

Probing Anomalous Structural Features in Polypurine Tract-Containing RNA–DNA Hybrids with Neomycin B[†]

Robert G. Brinson,^{‡,||,⊥} Kevin B. Turner,^{§,⊥} Hye Young Yi-Brunozzi,^{||} Stuart F. J. Le Grice,^{||} Daniele Fabris,^{*,§} and John P. Marino^{*,‡}

[‡]Center for Advanced Research in Biotechnology of the University of Maryland Biotechnology Institute and the National Institute of Standards and Technology, 9600 Gudelsky Drive, Rockville, Maryland 20850, [§]Department of Chemistry and Biochemistry, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, Maryland 21228, and ^{||}HIV Drug Resistance Program, National Cancer Institute, National Institutes of Health, Frederick, Maryland 20892. [⊥]These authors contributed equally to this work.

Received March 2, 2009; Revised Manuscript Received May 18, 2009

ABSTRACT: During (–)-strand DNA synthesis in retroviruses and *Saccharomyces cerevisiae* LTR retrotransposons, a purine rich region of the RNA template, known as the polypurine tract (PPT), is resistant to RNase H-mediated hydrolysis and subsequently serves as a primer for (+)-strand, DNA-dependent DNA synthesis. Although HIV-1 and Ty3 PPT sequences share no sequence similarity beyond the fact that both include runs of purine ribonucleotides, it has been suggested that these PPTs are processed by their cognate reverse transcriptases (RTs) through a common molecular mechanism. Here, we have used the aminoglycoside neomycin B (NB) to examine which structural features of the Ty3 PPT contribute to specific recognition and processing by its cognate RT. Using high-resolution NMR, direct infusion FTICR mass spectrometry, and isothermal titration calorimetry, we show that NB binds preferentially and selectively adjacent to the Ty3 3′ PPT–U3 cleavage junction and in an upstream 5′ region where the thumb subdomain of Ty3 RT putatively grips the substrate. Regions highlighted by NB on the Ty3 PPT are similar to those previously identified on the HIV-1 PPT sequence that are implicated as contact points for substrate binding by its RT. Our findings thus support the notion that common structural features of lentiviral and LTR-retrotransposon PPTs facilitate the interaction with their cognate RT.

During minus (–)-strand DNA synthesis, the retroviral and long terminal repeat (LTR)¹ retrotransposon RNA genomes are nonspecifically hydrolyzed by the RNase H domain of their respective reverse transcriptase (RT). In human immunodeficiency virus type 1 (HIV-1), two purine rich regions, known as the 3′ and the central polypurine tracts (PPT), are refractory to RNase H-mediated degradation, and the resulting PPT-containing

RNA–DNA hybrids subsequently serve as initiation sites for (+)-strand DNA synthesis. The PPT RNA primers are later removed from nascent DNA, and polymerization continues to produce a full-length DNA duplex. Precise removal of the (+)- and (–)-strand RNA primer is crucial for replication of retroviruses and LTR retrotransposons, since this defines the duplex termini that are subsequently recognized by the integration machinery of the particular element (1, 2).

Selective processing and utilization of PPTs as primers in reverse transcription are highly conserved across retroviruses and retrotransposons (3, 4). Retroviral PPTs are commonly defined by an interrupted A-tract followed by a G-tract (Figure 1). In contrast, LTR retrotransposon PPTs, such as those of Ty1 and Ty3 (derived from *Saccharomyces cerevisiae*), are shorter and share little sequence similarity with those of retroviruses beyond the fact that all the sequences contain a string of consecutive purine ribonucleotides. Despite these differences, it would not be an unreasonable assumption that PPTs are recognized and processed by their cognate reverse transcriptases (RTs) through a common molecular mechanism (3). The idea that similar structural motifs within the PPTs are recognized by their cognate RTs is supported by data from a recent mutagenic study of both HIV-1 and *Saccharomyces cerevisiae* LTR retrotransposon Ty3 RTs (5).

[†]This research was funded by the National Institutes of Health (Grant GM59107 to J.P.M. and Grant GM643208 to D.F.) and by the National Science Foundation (Grant CHE-0439067 to D.F.). S.F.J.L.G. was supported by the intramural research program of the Center for Cancer Research, NCI, NIH. R.G.B. acknowledges support from a NIST/NIH NRC postdoctoral fellowship.

*To whom correspondence should be addressed. J.P.M.: Center for Advanced Research in Biotechnology, 9600 Gudelsky Dr., Rockville, MD 20850; telephone, (240) 314-6160; fax, (240) 314-6255; e-mail, marino@umbi.umd.edu. D.F.: Department of Chemistry and Biochemistry, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21228; telephone, (410) 455-3053; fax, (410) 455-2608; e-mail, fabris@umbc.edu.

Abbreviations: PPT, polypurine tract; LTR, long terminal repeat; HIV-1, human immunodeficiency virus type 1; RT, reverse transcriptase; NB, neomycin B; NMR, nuclear magnetic resonance; MS, mass spectrometry; ESI-FTICR, electrospray ionization Fourier transform ion cyclotron; ITC, isothermal titration calorimetry; dF, 2,4-difluoro-5-methylbenzene; PAGE, polyacrylamide gel electrophoresis; SORI-CID, sustained off-resonance irradiation-collision-induced dissociation; CD, circular dichroism; CHEF, correlated harmonic excitation fields.

HIV -1 (HXB2)	UCUUAGCCACUUUUU <u>AAAAGAAAAGGGGGG</u> CUGGAAGGGCUAA
HIV -1 (NL43)	UCUUAGCCACUUUUU <u>AAAAGAAAAGGGGGG</u> CUGGAAGGGCUAA
HIV -2 (BEN)	UAUGUCACAUUUUU <u>AAAAGAAAAGGGGGG</u> CUGCAAGGGGAUGU
HIV -2 (NIHZ)	UAUGUCACAUUUUU <u>AAAAGAAAAGGGGGG</u> CUGGAAGGGGCUU
SIV (MND)	CUUCUCACAUUUUU <u>AAAAGAAAAGGGGGG</u> CUGGGAGGGGAUU
SIV (AGM)	UUUUUCGACUUUUU <u>AAAAGAAAAGGGGGG</u> CUGGAAGGGGAUU
CAEV lenti	AACACCAC <u>AAAAUAAAAA</u> AGGAGGACUGUGAGA
BIV lenti	UCCUAUUUUAAACUU <u>AAAAGGGGGG</u> CUGUGGGGCAGGG
Ty1 (retrotransposon)	CAUUU <u>UGGGUGU</u> AUGUU
Ty3 (retrotransposon)	CCCUGAGAGAGGGAAG <u>UGUUGUAUCUC</u>

FIGURE 1: Comparison of 3' PPT and flanking RNA sequences from selected lentiviruses and LTR retrotransposons. The PPT RNA sequence is underlined. See ref (4) for a more complete listing of LTR-containing retroelements.

HIV-1 RT, a p66–p51 heterodimer, is a multifunctional enzyme that exhibits DNA polymerase and RNase H activities. These two catalytic centers are spaced a distance equivalent to approximately 17 bp which correlates with the length (15 bp) of the HIV-1 PPT sequence (6–8). The crystal of an HIV-1 RT–PPT complex, determined at 3 Å resolution, indicated protein contacts with the substrate mostly along the sugar–phosphate backbone and also revealed weak or mispaired bases in the RNA–DNA hybrid (6). While the crystal structure provided important molecular insights into the RT–PPT complex, it did not reveal the entire PPT sequence, and the RNase H catalytic site was improperly oriented with respect to the scissile phosphate at the PPT–U3 junction. In contrast, Ty3 RT is a 55 kDa monomeric protein with catalytic centers similar to those of HIV-1 RT. Although a crystal structure of the Ty3 RT–PPT complex is not yet available, the length of the Ty3 PPT sequence (12 bp) also likely correlates with the distance between its DNA polymerase and RNase H catalytic centers.

Recent NMR studies of the HIV-1 and Ty3 PPTs have indicated that, in the absence of RT, these hybrids adopt regular A-form helical structures with standard Watson–Crick pairing. However, NMR analysis of these PPTs containing either the thymidine isostere, 2,4-difluoro-5-methylbenzene (dF), or base pair modifications revealed that single-site alterations induce periodic and long-range structural perturbations, presumably due to altered base stacking in the purine tracks that propagate along the hybrid helix (9). Anomalous structural features in the HIV-1 PPT were further highlighted by using the aminoglycoside neomycin B (NB) as a molecular probe (10). In this study, it was shown that both NB and RT interacted with the same regions of the HIV-1 PPT, namely, the PPT–U3 junction and the distal (rA)₄(dT)₄ tract. These findings highlighted the utility of NB as a probe of distortable regions in the PPT hybrid, which are likely a common feature of its recognition by NB and HIV-1 RT. Moreover, the results supported the proposed distance correlation between the RNase H active site and the thumb subdomain for proper alignment of RT on the PPT substrate.

As an extension of our studies of the interaction of archetypical nucleic acid ligands with the HIV-1 PPT (10), here we have utilized NB to examine structural features in the Ty3 PPT hybrid. Direct infusion electrospray ionization (ESI) (11, 12), Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) (13, 14), and high-resolution nuclear magnetic resonance (NMR) spectroscopy have been used to determine the binding mode, stoichiometry, and sites of NB interaction with the Ty3 PPT. This interaction was further assayed by isothermal titration calorimetry (ITC) to determine thermodynamic parameters and spectropolarimetry to assess global changes in helical geometry

of the hybrid. These combined approaches have identified two specific NB binding sites on the Ty3 PPT, i.e., proximal to the PPT–U3 junction and overlapping the putative Ty3 RT thumb binding site. Localization of NB binding sites and possible allosteric coupling between them are discussed in light of structural features that may guide Ty3 PPT recognition and subsequent processing by its cognate RT.

MATERIALS AND METHODS²

PPT Substrates. All ribo- and deoxyribonucleoside phosphoramidites were purchased from Glen Research Corp. (Sterling, VA). Oligonucleotides (20-mers) of DNA and RNA were synthesized using standard phosphoramidite chemistry on a PE Biosystems Expedite 8909 nucleic acid synthesizer. The individual strands were purified by preparative polyacrylamide gel electrophoresis (PAGE), dialyzed against sterile ddH₂O and then against either ESI [100 mM ammonium acetate (pH 7.0)] or NMR buffer [80 mM NaCl and 10 mM NaH₂PO₄/Na₂HPO₄ (pH 7.0)], and finally quantified spectrophotometrically. We annealed single strands by mixing equimolar amounts of the desired constructs, heating the solution to 90 °C for 3 min, and allowing the samples to slowly cool to room temperature.

Mass Spectrometry. All analyses were performed on a Bruker Daltonics (Billerica, MA) Apex III FTICR mass spectrometer equipped with a 7T actively shielded superconducting magnet and a nano-ESI source built in-house. The desolvation temperature, skimmer voltage, and other source parameters were optimized to allow examination of intact RNA noncovalent complexes as previously described (15–18). Analyte solutions were mixed with 2-propanol immediately before analysis to a final concentration of 10% (v/v) to assist desolvation. Typically, 5 μL samples were loaded into the nano-ESI needle, and a spray voltage of less than 1 kV was applied to the solution through a stainless steel wire inserted into the back of the needle. No solvent pumps were necessary, as the solution flow rate was dictated by the applied voltage and the size of the nano-ESI needle tip (typically ~1–2 μm). Spectra were acquired in negative ionization mode and processed using either XMASS 7.0.2 or Apex-Control 2.0 (Bruker Daltonics). Scans were completed in broadband mode that allowed for a typical 150000 Da resolving power at *m/z* 2000. Spectra were externally calibrated using a 1 mg/mL solution of CsI, which produced a series of peaks throughout the mass range of *m/z* 1000–6000 and provided a typical mass accuracy of 20 ppm or better across the range. Each analysis was performed a minimum of three times, and representative spectra are shown.

In tandem MS experiments, the precursor ion of interest was isolated in the FTICR cell using correlated RF sweeps (CHEF) (19) and then activated through sustained off-resonance irradiation-collision-induced dissociation (SORI-CID) (20). Frequency offsets below and above the resonant frequency of the precursor ion were sampled to avoid possible “blind spots” in the product spectra. Best results were achieved by using irradiation frequencies that were 600–2000 Hz below that of the

²Certain commercial equipment, instruments, and materials are identified in this paper to specify the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the material or equipment identified is necessarily the best available for the purpose.

precursor ion. Typical activation regimes were reached by applying off-resonant pulses for 250 ms, using 26–31 dB attenuation of the maximum power output allowed by the hardware. Argon was used as the collision gas in pulsed bursts of 100–250 ms, which resulted in momentary increases from 1×10^{-11} to 1×10^{-7} mbar of the background pressure measured by the instrument gauge located underneath the ion optics. No attempt was made to determine the actual pressure within the FTICR cell.

NMR Analysis. Single-stranded RNA and DNA oligonucleotides representing the 20-nucleotide Ty3 PPT hybrid were prepared and annealed as described above. The duplex was dialyzed against 80 mM NaCl and 10 mM $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ (pH 7.0). The final duplex concentration was 225 μM (61.9 nmol). NB was titrated in 0.25 equiv (15.5 nmol) up to a total of 2.0 equiv (123.8 nmol). After each addition, the complex was equilibrated for 20 min before ^1H NMR data were acquired using a Bruker Biospin DMX600 apparatus equipped with a triple-resonance, triple-axis pulsed field shielded gradient TXI HCN cryoprobe. All one-dimensional (1D) spectra were processed with TopSpin (Bruker Biospin, Billerica, MA). ^1H spectral assignments for the PPT hybrids were made using standard NOESY, DQF-COSY, and natural abundance ^{13}C HMQC experiments.

Isothermal Titration Calorimetry. Ty3 PPT duplexes were dialyzed in 80 mM NaCl and 10 mM $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ (pH 7.0). Separate solutions of NB were prepared using the same buffers. ITC experiments were performed at 40 °C on a MicroCal VP-ITC instrument (MicroCal, Inc., Northampton, MA). In a typical experiment, 5 μL aliquots of 900 μM NB were injected from a 275 μL rotating syringe (313 rpm) into an isothermal chamber containing 1.42 mL of 20 μM Ty3 PPT duplexes. The initial delay before the first injection was 180 s. The duration of each 5 μL injection was 10 s with a 240 s delay between each injection. As a control, NB was injected into each buffer alone to take into account the heats of dilution. Each injection generated a heat burst curve (microcalories per second versus seconds) that was integrated using Origin 7.0 (OriginLab Corp., Northampton, MA) to yield the heat associated with each injection. After the heats of dilution were subtracted, the results were plotted as kilocalories per mole of injection versus molar ratio. The plot was fit with a model of either one set or two sets of equivalent sites using Origin 7.0.

RESULTS

Archetypical nucleic acid ligands with well-characterized binding determinants can be exploited to probe the structural features that mediate specific recognition of PPT substrates by the cognate RT protein (10). In particular, the aminoglycoside neomycin B (NB) was found to be capable of establishing stable interactions at the PPT–U3 cleavage junction and within the DNA polymerase primer grip region of the HIV-1 PPT, which NMR data support as being in direct contact with HIV-1 RT in the cognate assembly. In our investigation, the same molecular probes, consisting of a general intercalator (ethidium bromide), a minor groove binder (distamycin A), a mixed-mode intercalator/minor groove binder (mitoxantrone), and a multifunctional polycationic aminoglycoside (NB), were used to examine structural anomalies of the Ty3 PPT-containing RNA–DNA hybrid (Figure 2a, Ty3 PPT_{wt}) by ESI-FTICR mass spectrometry (Table 1 of the Supporting Information). In a clear concentration-dependent fashion, increasing amounts of ligand resulted

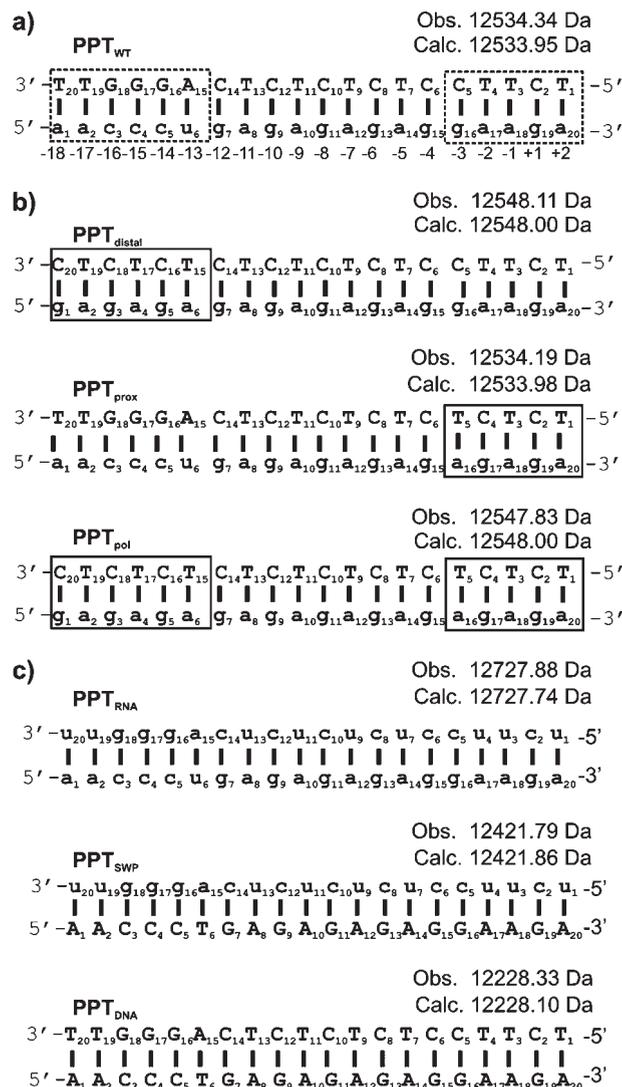


FIGURE 2: Ty3 PPT-containing hybrid duplex sequences. (a) Wild-type hybrid from the *S. cerevisiae* LTR retrotransposon Ty3, denoted as PPT_{wt}. The dotted boxes indicate regions that were mutated in PPT_{distal}, PPT_{prox}, and PPT_{pol}. (b) Mutated sequences. For PPT_{distal}, the region proposed to be contacted by the Ty3 RT thumb was mutated to an rG·rA polymer to remove the distal NB site. For PPT_{prox}, the PPT U3 cleavage junction was mutated to an rG·rA polymer to remove the proximal neomycin binding site. For PPT_{pol}, both NB sites were removed to create an rG·rA 20-mer. Solid boxes indicate regions that were mutated in each sample. (c) Altered “strandedness”: PPT_{RNA}, both strands of RNA; PPT_{DNA}, both strands of DNA; PPT_{SWP}, sequences of the RNA and DNA strands swapped. Sequences are numbered from 5' to 3', and biological numbering relative to the scissile –1g/+1a phosphodiester bond is indicated. Nucleotides in upper- and lowercase denote DNA and RNA, respectively. The monoisotopic mass observed by nanospray-FTICR analysis for each duplex construct is reported together with the respective mass calculated from the sequence.

in the binding of multiple units of ethidium bromide and mitoxantrone. For these ligands, a maximum stoichiometry of three units was observed, consistent with the number of possible binding sites present between adjacent GC pairs, which are particularly favorable to intercalation (21, 22). The minor groove binder distamycin A afforded lower stoichiometries and reduced affinity, which is not surprising as the Ty3 PPT lacks contiguous AT base pairs preferred by this class (23). In analogy with HIV-1 PPT, Ty3 PPT_{wt} bound 2 equiv of NB (Figure 1a of the Supporting Information). However, while the HIV-1 sites

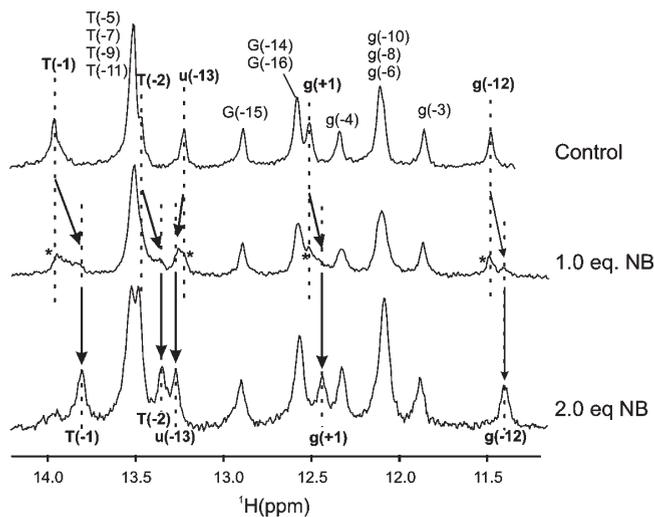


FIGURE 4: 1D water flip-back Watergate ^1H NMR spectra of the imino region of the PPT_{wt} duplex after titration with NB at 10°C in 80 mM NaCl and 10 mM $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ (pH 7.0). The concentration of the PPT_{wt} duplex was $225\ \mu\text{M}$. 1D ^1H NMR spectra are shown in the absence (top trace) and presence of 1.0 equiv (middle trace) and 2.0 equiv (bottom trace) of NB. Imino resonances for T(-1), T(-2), u(-13), g(+1), and g(-12), where chemical shift changes can be tracked, are highlighted in bold. Dotted lines and arrows indicate the shifts in position of the resonances.

simultaneously eliminated by preparing a rG·rA polymer (PPT_{pol}). In this way, if each site bound NB independently, the affinity for $\text{PPT}_{\text{distal}}$ and PPT_{prox} would be expected to be very similar to the affinity of the corresponding site in PPT_{wt} . However, if the features responsible for binding were coupled, affinities for the remaining sites might be weaker or nonexistent compared to that for the wild type. For the PPT_{pol} construct from which both sites had been removed, only nonspecific binding would be expected because of the general affinity of NB for A-form helices.

Initial screening by ESI-FTICR mass spectrometry (data not shown) found that the three mutant Ty3 PPT constructs bound 1 equiv of NB, in contrast to Ty3 PPT_{wt} which bound 2 equiv. Submitting the 1:1 NB- $\text{PPT}_{\text{distal}}$ and NB- PPT_{prox} complexes to MS/MS demonstrated that mutation of the distal and proximal regions eliminated interactions exclusively at these positions (Figure 5). For PPT_{pol} , a 1:1 complex with NB was still observed, indicating that the double mutation did not fully eliminate NB binding. However, MS/MS analysis was unable to locate a defined binding site, which can be attributed either to the limited stability of the interaction or to multiple low-affinity contact points distributed over the entire construct.

NMR-monitored NB titrations with $\text{PPT}_{\text{distal}}$ and PPT_{prox} confirmed that the analogous wild-type binding positions were retained in the mutated samples (Figure 6). In $\text{PPT}_{\text{distal}}$, chemical shift perturbations of the T3, T4, and g19 imino resonances indicated that the binding site adjacent to the PPT-U3 cleavage junction was preserved. Conversely, in PPT_{prox} , chemical shift perturbations of the u6 and g7 imino resonances indicated that the binding site potentially contacted by the thumb subdomain of Ty3 RT was maintained. However, in both cases, NB binding at the remaining wild-type sites exhibited intermediate rather than slow exchange kinetics on the NMR time scale. The apparent affinity for the remaining NB wild-type sites in these two mutants was also weaker, with 2 equiv needed to achieve saturation. A second site(s) was also observed in both $\text{PPT}_{\text{distal}}$ and PPT_{prox}

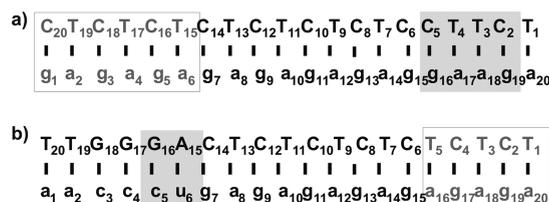


FIGURE 5: Sequences of the (a) $\text{PPT}_{\text{distal}}$ and (b) PPT_{prox} hybrids (gray nucleotides indicate mutated bases) with the gray boxes highlighting binding sites identified using MS/MS footprinting.

at the respective mutated end of the construct, with chemical shift perturbations observed for T19 and g3 imino resonances of $\text{PPT}_{\text{distal}}$ and imino resonances g17, T3, and g19 of PPT_{prox} . However, while NB interacted with similar affinity for these site(s) when compared to the wild-type sites, binding produced general broadening of NMR signals, consistent with a faster off rate (k_{off}) and/or intermediate conformational exchange at the binding interface on the NMR time scale.

These results highlight a degree of allosteric coupling between the two sites of Ty3 PPT_{wt} . If these sites were acting independently from each other, the wild-type site remaining in $\text{PPT}_{\text{distal}}$ or PPT_{prox} constructs would exhibit a very similar affinity to their PPT_{wt} counterpart. Instead, replacing the proximal or distal regions of the PPT with d(CT):r(ga) repeats not only repressed binding of a second NB equivalent but also altered the affinity and binding kinetics of the remaining site. This suggests that long-range structural effects couple the wild-type binding sites, likely through alteration in the stacking of the repeating d(TCT):r(AGA) units of the PPT.

Structural Determinants of NB- PPT_{wt} Binding. The role of helical geometry in selective NB binding was explored by synthesizing PPT hybrids in which the RNA and DNA strands were exchanged (PPT_{swp}) or in which both strands were RNA (PPT_{RNA}) or DNA (PPT_{DNA}) (Figure 2c). These constructs allowed a comparison of interactions of NB with Ty3 PPT hybrids that presented a canonical A-form helical geometry (PPT_{RNA}), a combination of A- and B-form geometries (PPT_{wt} and PPT_{swp}), or a canonical B-form geometry (PPT_{DNA}). Formation of noncovalent complexes between NB and PPT constructs was first monitored directly by ESI-FTICR mass spectrometry to determine binding stoichiometries and affinities. In these experiments, we treated equimolar mixtures of the PPT constructs with NB to directly compare binding specificity. This approach promotes competition among substrates and yields distinctive assembly distributions that reflect the strength of the respective interactions. Because of the similar charging properties of the various species in the mixture, signal intensities correlate well with relative abundances in solution, which can be readily translated into relative scales of binding affinities (see Materials and Methods).

The competition experiment depicted in Figure 7 was carried out by titration of $5\ \mu\text{M}$ solutions of each PPT substrate. ESI-FTICR analysis of the initial mixture provided fully resolved signals for all four substrates together with typical cation adducts observed in nucleic acid analysis (Figure 7a). Noncovalent assemblies containing up to two NB units were readily detected for PPT_{wt} , with up to one ligand for the mutant duplexes after addition of 1 equiv of NB per substrate (Figure 7b; see Materials and Methods). Increasing the NB concentration 10-fold increased the population of the bound states of all substrates but did not affect the maximum stoichiometries (Figure 7c). The

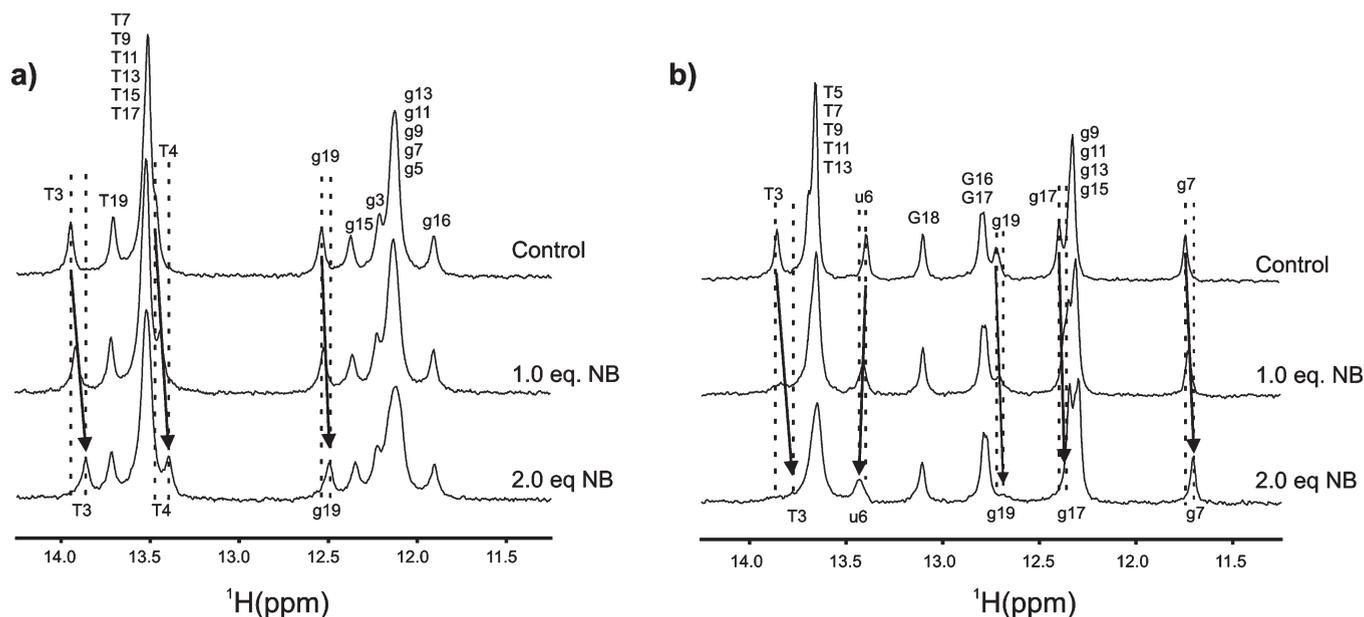


FIGURE 6: 1D water flip-back Watagate ^1H NMR spectra of the imino region of $\text{PPT}_{\text{distal}}$ and PPT_{prox} . (a) $\text{PPT}_{\text{distal}}$ at 10 °C: control (top trace), 1.0 equiv of NB (middle trace), and 2.0 equiv of NB (bottom trace). (b) PPT_{prox} at 30 °C: control (top trace), 1.0 equiv of NB (middle trace), and 2.0 equiv of NB (bottom trace). Dotted lines and arrows indicate the shift in position of the resonances.

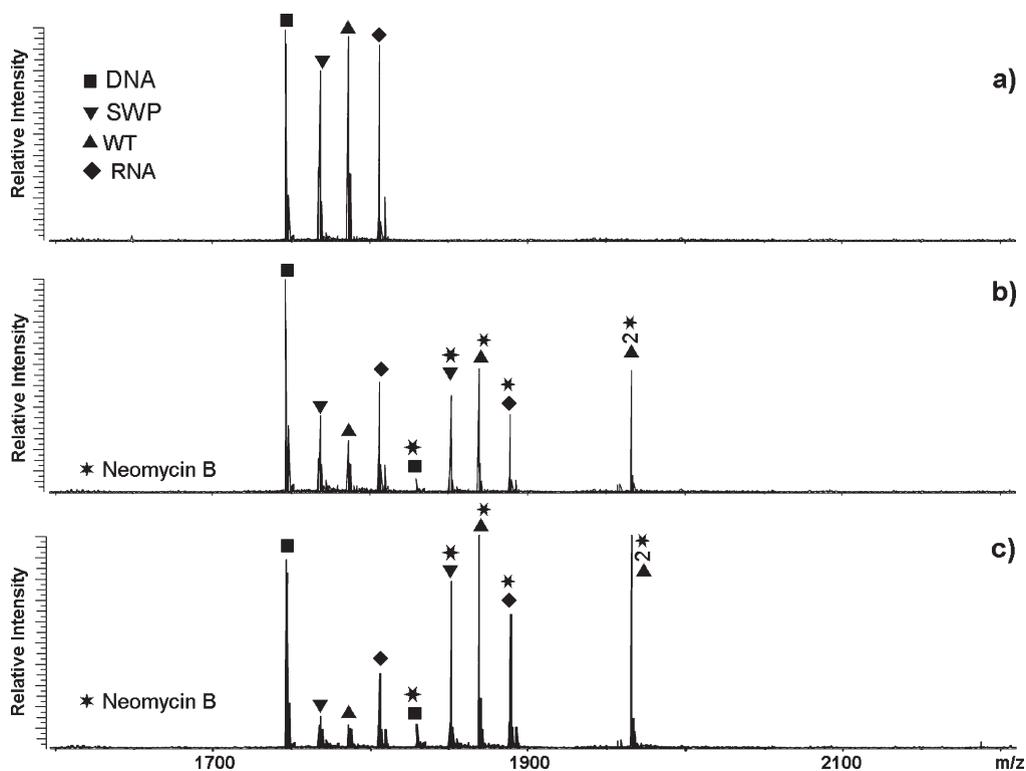


FIGURE 7: Nanospray FTICR mass spectra of (a) the initial equimolar mixture of the four PPT substrates ($5\ \mu\text{M}$ each), (b) products obtained after a 1-fold per substrate addition of NB, and (c) the assemblies after 10-fold ligand addition (see Materials and Methods): (■) PPT_{DNA} , (▼) PPT_{SWP} , (▲) PPT_{WT} , and (◆) PPT_{RNA} . Stars denote the positions of bound ligand. The high resolution and accuracy afforded by this analytical platform enabled unambiguous identification of the different species in solution according to their unique molecular masses (see also Figure 2).

patterns of assembly formation and initial substrate consumption are consistent with the following relative scale of binding affinities: $\text{PPT}_{\text{WT}} > \text{PPT}_{\text{RNA}} \approx \text{PPT}_{\text{SWP}} > \text{PPT}_{\text{DNA}}$.

NMR-detected NB titrations with PPT_{SWP} revealed that the binding sites were very similar to PPT_{WT} . Selective perturbation and/or broadening of chemical shifts was observed for the imino resonances of u3, u4, T6, and G7 (Figure 8a). These positions are analogous to the wild-type binding sites, suggesting NB

recognizes a helical distortion at the junctions at either end of the A·G·A tract. However, in contrast to PPT_{WT} which showed two-state binding kinetics, binding to the analogous sites in PPT_{SWP} exhibited intermediate to fast exchange kinetics on the NMR time scale. To test whether NB recognition of these sites is specific to RNA–DNA hybrids, NMR was used to monitor the binding of NB to PPT_{RNA} . In this titration, NB weakly bound to one site on PPT_{RNA} , indicated by perturbation of imino

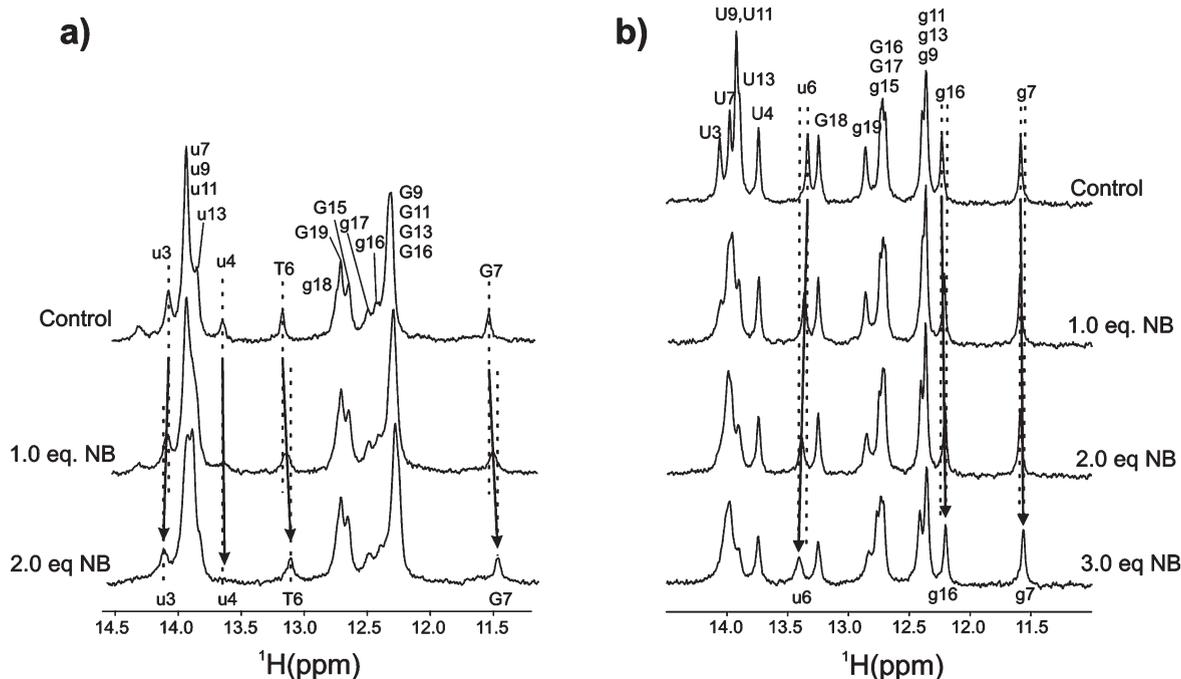


FIGURE 8: 1D water flip-back Watergate ^1H NMR spectra of the imino region of the PPT_{swp} and PPT_{RNA}. (a) PPT_{swp} at 10°C: control (top trace), 1.0 equiv of NB (middle trace), and 2.0 equiv of NB (bottom trace). (b) PPT_{RNA} at 30°C: control (top trace), 1.0 equiv of NB (top middle trace), 2.0 equiv of NB (bottom middle trace), and 3.0 equiv of NB (bottom trace). Dotted lines and arrows indicate the shifts in position of the resonances.

resonances g16, u6, and g7 (Figure 8b), which never saturated even at a 4:1 NB:PPT_{RNA} stoichiometry. Instead, additional nonspecific and sequential binding events occurred with the titration of each equivalent of NB, leading to general broadening of all the uridine (~14.0 ppm) and guanosine (~12.75 ppm) imino resonances.

Further analysis of chemical shift perturbations of base protons assigned by comparison of 2D NOESY experiments acquired for the PPT and PPT–NB complexes suggests a common minor groove binding mode. For example, if NB bound in the major groove, as has been commonly observed, a chemical shift perturbation would have been observed for the cytosine amino protons that project into the major groove. In contrast, if binding occurred in the minor groove, then greater chemical shift perturbation would be expected for the adenosine H2 resonances, which project into the minor groove. For PPT_{wt}, adenosine H2 protons of A15, a8, a17, and a18 all exhibited chemical shift perturbations, while the corresponding amino protons on C2, C5, C14, and c5 exhibited little chemical shift (Table 2 of the Supporting Information). Similar chemical shift perturbations were observed for PPT_{swp} and PPT_{RNA}, suggesting that the differences observed during the titrations were not due to differences in the mode of groove binding.

Effect of NB Binding on Ty3 PPT Helical Geometry. Since NB binding is known to shift helices toward a more A-form geometry (27–29), circular dichroism (CD) spectra were collected for PPT_{wt}, PPT_{swp}, and PPT_{RNA} to detect any shift in global helical structure (Figure 9). When the RNA strand in an RNA–DNA hybrid is purine rich, as is the case for PPT_{wt}, the helix, while neither a pure A- or B-form, is more stable and tends to adopt a more A-form helical geometry, with a characteristic minimum at 215 nm and a maximum at 270 nm (30). In contrast, if the RNA strand of an RNA–DNA hybrid is rich in pyrimidines (i.e., PPT_{swp}), the helical geometry shifts to a more mixed state and thus displays the characteristic minimum at 215 nm and

a maximum at 270 nm indicative of A-form geometry, as well as a local minimum at 240 nm characteristic of B-form geometry. CD analysis of PPT_{RNA}, as anticipated for an RNA duplex, shows the characteristic maxima and minima of a pure A-form helical geometry. Adding 5.0 equiv of NB to PPT_{wt} and PPT_{RNA} resulted in an increase in the minimum at 215 nm and the maximum at 270 nm indicative of a shift to a more pure A-form geometry. While the minimum at 215 nm increased upon addition of NB to PPT_{swp}, the maximum at 270 nm remained relatively unchanged, suggesting that a shift to a more A-form geometry is energetically less favorable for the PPT_{swp} hybrid and not required for NB binding.

Consistent with the CD observations, 2D DQF-COSY experiments conducted with PPT_{wt} and PPT_{swp} hybrids in the absence and presence of NB were consistent with a lack of a shift toward a more A-form geometry in the case of the PPT_{swp} substrate (Figure 2 of the Supporting Information). For a helix with A-form geometry, the ribose sugar pucker is in a C3'-endo conformation and the coupling constant between H1' and H2' resonances is less than 2.0 Hz. A cross-peak in the H1'–H2' region of the DQF-COSY spectrum is therefore normally not observed. For a hybrid A-form–B-form helix, ribose sugars can be a mixture of C2'-endo/C3'-endo pucker and exhibit H1'–H2' coupling constants that can be greater than 2.0 Hz and result in the detection of H1'–H2' COSY cross-peaks. The observation of sugar pucker switches in RNA–DNA hybrids has previously been described (9, 26). For the PPT_{swp} construct, the H1'–H2' region in the DQF-COSY spectrum for PPT_{swp} contained eight resonances, consistent with a helical geometry that has greater B-like character. In contrast, PPT_{wt} has fewer cross-peaks in this region, as expected for a hybrid with a purine rich RNA strand (26). Adding NB to PPT_{swp} perturbed the chemical shifts instead of changing the number of observed peaks, as might be expected if binding resulted in a structural shift toward a more A-form helical geometry.

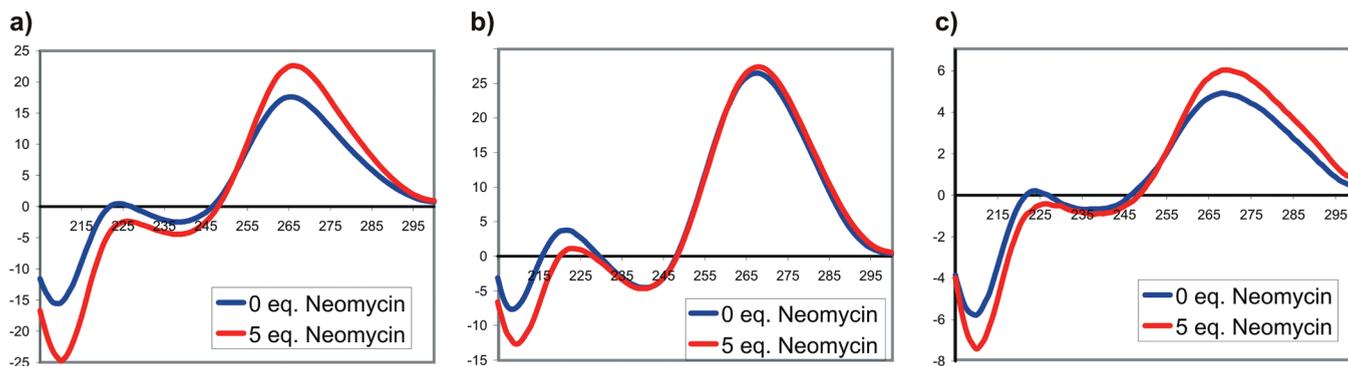


FIGURE 9: CD spectra with 5.0 equiv of NB (red trace) and without NB (blue trace) of (a) PPT_{RNA}, (b) PPT_{swp}, and (c) PPT_{WT} hybrid samples.

Thermodynamics of Binding of NB to Ty3 PPT_{WT}. On the basis of the similar mode of binding of NB to PPT_{WT}, PPT_{swp}, and PPT_{RNA}, we investigated the thermodynamic characteristics of these interactions by ITC to determine if there were discriminating signatures for each. The thermodynamic parameters for binding equilibria involving both protein (31) and nucleic acid (32) substrates are commonly determined using ITC. We measured these parameters for the PPT samples using typical NMR salt conditions [80 mM sodium chloride and 10 mM sodium phosphate (pH 7.0)] at 40 °C.

ITC binding isotherms (Figure 3 of the Supporting Information) in each case were monophasic and, after integration, could be fit to one set of binding equilibria to derive the corresponding values of the association constant (K_a), observed binding enthalpy (ΔH_{obs}), stoichiometry (N), and observed entropy ($T\Delta S_{\text{obs}}$), which are summarized in Table 1. This analysis reveals a remarkable similarity in thermodynamic characteristics among the three PPT samples. Each hybrid displays a NB:PPT stoichiometry of ~ 2.00 , which is consistent with the FTICR-MS and NMR data. In addition, there is very little variation in the association constants (average $K_a \sim 3.7 \times 10^5 \text{ M}^{-1}$), enthalpies, and entropic contributions. While the ITC data indicated that NB bound PPT_{RNA} with similar affinity and stoichiometry when compared to PPT_{WT} or PPT_{swp}, MS and NMR data indicated very little specificity for the PPT_{RNA} substrate. However, ITC did reveal that the entropic cost for binding to PPT_{RNA} was significantly higher and suggested a different mode of NB binding of PPT_{RNA} that likely involves the displacement of more Na^+ ions from the sugar–phosphate backbone.

ITC was also used to characterize the association of NB with the PPT_{distal}, PPT_{prox}, and PPT_{pol} hybrids, and association constants for each of these samples were also in the micromolar range (data not shown). In contrast to NMR which allowed discrimination between wild-type and mutant “end effects”, ITC failed to distinguish between these two binding modes. PPT_{pol}, for example, which does not contain either of the wild-type binding sites, still exhibited a stoichiometry of 2 because of association of NB with the ends of the helices. While further analysis to derive the degree of ligand protonation upon complexation, heat capacity changes, and intrinsic enthalpic and intrinsic entropic contributions could be used to distinguish these binding modes (27, 33–35), these approaches were not pursued since discrimination could be obtained using MS/MS and NMR methods.

DISCUSSION

The mechanism of PPT recognition by RT and its resistance to RNase H-mediated hydrolysis has been the subject of intensive

Table 1: Thermodynamic Parameters and Association Constants Obtained from ITC Titrations of the PPT_{WT}, PPT_{swp}, and PPT_{RNA} Hybrids with NB at 40 °C

sample	N	$K (\times 10^5 \text{ M}^{-1})$	ΔH (kcal/mol)	$T\Delta S$ (kcal/mol)
PPT _{WT}	2.60 ± 0.009	2.05 ± 0.04	-13.03 ± 0.0059	-5.41
PPT _{swp}	2.24 ± 0.011	3.06 ± 0.10	-16.5 ± 0.11	-8.64
PPT _{RNA}	2.07 ± 0.006	6.07 ± 0.17	-22.7 ± 0.091	-14.0

study. Despite the lack of sequence similarity between the PPT sequence of retroviruses and LTR-containing retrotransposons, a common molecular mechanism of enzyme recognition and processing has been postulated. In this study, we examined the interaction of NB with the Ty3 PPT, for which there is little structural information about RT binding. What is immediately apparent from our data is that, despite significant differences in PPT sequence, NB selectively binds regions in Ty3 PPT_{WT} (the 3' PPT–U3 junction and distal 5' RT thumb binding region) that match those observed in the analogous RNA–DNA HIV-1 hybrid (10).

In contrast to the HIV-1 PPT, the two distinct NB binding sites of the Ty3 PPT have very similar affinity and are occupied simultaneously in both the FTICR-MS and NMR titrations. Moreover, the sites display significant allosteric coupling, as the binding affinity is reduced for the remaining wild-type positions in the PPT_{distal} and PPT_{prox} constructs. If the sites were functionally independent, each should exhibit a binding affinity similar to that of PPT_{WT}. While coupling between NB binding sites on the Ty3 PPT was unexpected, propagation of structural signals has been observed along the rA·rG·rA symmetric steps of the PPT with both dF base analogue (9) and LNA-substituted PPTs (36).

Surprisingly, helical geometry plays only a modest role in modulating the selectivity of NB binding to the Ty3 PPT. Experiments with hybrids in which the RNA and DNA strands were exchanged (PPT_{swp}), both strands were RNA (PPT_{RNA}), or both were DNA (PPT_{DNA}) demonstrated that all bound NB with relatively high affinity and that the wild-type sites were maintained in the PPT_{swp} construct, despite a shift to a more B-form helical geometry in this substrate. NB binding to analogous positions in the wild-type and swapped constructs suggests that it recognizes a helical distortion at the junctions between the A·G·A tract and canonical sequence that can be adopted in both constructs. Nonetheless, the backbone composition did influence the binding kinetics at these sites with only the binding to sites on the PPT_{WT} showing higher, two-state binding kinetics by NMR.

In summary, results from this study of the Ty3 PPT further demonstrate the utility of NB as a molecular probe of deformable structural features in RNA–DNA hybrids. For both Ty3 and HIV-1 PPT hybrids, NB highlighted similar regions within the sequences where the cognate RT is predicted to make contacts. Nonetheless, differences in the NB binding modes for the HIV-1 and Ty3 PPT suggest subtle differences in how the two sites may function in concert to mediate recognition. While the exact structural details of the RT interaction are likely to be different for Ty3 and HIV-1, these data suggest that similar structural motifs are present in these PPTs, which may support a common mechanism for RT processing of PPTs in both retroviruses and LTR retrotransposons.

SUPPORTING INFORMATION AVAILABLE

A summary of the stoichiometries and relative binding affinities obtained by titrating mixtures of the four PPT substrates with increasing amounts of each individual ligand (Table 1), perturbations in chemical shifts assigned to the base proton of the PPT hybrid samples due to the addition and binding of NB (Table 2), nanospray-FTICR mass spectra of the Ty3 and HIV-1 PPT hybrids in the presence of NB (Figure 1), an overlay of the H1'–H2' region of a DQF-COSY spectrum applied to Ty3 PPT hybrids in the absence and presence of NB (Figure 2), and the ITC titration of selected PPTs with NB (Figure 3). This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Champoux, J. J. (1993) Roles of ribonuclease H in reverse transcription. In *Reverse Transcriptase* (Skalka, A. M., and Goff, S. P., Eds.) pp 103–118, Cold Spring Harbor Laboratory Press, Plainview, NY.
- Telesnitsky, A., and Goff, S. P. (1997) *Reverse transcriptase and the generation of retroviral DNA*, Cold Spring Harbor Laboratory Press, Plainview, NY.
- Wilhelm, M., and Wilhelm, F. X. (2001) Reverse transcription of retroviruses and LTR retrotransposons. *Cell. Mol. Life Sci.* 58, 1246–1262.
- Rausch, J. W., and Le Grice, S. F. (2004) 'Binding, bending and bonding': Polypurine tract-primed initiation of plus-strand DNA synthesis in human immunodeficiency virus. *Int. J. Biochem. Cell Biol.* 36, 1752–1766.
- Bibillo, A., Lener, D., Tewari, A., and Le Grice, S. F. (2005) Interaction of the Ty3 reverse transcriptase thumb subdomain with template-primer. *J. Biol. Chem.* 280, 30282–30290.
- Sarafianos, S. G., Das, K., Tantillo, C., Clark, A. D. Jr., Ding, J., Whitcomb, J. M., Boyer, P. L., Hughes, S. H., and Arnold, E. (2001) Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J.* 20, 1449–1461.
- Jacobo-Molina, A., Ding, J. P., Nanni, R. G., Clark, A. D., Lu, X. D., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., Hizi, A., Hughes, S. H., and Arnold, E. (1993) Crystal structure of human-immunodeficiency-virus type-1 reverse transcriptase complexed with double stranded DNA at 3.0 angstrom resolution shows bend DNA. *Proc. Natl. Acad. Sci. U.S.A.* 90, 6320–6324.
- Huang, H. F., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: Implications for drug resistance. *Science* 282, 1669–1675.
- Yi-Brunozzi, H. Y., Brinson, R. G., Brabazon, D. M., Lener, D., Le Grice, S. F., and Marino, J. P. (2008) High-resolution NMR analysis of the conformations of native and base analog substituted retroviral and LTR-retrotransposon PPT primers. *Chem. Biol.* 15, 254–262.
- Turner, K. B., Brinson, R. G., Yi-Brunozzi, H. Y., Rausch, J. W., Miller, J. T., Le Grice, S. F., Marino, J. P., and Fabris, D. (2008) Structural probing of the HIV-1 polypurine tract RNA:DNA hybrid using classic nucleic acid ligands. *Nucleic Acids Res.* 36, 2799–2810.
- Aleksandrov, M., Gall, L., Krasnov, V., Nikolaev, V., and Pavlenko, V. (1984) Extraction of ions from solution under atmospheric pressure: A method of mass spectrometric analysis of bioorganic compounds. *Dokl. Akad. Nauk* 277, 379–383.
- Yamashita, M., and Fenn, J. B. (1984) Electrospray Ion-source: Another variation on the free-jet theme. *J. Phys. Chem.* 88, 4451–4459.
- Comisarow, M. B. (1978) Theory of Fourier-transform ion cyclotron resonance mass spectrometry 2. Signal modeling for ion-cyclotron resonance. *J. Chem. Phys.* 69, 4097–4104.
- Marshall, A. G., Hendrickson, C. L., and Jackson, G. S. (1998) Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.* 17, 1–35.
- Hagan, N. A., and Fabris, D. (2007) Dissecting the protein-RNA and RNA-RNA interactions in the nucleocapsid-mediated dimerization and isomerization of HIV-1 stemloop 1. *J. Mol. Biol.* 365, 396–410.
- Hagan, N., and Fabris, D. (2003) Direct mass spectrometric determination of the stoichiometry and binding affinity of the complexes between nucleocapsid protein and RNA stem-loop hairpins of the HIV-1 psi-recognition element. *Biochemistry* 42, 10736–10745.
- Turner, K. B., Hagan, N. A., Kohlway, A. S., and Fabris, D. (2006) Mapping noncovalent ligand binding to stemloop domains of the HIV-1 packaging signal by tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* 17, 1401–1411.
- Turner, K. B., Hagan, N. A., and Fabris, D. (2006) Inhibitory effects of archetypical nucleic acid ligands on the interactions of HIV-1 nucleocapsid protein with elements of Psi-RNA. *Nucleic Acids Res.* 34, 1305–1316.
- de Koning, L. J., Nibbering, N. M. M., van Orden, S. L., and Laukien, F. H. (1997) Mass selection of ions in a Fourier transform ion cyclotron resonance trap using correlated harmonic excitation fields (CHEF). *Int. J. Mass Spectrom.* 165, 209–219.
- Gauthier, J. W., Trautman, T. R., and Jacobson, D. B. (1991) Sustained off-resonance irradiation of collision-activated dissociation involving Fourier transform mass spectrometry: Collision dissociation technique that emulates infrared multiphoton dissociation. *Anal. Chim. Acta* 246, 211–225.
- Krugh, T. R., and Reinhardt, C. G. (1975) Evidence for sequence preferences in intercalative binding of ethidium-bromide to dinucleoside monophosphates. *J. Mol. Biol.* 97, 133–162.
- Hardwick, J. M., Vonsprecken, R. S., Yielding, K. L., and Yielding, L. W. (1984) Ethidium binding sites on plasmid DNA determined by photoaffinity labeling. *J. Biol. Chem.* 259, 1090–1097.
- Kopka, M., Yoon, C., Goodsell, D., Pjura, P., and Dickerson, R. (1985) The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc. Natl. Acad. Sci. U.S.A.* 82, 1376–1380.
- Hunt, D. F., Shabanowitz, J., Yates, J. R., Zhu, N. Z., Russell, D. H., and Castro, M. E. (1987) Tandem quadrupole Fourier transform mass spectrometry of oligopeptides and small proteins. *Proc. Natl. Acad. Sci. U.S.A.* 84, 620–623.
- McLafferty, F. W. (1981) Tandem Mass Spectrometry. *Science* 214, 280–287.
- Yi-Brunozzi, H. Y., Brabazon, D. M., Lener, D., Le Grice, S. F., and Marino, J. P. (2005) A ribose sugar conformational switch in the LTR-retrotransposon Ty3 polypurine tract-containing RNA/DNA hybrid. *J. Am. Chem. Soc.* 127, 16344–16345.
- Li, T. K., Barbieri, C. M., Lin, H. C., Rabson, A. B., Yang, G., Fan, Y., Gaffney, B. L., Jones, R. A., and Pilch, D. S. (2004) Drug targeting of HIV-1 RNA·DNA hybrid structures: Thermodynamics of recognition and impact on reverse transcriptase-mediated ribonuclease H activity and viral replication. *Biochemistry* 43, 9732–9742.
- Robinson, H., and Wang, A. H. (1996) Neomycin, spermine and hexaamminecobalt (III) share common structural motifs in converting B- to A-DNA. *Nucleic Acids Res.* 24, 676–682.
- Arya, D. P., Xue, L., and Willis, B. (2003) Aminoglycoside (neomycin) preference is for A-form nucleic acids, not just RNA: Results from a competition dialysis study. *J. Am. Chem. Soc.* 125, 10148–10149.
- Gyi, J. I., Conn, G. L., Lane, A. N., and Brown, T. (1996) Comparison of the thermodynamic stabilities and solution conformations of DNA·RNA hybrids containing purine-rich and pyrimidine-rich strands with DNA and RNA duplexes. *Biochemistry* 35, 12538–12548.
- Perozzo, R., Folkers, G., and Scapozza, L. (2004) Thermodynamics of protein-ligand interactions: History, presence, and future aspects. *J. Recept. Signal Transduction* 24, 1–52.
- Haq, I., Jenkins, T. C., Chowdhry, B. Z., Ren, J. S., and Chaires, J. B. (2000) Parsing free energies of drug-DNA interactions. *Methods Enzymol.* 323, 373–405.
- Barbieri, C. M., Li, T. K., Guo, S., Wang, G., Shalloo, A. J., Pan, W., Yang, G., Gaffney, B. L., Jones, R. A., and Pilch, D. S. (2003)

- Aminoglycoside complexation with a DNA·RNA hybrid duplex: The thermodynamics of recognition and inhibition of RNA processing enzymes. *J. Am. Chem. Soc.* 125, 6469–6477.
- (34) Kaul, M., Barbieri, C. M., Kerrigan, J. E., and Pilch, D. S. (2003) Coupling of drug protonation to the specific binding of aminoglycosides to the A site of 16 S rRNA: Elucidation of the number of drug amino groups involved and their identities. *J. Mol. Biol.* 326, 1373–1387.
- (35) Barbieri, C. M., Srinivasan, A. R., and Pilch, D. S. (2004) Deciphering the origins of observed heat capacity changes for aminoglycoside binding to prokaryotic and eukaryotic ribosomal RNA a-sites: A calorimetric, computational, and osmotic stress study. *J. Am. Chem. Soc.* 126, 14380–14388.
- (36) Dash, C., Marino, J. P., and Le Grice, S. F. (2006) Examining Ty3 polypurine tract structure and function by nucleoside analog interference. *J. Biol. Chem.* 281, 2773–2783.