. University of, Göttingen; Göttingen, Ger: 2008.
37. Spek, AL. PLATON, A Multipurpose Crystallographic Tool. Utrecht. Univ; Utrecht, Netherlands: 2008.
38. Santner T, Hartl M, Bister K, Micura R. Efficient Access to 3'-Terminal Azide-Modified RNA for Inverse Click-Labeling Patterns. 2014
39. Paredes E, Das SR. Click chemistry for rapid labeling and ligation of RNA. *Chembiochem*. 2011; 12:1–2. <http://dx.doi.org/10.1002/cbic.201000466>.
40. Cressey PB, Eskandari A, Bruno PM, Lu C, Hemann MT. The Potent Inhibitory Effect of a Naproxen-Appended Cobalt (III) - Cyclam Complex on Cancer Stem Cells. 2016; 2139:1–2. <http://dx.doi.org/10.1002/cbic.201600368>.
41. Yoo J, Reichert DE, Welch MJ. Comparative in vivo behavior studies of cyclen-based copper-64 complexes: regioselective synthesis, X-ray structure, radiochemistry, log P, and biodistribution. *J Med Chem*. 2004; 47:6625–6637. <http://dx.doi.org/10.1021/jm0496990>. [PubMed: 15588098]
42. NIST. Critical Stability Constants of Metal Complexes Database. 1993
43. Ozay H, Baran Y, Miyamae H. The stabilities and formation kinetics of some macrocycles with copper(II): crystal structures of some pendant arm macrocycles. *J Coord Chem*. 2011; 64:1–2. <http://dx.doi.org/10.1080/00958972.2011.574128>.
44. Fuess H, M A, Ivaniková R, Svoboda I. trans-Dichloro-(1,4,8,11-tetra-azacyclo-tetra-decane)-cobalt(III) chloride. *Crystallogr Commun*. 2006;m1553–m1554.
45. Bertini, I., Luchinat, C. NMR of Paramagnetic Molecules in Biological Systems. The Benjamin/Cummings Publishing Company, Inc., Department of Chemistry. University of Florence; Italy: 1986.
46. Caneschi A, Dei A, Gatteschi D, Tangoulis V. Antiferromagnetic coupling in a six-coordinate high spin cobalt(II)-semiquinonato complex. *Inorg Chem*. 2002; 41:3508–3512. [http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1521-3773\(20000103\)39:1%3C246::AID-ANIE246%3E3.0.CO;2-B/full](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1521-3773(20000103)39:1%3C246::AID-ANIE246%3E3.0.CO;2-B/full). [PubMed: 12079471]
47. Jones, RD., Summerville, DA., Basolo, F. Synthetic Oxygen Carriers Related to Biological Systems. 1979. <http://dx.doi.org/10.1021/cr60318a002>
48. D'Souza V, Summers MF. How retroviruses select their genomes. *Nat Rev Microbiol*. 2005; 3:1–2. <http://dx.doi.org/10.1038/nrmicro1210>.
49. Gramlich PME, Warncke S, Gierlich J, Carell T. Click-click-click: single to triple modification of DNA. *Angew Chem Int Ed*. 2008; 47:1–2. <http://dx.doi.org/10.1002/anie.200705664>.
50. Bosnich MLTB, Poon CK. Complexes of cobalt(III) with a cyclic tetradentate secondary amine. *Inorg Chim Acta*. 1965; 4:1–2. [http://dx.doi.org/10.1016/S0020-1693\(00\)93250-X](http://dx.doi.org/10.1016/S0020-1693(00)93250-X).
51. Davies NB. Reactions of L-ascorbic acid with transition metal complexes. *Polyhedron*. 1992; 11:285–321.
52. Berkowitz RD, Luban J, Goff SP. Specific binding of human immunodeficiency virus type 1 gag polypeptide and nucleocapsid protein to viral RNAs detected by RNA mobility shift assays. *J Virol*. 1993; 67:7190–7200. doi:0022-538X/93/127190-11\$02.00/0. [PubMed: 8230441]
53. Clever J, Sasseti C, Parslow TG. RNA secondary structure and binding sites for gag gene products in the 5' packaging signal of human immunodeficiency virus type 1. *J Virol*. 1995; 69:2101–2109. [PubMed: 7884856]
54. South TL, Summers MF. Zinc- and sequence-dependent binding to nucleic acids by the N-terminal zinc finger of the HIV-1 nucleocapsid protein: NMR structure of the complex with the Psi-site analog, dACGCC. *Protein Sci*. 1993; 2:1–2. <http://dx.doi.org/10.1002/pro.5560020102>.

55. Morellet N, Déméné H, Teilleux V, Huynh-Dinh T, de Rocquigny H, Fournié-Zaluski MC, Roques BP. Structure of the complex between the HIV-1 nucleocapsid protein NCp7 and the single-stranded pentanucleotide d(ACGCC). *J Mol Biol.* 1998; 283:1–2. <http://dx.doi.org/10.1006/jmbi.1998.2098>. [PubMed: 9761668]
56. Bertini I, Luchinat C, Parigi G. Magnetic susceptibility in paramagnetic NMR. *Prog Nucl Magn Reson Spectrosc.* 2002; 40:1–2. [http://dx.doi.org/10.1016/S0079-6565\(02\)00002-X](http://dx.doi.org/10.1016/S0079-6565(02)00002-X).
57. Shubsda MF, Paoletti AC, Hudson BS, Borer PN. Affinities of packaging domain loops in HIV-1 RNA for the nucleocapsid protein. *Biochemistry.* 2002; 41:1–2. <http://dx.doi.org/10.1021/bi016045+> [PubMed: 11771996]
58. Anton H, Taha N, Boutant E, Richert L, Khatter H, Klaholz B, Rondé P, Réal E, De Rocquigny H, Mély Y. Investigating the cellular distribution and interactions of HIV-1 nucleocapsid protein by quantitative fluorescence. *Microscopy.* 2015; 1:1–2. <http://dx.doi.org/10.1371/journal.pone.0116921>.

Fig. 1. Paramagnetic NMR probe, **1**, and Ortep representation, showing 50% thermal ellipsoids and selected atom labels. Hydrogen atoms were omitted for clarity. A molecule of methanol was also present in the unit cell.

Fig. 2.

(A) Amino acid sequence of NCp7. The reported NMR structure, as well the computational studies described in this work, are based on the amino acid residues 13–53 that encompass the two zinc knuckle domains. (B) A snapshot from the MD simulation highlighting the attached cobalt probe in the NCp7-**RNA1** complex. (C) Predicted proximity of the protein's backbone to the paramagnetic probe, representative probe conformations. In (B) and (C), surface representation is used for the protein, licorice for the RNA/probe with the cobalt ion as van der Waals sphere.

Fig. 3.

(A) Overlay of ^{15}N -HSQC spectrum for NCp7-**RNA2** (red) and NCp7-**RNA1** (blue). The residues experiencing the strongest attenuation are shown in boxes. (B) Ratio of peak volume change of the cross peak intensities. Black circles represent the residues which are lowered in intensity upon binding of the metal-free RNA (F16, W37, K38, M46.) These are excluded from the analysis, in addition to P31. Red diamonds represent the residues (N17, K26, G35) which were excessively broadened due to proximity to the metal-RNA and also removed from the peak volume change analysis. (C) Protein surface map, based on the previously reported NMR structure [54,55]. The amino acid residues are color-coded based on the extent of attenuation of the HSQC cross peak intensities. (D) Protein backbone map, the backbone is color-coded based on the extent of attenuation of the HSQC cross peak intensities.

Scheme 1.

Synthesis of the 'clickable' NMR probe **1**; (a) Boc_2O , Et_3N , CH_2Cl_2 ; (b) propargyl bromide, K_2CO_3 , MeCN ; (c) TFA , CH_2Cl_2 ; (d) CoCl_2 , EtOH , 75°C .

Scheme 2.

Solid phase synthesis of the RNA pentanucleotide modified with **1** using click chemistry, and unmodified RNA strand.