Affinities of Packaging Domain Loops in HIV-1 RNA for the Nucleocapsid Protein†

Michael F. Shubsda, Andrew C. Paoletti, Bruce S. Hudson,* and Philip N. Borer*

Department of Chemistry, Graduate Program in Structural Biology, Biochemistry, and Biophysics, Syracuse University, Syracuse, New York 13244-4100

Received December 14, 2001; Revised Manuscript Received February 9, 2002

ABSTRACT: To design anti-nucleocapsid drugs, it is useful to know the affinities the protein has for its natural substrates under physiological conditions. Dissociation equilibrium constants are reported for seven RNA stem-loops bound to the mature HIV-1 nucleocapsid protein, NCp7. The loops include SL1, SL2, SL3, and SL4 from the major packaging domain of genomic RNA. The binding assay is based on quenching the fluorescence of tryptophan-37 in the protein by G residues in the single-stranded loops. Tightly bound RNA molecules quench nearly all the fluorescence of freshly purified NCp7 in 0.2 M NaCl. In contrast, when the GGAG-tetraloop of tight-binding SL3 is replaced with UUCG or GAUA, quenching is almost nil, indicating very low affinity. Interpreting fluorescence titrations in terms of a rapidly equilibrating 1:1 complex explains nearly all of the experimental variance for the loops. Analyzed in this way, the highest affinities are for 20mer SL3 and 19mer SL2 hairpin constructs ($K_d = 28 \pm 3$ and 23 ± 2 nM, respectively). The 20mer stem-UUCG-loop and GAUA-loop constructs have <0.5% of the affinity for NCp7 relative to SL3. Affinities relative to SL3 for the other stem-loops are the following: 10% for a 16mer construct to model SL4, 30% for a 27mer model of the 9-residue apical loop of SL1, and 20% for a 23mer model of a 1×3 asymmetric internal loop in SL1. A 154mer construct that includes all four stem-loops binds tightly to NCp7, with the equivalent of three NCp7 molecules bound with high affinity per RNA; it is also possible that two strong sites and several weaker ones combine to give the appearance of three strong sites.

The nucleocapsid domain (NC)[1] of the *gag* and *gag-pol* polyproteins of the HIV-1 virus is known to play a significant role in the life cycle of the retrovirus (*1*−*4*). NC is involved in the selection of unspliced genomic RNA from the host cell during replication. The polyproteins are cleaved after the virus particle buds off the host, and the mature 55 amino acid nucleocapsid protein (NCp7) is produced. Both the NC domain and mature NCp7 have two CCHC zinc finger motifs common to retroviral nucleocapsid proteins (*5*), so NCp7 is frequently used as a model for the NC domain. NCp7 is involved in several other viral functions, including dimerization of the viral genome (*6*−*10*), coating and protecting the viral genome in mature viral particles (*11*), and annealing the tRNA primer to begin reverse transcription (*12*).

The primary packaging domain within the 5′-leader of the HIV-1 viral genome includes four stem-loop structures, shown in Figure 1. Other overall structures have been suggested for this region, but the four stem-loops are common to nearly all (*13*−*18*). This region of the RNA genome probably interacts with multiple NC domains during packaging—about 1500 *gag* and *gag-pol* precursors are enclosed in each virus particle (*19*). High-resolution structures have been published for SL1 (*20*−*24*), SL2 (*25*), SL3 (*16, 17*), and SL4 (*26*). SL3 is the major recognition site for NC (*15, 27*−*34*); however, NCp7 also binds to SL1, SL2, SL4, and mutated 5′-leaders lacking SL3 (*15, 28, 31, 35*). HIV-1 RNA with SL3 deletions is packaged at ∼10% of the efficiency of the wild-type (*28*). All of the high-resolution complexes of nucleic acids with NC-related peptides have G residues bound in nearly identical configurations in a hydrophobic cleft of the zinc fingers (*35*−*38*); strong interaction sites appear to require unpaired G residues (*37*). It has been shown that SL3 and SL2 have similar affinity for NCp7 by isothermal titration calorimetry (*35*). Studies on NCp7 binding to nucleic acids have been performed by a variety of other techniques including polyacrylamide gel electrophoresis (*39, 40*), NMR (*35, 36, 38*), surface plasmon resonance (*41, 42*), and fluorescence (*4, 42*−*44*). These studies have shown little consensus in terms of the binding stoichiometry and affinity, being conducted with a wide variety of RNA and DNA sequences and ionic conditions.

† Supported in part by Grant 02740-30-RGT from the American Foundation for AIDS Research (to P.N.B. and B.S.H.), by NIH Grant GM32691 (to P.N.B.), and by Syracuse University.

* Corresponding author. P.N.B. e-mail: pnborer@syr.edu, phone: 315 443 5925. B.S.H. e-mail: bshudson@syr.edu, phone: 315 443 5805, fax: 315 443 4070.

[1] Abbreviations: NCp7, the 55 amino acid nucleocapsid protein of HIV-1, human immunodeficiency virus 1; SL1, SL2, SL3, and SL4, short stem-loop segments of RNA from the 5′-leader region of HIV-1; NC, nucleocapsid; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; PEG-8000, poly(ethylene glycol) average molecular weight 8000; P, protein; R, RNA; B, RNA/protein complex; concentrations are denoted by boldface italic font, e.g., $\boldsymbol{P}_t$ indicates the total concentration of protein; W, tryptophan; W37, tryptophan residue 37 of NCp7; W|G, tryptophan|guanine stacking in the RNA/protein complex; ITC, isothermal titration calorimetry.

10.1021/bi016045+ CCC: $22.00
Published on Web 03/28/2002

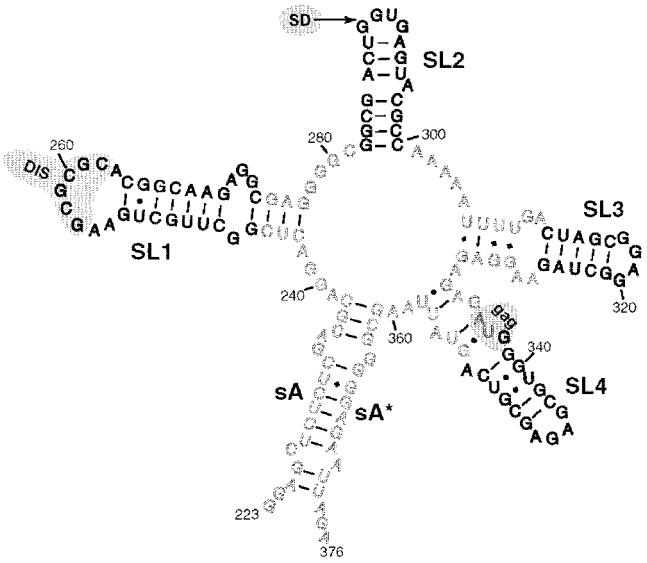

FIGURE 1: A 154mer representing the major packaging domain of HIV-1 genomic RNA [adapted from (*13*)]. Dark fonts mark stem-loops examined in this work.

FIGURE 2: Stem-loops used in this study. For SL3-UUCG and -GAUA, the GGAG tetraloop is replaced with the stated loop sequence.

The present work gives a unified view of NCp7 binding to individual RNA stem-loops from the domain of the 5′-leader that provides most of the packaging specificity in HIV-1. The reference for comparison is the primary packaging determinant SL3. Near-physiological ionic conditions were chosen; almost full quenching of tryptophan-37 (W37) fluorescence due to SL3 binding occurs at ionic strengths above 0.15 M. The second finger in the NCp7 sequence contains W37, which stacks strongly upon G residues in the complexes (*35*, *36*, *38*, *44*). Fluorescence titrations were fitted to a theoretical 1:1 model where it was assumed that the fluorescence of the complex is zero. The dissociation constants presented here were determined by finding the K_d value that gives the best fit of the model to the measured fluorescence data.

The significance of the work is in mapping the affinities of the wild-type interaction sites for the NC domain in the major packaging domain of the 5′-leader RNA. This is important in designing and testing anti-nucleocapsid drugs. Because NC participates in many viral functions, such drugs could exert effects at several stages of the viral life cycle. The current results also add to our understanding of RNA−protein interactions, highlighting problems that may occur when these studies are conducted at low ionic strength. Finally, we demonstrate that multiple NCp7 proteins interact with the major packaging domain, and that a G-X-G loop sequence in the RNA is not required for reasonable affinity.

EXPERIMENTAL PROCEDURES

NCp7 Preparation. The 55 amino acid NCp7 protein of the pNL4-3 isolate of HIV-1 was overexpressed in BL21-(DE3)pLysS *E. coli* containing the pRD2 plasmid that was a generous gift from Dr. Michael Summers (University of Maryland, Baltimore County). The cells were grown and harvested and the protein was isolated using a protocol similar to that of Lee et al. (*45*). Cell lysis was accomplished by using the B-PER II reagent (Pierce) and grinding the cells with a mortar and pestle. Preparations from 0.5 L cultures were usually conducted, although a 2 L preparation was used to isolate protein to verify zinc finger formation by NMR (*37*). Protein was purified to homogeneity as judged by SDS−PAGE (15% gels). The NCp7 concentration was determined by UV in the storage buffer (50 mM Tris, pH 8.0, 10% glycerol, 0.1 M NaCl, 0.1 mM $ZnCl_2$, 10 mM β-mercaptoethanol), using an extinction coefficient at 280 nm of 6050 cm^{-1} M^{-1} (*46*). As a positive control for each new protein preparation, fluorescence titrations with SL3 were repeated daily.

RNA. The RNA stem-loops used in this study, shown in Figure 2, were purchased from Dharmacon Research Inc. as unpurified 2′-ACE-protected oligonucleotides. Using procedures recommended by Dharmacon (*47*), the RNA was purified by HPLC on a Dionex 4 × 250 mm DNAPac PA-100 column, using a 5−150 mM sodium perchlorate gradient in 10 mM Tris, 2% acetonitrile, pH 8, buffer, followed by desalting with a Waters reverse-phase C18 Sep Pak cartridge. The purified RNA was then deprotected and dried. RNA samples were redissolved in Milli-Q purified water (Millipore Corp.) Prior to using the RNA in titrations, samples were disaggregated by heating briefly to 85 °C. The concentration of the RNA was determined by UV using an extinction coefficient of 160 000 cm^{-1} M^{-1} (160 A_{260} units/μmol) for SL3 (*39*) in oligo buffer (5 mM sodium phosphate, pH 7.0, 25 mM sodium chloride). Extinction coefficients for the other RNA sequences were estimated at ($N \times 8$) A_{260} units/μmol, where N = chain length.

Protein Binding Assay. A 1 mL aliquot of 0.3 μM NCp7 in the fluorescence buffer [5 mM sodium phosphate, pH 7.0, 200 mM sodium chloride, 0.1 mM zinc chloride, and 0.01% poly(ethylene glycol) (PEG-8000; to prevent protein from sticking to the cuvette walls)] was placed in a 10 × 10 mm path length quartz fluorescence cuvette (NSG precision cells) along with an 8 × 1.5 mm stir bar. The initial fluorescence was measured at 350 nm with a Hitachi F-4010 fluorescence spectrophotometer using a 290 nm excitation wavelength, a 5 nm excitation band-pass, and either a 3 or a 5 nm emission band-pass. The stir bar was small enough not to interfere with the light path of the spectrometer. Then aliquots of a ~100 μM RNA solution in oligo buffer were added to the protein sample to perform the titration. Addition of RNA solution results in less than 1% change in the ionic environment in the cuvette. After the addition of each aliquot of RNA to the protein, the solution in the cuvette was mixed briefly and the fluorescence measured.

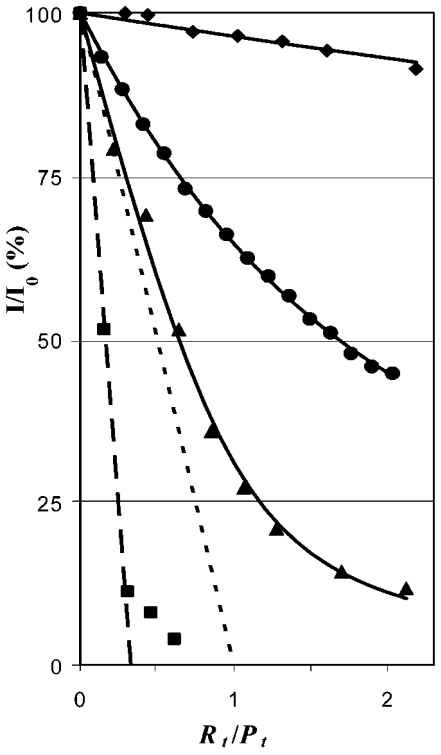

FIGURE 3: Titrations of 0.3 μM NCp7 protein (P) with RNA (R) packaging signal components. Experimental data for (◆) SL3-UUCG, (●) SL4, (▲) SL3, and (■) 154mer full domain; solid lines represent calculated fits for 1:1 complexes (see text). Short dashed line, ideal R_1P_1 complex; long dashes, ideal R_1P_3 complex.

The fluorescence titration curve was fit to a model assuming 1:1 stoichiometry for the ratio of NCp7 to bound nucleic acid in the equilibrium P + R = B, where P is the protein, NCp7, R is the RNA, and B is the bound complex. The dissociation constant, K_d, of the equilibrium is $K_d = \boldsymbol{PR}/\boldsymbol{B}$ (concentrations are denoted by boldface italic font). The total concentrations of protein and nucleic acid are $\boldsymbol{P}_t = \boldsymbol{P} + \boldsymbol{B}$ and $\boldsymbol{R}_t = \boldsymbol{R} + \boldsymbol{B}$. The titration curves are fitted to the equation:

$$(I - I_\infty)/I_0 = \{(\boldsymbol{R}_t - \boldsymbol{P}_t + K_d) + [(\boldsymbol{R}_t - \boldsymbol{P}_t + K_d)^2 + 4\boldsymbol{P}_t K_d]^{1/2}\}/2\boldsymbol{P}_t$$

where I_0 is the intensity at $\boldsymbol{R} = 0$ and I_∞ is the limiting intensity at saturation; in this work, we set $I_\infty = 0$ unless indicated otherwise. The Goal Seek function in Microsoft Excel was used to find the value of K_d for each titration that minimizes the sum of the squared deviations between the theoretical and experimental curves.

RESULTS

Figure 3 shows representative titration curves for quenching W37 fluorescence in the NCp7 protein upon binding RNA for three of the stem-loops, and the full 154mer packaging domain. The titration curves illustrated are for SL3, SL4, and SL3-UUCG (a 20mer with the SL3 stem and a UUCG tetraloop replacing the wild-type GGAG). These curves typify sequences with high, intermediate, and low affinities. The fluorescence intensity decreases as RNA is added to the protein. Fits with $I_\infty = 0$ gave the dissociation constant values in Table 1. The affinity relative to that of SL3 is also shown. (Supporting Information includes titration data for SL2, SL1a, and SL1i.)

Table 1: Dissociation Constant, Relative Affinity, and Decrease in Fluorescence for RNA−NCp7 Complexes

RNA	K_d[a] (nM)	RA[b] (%)	ΔF[c] (%)
SL1a	100 ± 10	30	75
SL1i	140 ± 20	20	70
SL2	23 ± 2	120	95
SL3	28 ± 3	100	90
SL4	320 ± 30	10	60
SL3-UUCG	~7500	~0.4	10
SL3-GAUA	~10000	~0.3	5

[a] Errors estimated as standard deviations of the mean; multiply by 1.96 for 95% confidence limits. [b] Affinity for NCp7 relative to SL3 (rounded to one significant digit). [c] 100 × (1− I/I_0) at $\boldsymbol{R}_t/\boldsymbol{P}_t = 2$.

The high-affinity RNA loops studied, SL3 and SL2, quench ~90% of the Trp fluorescence when the molar amount of RNA is twice that of protein. The binding profiles for SL2 and SL3 nearly approach that for a 1:1 complex with an infinite binding constant, $(K_d)^{-1}$. The 154mer (Figure 3) and higher affinity stem-loops (manuscript in preparation) quench the fluorescence to essentially the background of the buffer. This confirms that the fluorescence from the complex is negligible, so mechanisms other than W|G stacking need not be considered to account for the quenching. The K_d values providing the best fits were 23 nM for SL2 and 28 nM for SL3 (see Table 1).

The low-affinity RNA hairpins reported, SL3-UUCG and SL3-GAUA, show <10% decrease in fluorescence intensity, indicating that these hairpins bind very little to the W-site in NCp7. SL3-GAUA has the lowest affinity for NCp7 of the ~40 SL3 variants we have examined to date (manuscript in preparation); such a molecule is useful in showing that inner filter effects are negligible and for competition experiments. The apparent affinities of NCp7 for these SL3 variants are 20−400 times less than for the wild-type loops assuming that $I_\infty = 0$ (Table 1). Adding SL3-GAUA did not significantly change the fluorescence of mixtures of NCp7 and SL3 (data not shown). Thus, SL3-GAUA does not compete with wild-type SL3 for the W|G binding site. This has important implications for our work as it implies that nonspecific binding to the stem or loop is insignificant at 0.2 M NaCl. Also, since the UUCG tetraloop has a low affinity for the W|G site, this loop could be used to construct a stable hairpin (*48*) to determine the protein affinity for the internal loop of SL1 (see SL1i in Discussion).

The other hairpins studied (SL1a, SL1i, and SL4) have intermediate affinity, with K_d values about 5−10 times less than for SL3 and SL2. Under the titration conditions used here, the fluorescence falls to 60−75% of I_0 at $\boldsymbol{R}_t/\boldsymbol{P}_t = 2$. The SL1a 27mer (Figure 2) models the apical part of SL1; a 23mer deleting the terminal unpaired bases has a similar K_d value (data not shown). The SL1i 23mer has an apical UNCG tetraloop, and a 1 × 3 asymmetric internal loop to model the unpaired G residues in the SL1 stem. The stem-loops exhibit UV-absorbance vs temperature profiles with T_m >70 °C except for SL4 (~55 °C) in 0.2 M NaCl (data not shown). The K_d values determined for SL1a, SL1i, and SL4 are listed in Table 1.

The accuracy of the K_d determinations was assessed for 11 SL3 datasets all at $\boldsymbol{P} = 0.3$ μM, and a triplicate analysis

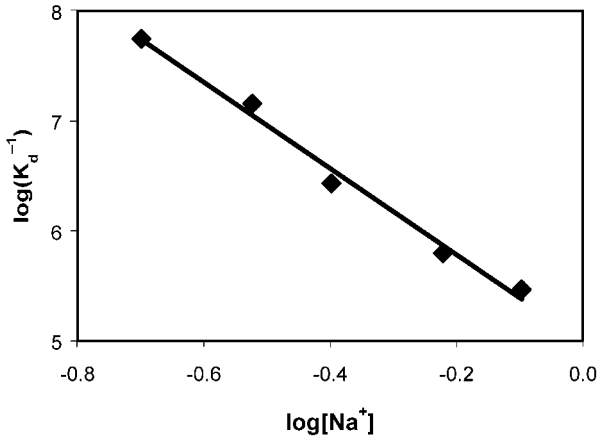

FIGURE 4: Salt-dependence of binding for SL3 to NCp7.

of NCp7 binding to SL4, where the protein concentration was varied over a factor of 4. Standard deviations of the mean were 9.8% for SL3 and 8.4% for SL4. Replicate error analyses were not conducted for the other RNA molecules, but error estimates of ±10% are reasonable for all of the entries in Table 1.

The salt-dependence of binding was determined for SL3 over the range from 0.2 to 0.8 M NaCl. The data are presented in Figure 4. Increasing salt concentration causes a dramatic decrease in K_d. It was not possible to obtain reproducible titrations below 0.15 M NaCl.

DISCUSSION

Analysis of the Model. The fluorescence emission of NCp7 is quenched when W37 is in close proximity to nucleic acid bases (*35*, *36*, *44*). Thus, fluorescence can be a very sensitive assay when there is strong interaction of a nucleobase with the second zinc finger in the protein. The interaction should be nearly optimal when there is tryptophan|guanine stacking, as in the high-resolution NCp7−nucleic acid complexes observed to date (*35*, *36*, *38*). In the studies with RNA loops, a single-stranded G-X-G sequence makes hydrogen-bond and hydrophobic contacts, where each G binds a separate zinc finger in NCp7 (*35*, *36*). In addition, there are primarily electrostatic interactions between an N-terminal 3-10 helix and the RNA stem. Thus, the binding site covers a substantial part of this small protein's surface.

Binding isotherms reported here for NCp7 and the short stem-loops were interpreted by assuming a rapid equilibrium between the free species and a 1:1 complex, where the latter has fully quenched fluorescence. Only a single parameter, K_d, was necessary to explain the systematic variance in titrations of individual stem-loops, even for the weakly bound species. This analysis requires that there be (1) negligible residual fluorescence in the complex ($I_\infty \cong 0$) and (2) only weak second-site binding compared to binding at the W37 site. With regard to the first requirement, highly efficient quenching occurs for SL3, SL2, and the 154mer (up to 96% at $\boldsymbol{R}_t/\boldsymbol{P}_t = 2$, Figure 3). The second requirement is addressed by the experiments with SL3-UUCG and -GAUA, which apparently have a very small affinity for NCp7. SL3-GAUA exhibits the lowest affinity for NCp7 of more than 40 RNA and DNA loops we have examined to date (manuscript in preparation). The small degree of quenching indicates that no significant interaction occurs with W37 under the conditions used here. Furthermore, the competition experiments show that second-site binding either is negligible or at least does not affect binding at the primary site. It is worth emphasizing that these experiments were conducted in 0.2 M NaCl, which will reduce nonspecific electrostatic interactions compared to the low ionic strengths used in many previous studies (NCp7·Zn_2 has a +9 charge and SL3 has −19 at neutral pH). Titrations were usually not extended to saturation for low- and intermediate-affinity complexes because of concerns that a large excess of RNA may favor unusual stoichiometries. Accurate binding constants for low-affinity complexes are difficult to obtain unless one can be sure that the nature of the interaction remains constant over a concentration range that is 50−1000 times larger than used in this study. Furthermore, our primary interest is to characterize high-affinity complexes to aid in designing strategies to block their formation.

Strong Binding Sites. Of the short stem-loops examined here, SL2 and SL3 have the strongest affinity for NCp7 and nearly complete fluorescence quenching (>90% at $\boldsymbol{R}_t/\boldsymbol{P}_t = 2$). These two cases support the 1:1 model, closely following the limiting theoretical line for one protein per RNA (Figure 3, dotted line), that corresponds to $K_d = 0$ for the given stoichiometry. The relative position of the measurements compared to the limiting line has important implications for the model to be used. If the data points fall below the line, a stoichiometry of more than 1 protein per RNA is indicated. This did in fact occur for the 154mer RNA, which includes all four of the stem-loops; the rapid fluorescence decrease in its titration is consistent with at least three proteins binding per RNA (RP_n, $n \geq 3$; the fluorescence intensity extrapolates to zero at $\boldsymbol{R}_t/\boldsymbol{P}_t \cong 0.33$). It is possible that two strong sites and several weaker ones combine to give the appearance of three strong sites. It should be cautioned that R_nP complexes involving two or more RNA molecules per protein ($n \geq 2$) could exhibit an apparent 1:1 stoichiometry in fluorescence titrations. After the W37 site is saturated with one RNA molecule, no additional quenching would be observed upon binding additional RNA oligomers.

Intermediate Affinity Sites. While data points for SL1a, SL1i, and SL4 lie further above the limiting 1:1 line than SL3, they still fit well to a model with 1:1 stoichiometry. The SL1 stem-loops show a similar degree of quenching (~70% at $\boldsymbol{R}_t/\boldsymbol{P}_t = 2$) to the strong sites. While SL4 has about 60% quenching at $\boldsymbol{R}_t/\boldsymbol{P}_t = 2$, the fit obtained from the 1:1 model is reasonable (as shown in Figure 3), with an order of magnitude less affinity than SL3. A titration for SL4 was extended to $\boldsymbol{R}_t/\boldsymbol{P}_t = 10$, at which $I = 10\%$ of I_0, and similar values of K_d for the shorter titrations (data not shown). Instead, if one assumed that SL4 saturated at $I_\infty = 25\%$ of I_0 in Figure 3, the best fit of K_d increases the affinity from 320 nM (Table 1) to ~160 nM, and deviations from the best-fit line become noticeably worse. Therefore, it is generally not necessary to extend the titrations to saturation as long as one observes that the tightest binding oligomers have $I_\infty \cong 0$. The SL1i loop is the first demonstration that a linear G-X-G loop sequence is not required for reasonably tight binding to NCp7.

A UNCG tetraloop was substituted for the wild-type apical loop in the SL1i model. As might be expected, the SL3-UUCG tetraloop was shown to bind very weakly to NCp7, since the U and G are known to form an unusual sheared

base-pair in these loops (*49*). That leaves only the central U-C of the loop to interact with NCp7. The UACG loop chosen for the SL1i construct is slightly different, but also has low affinity for NCp7. We determined that $K_d = 1300$ for the SL1i 23mer when all of the internal loop bases are adenines (see Figure 2).

Comparison with Literature K_d Values. There is little consistency among the published values for the binding of NCp7 to the four individual RNA stem-loops in the major packaging domain, and there are differences from the K_d values reported here. A major source of the discrepancy may lie in the salt-dependence of the NCp7−RNA fluorescence changes. The salt dependence for binding DNA oligomers to a longer, 72-residue version of NCp7 has been described by Vuilleumier et al. (*43*). They reported that affinity decreases 2000-fold over the range from 0.1 to 1 M NaCl for an SL3-DNA analogue. A somewhat larger effect was obtained for NCp7 binding to SL3-RNA as a function of NaCl concentration (see Figure 4). When the data are analyzed as suggested by Vuilleumier et al. (*43*) as

$$\log(K_d^{-1}) = -m'\Psi \log [Na^+] + \log[K_d^{-1}(1\ M)]$$

the plot has a straight line with slope $= -m'\Psi$, where m' is the number of ion pairs between R and P, and Ψ is the fraction of sodium ions bound on average to R in the absence of P. If Ψ is estimated at 0.7−0.8 (*43*, *50*), there are five to six ion pairs in the SL3−NCp7 complex. In remarkable agreement, the NMR structure predicts six salt-bridge interactions between basic side chains of the protein and the RNA phosphates (*36*).

However, it was difficult to replicate SL3−NCp7 titrations below ∼0.15 M NaCl even though at 0.2 M NaCl SL3 titrations are highly reproducible. In some low-salt titrations with SL3, there was a marked increase in the residual fluorescence at $\boldsymbol{R}_t/\boldsymbol{P}_t = 2$ (only 20−40% quenching), as well as a decrease in I_0 (about 60−70% of the value in 0.2 M NaCl). A model involving two or more binding sites with distinct equilibrium constants is required to fit these low-salt titration curves. The low-salt regime may be dominated by nonspecific interactions between these highly charged molecules. At early stages in the titrations, the protein concentration greatly exceeds the RNA concentration. This could favor transient formation of RP_2 complexes for which solubility may depend on ionic conditions (such complexes would be near electrical neutrality). Experiments in progress are designed to investigate this phenomenon. Nevertheless, the simple expedient of comparing affinities at 0.2 M NaCl should be sufficient to distinguish trends in affinity that are helpful in designing anti-NC agents.

Other differences in reported K_d values may be attributed to differences in analytical technique, the stoichiometric model chosen to analyze the data, and our observation that NCp7 protein samples degrade over time. Protein samples used within a few days of isolation from *E. coli* gave consistent values of I_0 and I_∞, in contrast to samples that had been stored for 1 week, even in the cold and the presence of Zn^{2+} and reducing agent. A particular symptom is the increase in I_∞ with time in SL3 titrations. The source of this loss in activity is currently under investigation.

Shubsda et al. (*39*) in a previous report used nondenaturing PAGE to observe SL3 binding to NCp7. The analysis found that the best fit to the band patterns was with $(R_2P)_x$ stoichiometries. Perhaps the most important explanation for the difference is that the gel study was conducted under low-salt conditions; also, it was not done with fresh protein. Another recent study reported native gel experiments with SL4−NCp7 titrations; extra bands were observed that suggest multiple species in the gel (*40*). There are other issues with native gels that are difficult to control: (1) free protein, with its high positive charge, migrates rapidly toward the top of the loading well (away from the gel), depleting NCp7 from the complex; (2) dilution in the well alters the rapidly equilibrating distribution of species before they enter the gel—dilution is difficult to manage in a reproducible manner; (3) different ionic conditions exist in the sample tube, the well, and the gel; and (4) the rate constants for interchange among the species are dramatically different in solution and in the gel (*51*, *52*).

In a recent paper, Maki et al. (*44*) report tight binding of SL2, SL3, and SL4 to NCp7 with an RP_2 stoichiometry, and to a 20mer SL3 construct with an abasic loop at R_1P_1. Stoichiometry and binding constants were determined by monitoring W37 fluorescence titrations in low salt (10 mM phosphate, pH 7) and 1 μM NCp7. Aside from the difference in buffer conditions, our study was conducted in a very similar manner. The authors apparently derive the stoichiometry from an extrapolation similar to that described in the discussion of Figure 3, above. An RP_2 stoichiometry would produce an intercept at $\boldsymbol{R}_t/\boldsymbol{P}_t = 0.5$, which is clearly untenable for our data. Unfortunately, the original fluorescence data were not presented (*44*).

Other experiments used nitrocellulose filter-binding assays, all assuming 1:1 stoichiometry to analyze the data. Clever et al. (*15*) measured affinities at ∼0.1 M ionic strength for several packaging domain fragments to p15 fused at its C-terminus to bacterial glutathione *S*-transferase (p15 is the immediate precursor from which NCp7 is cleaved in maturing virus particles). Two categories of binding were established, with an SL2 fragment binding at K_d ∼400 nM, and SL3, SL4, and a full-length SL1 exhibiting K_d ∼200 nM. All of the RNA fragments were considerably longer than those in our current report; therefore, electrostatic effects on binding will be different. In a recent study by Berglund et al. (*53*), a K_d of 1450 nM for a 27mer SL3 derivative is reported. This is roughly 40-fold weaker binding than we observe. The experiments used 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM $MgCl_2$, 30 μM $ZnCl_2$, and 10 mM 1,4-dithiothreitol. Again, a source of the lower affinity observed in these studies may be that the ionic strength was less than the low-salt threshold discussed above.

In an elegant study of exchange cross-peaks in 2D NOE spectra, Amarasinghe et al. (*35*) demonstrated that NCp7 complexes with SL3 and SL2 indeed have 1:1 stoichiometry, and that equilibration is rapid among the bound and free species (nearly equimolar) at NMR concentrations (*35*). They also reported measurements by isothermal titration calorimetry (ITC), giving K_d values of 170 nM for SL3 and 110 nM for SL2 binding to NCp7. ITC was used to measure the enthalpy of binding; then a 1:1 model was used to fit the data and produce K_d values. A more recent ITC study showed that a 14mer SL4 construct exhibited such a small enthalpy of binding to NCp7 that K_d could not be determined reliably (*40*). These studies were done at low ionic strength (25 mM

acetate, pH 6.5, 25 mM NaCl, 0.1 mM $ZnCl_2$, and 0.1 mM β-mercaptoethanol). We note that the entropy contribution often dominates the stabilization of complexes between charged species, and the enthalpy contribution may be small (*54*−*57*). This is a consequence of energetically favorable ion−water and ion−counterion interactions in the free state being replaced by almost equally favorable ion−ion interactions in the bound form. With highly charged species, such as those examined here, one expects a large entropy component, with the release of water and counterions to the bulk. This often occurs in the interaction of charged compounds in the grooves of DNA (*58*−*60*). A small enthalpy for complex formation in low ionic strength buffer may be quite reasonable for the SL4−NCp7 complex.

It is noteworthy that three-dimensional structures have been determined for SL3−NCp7 (*36*) and SL2−NCp7 (*35*) complexes in this solution (∼50 mM in small ions). Structures could not have been determined if the RNA loop−NC complexes were unstable relative to electrostatic binding. How can we reconcile the NMR and ITC observations with our results? Perhaps nonspecific electrostatic binding is a more serious problem for low protein, RNA, and salt concentrations. As mentioned earlier, fluorescence titrations at [NaCl] ≤0.15 M were not reproducible. The ITC experiments used 30−500 times the $\boldsymbol{P}_t$ and $\boldsymbol{R}_t$ concentration in our experiments, and NMR concentrations are several thousand times larger than for fluorescence. Such a large concentration increase may have two effects. Simple mass action favors complex formation, of course, but there is no reason for that to distinguish between a 1:1 complex with specific loop−protein interaction and one involving nonspecific electrostatic binding. A second effect may involve the alterations in ionic strength that accompany the introduction of highly charged ions at millimolar concentrations. For example, an SL3 complex with R_1P_1 stoichiometry would have a −10 formal charge. The complex and its counterions would contribute nearly 0.1 M to the ionic strength in a 2 mM NMR sample.

Conclusion. RNA stem-loops from the major packaging specificity region of HIV-1 form complexes with NCp7 that have $K_d = 20-300$ nM at ionic strengths near physiological conditions. Knowing the strength of interactions with wild-type loops is an important step in designing drugs to interfere with nucleocapsid functions in the viral life cycle. Variations occur among the loop sequences, with SL3 and SL2 having the highest affinity. In addition to the G-X-G sequences in the four apical loops, it was found that the internal loop in SL1 also binds NCp7, and that several of the proteins can bind simultaneously to a sequence encompassing the major packaging domain. The sequence dependence of binding and lack of affinity for UNCG and GAUA tetraloops indicate that a large part of the binding free energy must be due to interaction of the protein with the loop. Raising the ionic strength to 0.2 M reduces nonspecific binding, leads to full quenching of Trp-37 fluorescence, a 1:1 stoichiometry, and is consistent with recent NMR-based structures of SL2 and SL3 complexes. The binding assay at the ionic strength described here provides a simple and reliable method to establish structure/free energy relationships. It should also be useful for surveying the competitive potential of small molecules aimed at blocking the RNA−nucleocapsid interaction.

ACKNOWLEDGMENT

Michael Cavaluzzi purified some NCp7 samples and, together with Deborah Kerwood, performed NMR analysis of the protein's integrity. We gratefully acknowledge Roberto de Guzman for advice in purifying the NCp7 protein. We also thank Profs. James Dabrowiak and Watson Lees (Syracuse University) for helpful discussions.

SUPPORTING INFORMATION AVAILABLE

A figure illustrating titration curves for 27mer SL1a, 23mer SL1i, and 19mer SL2 constructs is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Coffin, J. M., Hughes, S. H., and Varmus, H. E. (1997) *Retroviruses*, Cold Spring Harbor Laboratory Press, Plainview, NY.
2. Omichinski, J. G., Clore, G. M., Sakaguchi, K., Appella, E., and Gronenborn, A. M. (1991) *FEBS Lett. 292*, 25−30.
3. Morellet, N., Jullian, N., De Rocquigny, H., Maigret, B., Darlix, J. L., and Roques, B. P. (1992) *EMBO J. 11*, 3059−3065.
4. Summers, M. F., Henderson, L. E., Chance, M. R., Bess, J. W., South, T. L., Blake, P. R., Sagi, I., Perez-Alvarado, G., Sowder, R. C., Hare, D. R., et al. (1992) *Protein Sci. 1*, 563−574.
5. Berkowitz, R., Fisher, J., and Goff, S. P. (1996) *Curr. Top. Microbiol. Immunol. 214*, 177−218.
6. Darlix, J. L., Gabus, C., Nugeyre, M. T., Clavel, F., and Barre-Sinoussi, F. (1990) *J. Mol. Biol. 216*, 689−699.
7. Muriaux, D., De Rocquigny, H., Roques, B. P., and Paoletti, J. (1996) *J. Biol. Chem. 271*, 33686−33692.
8. De Rocquigny, H., Gabus, C., Vincent, A., Fournie-Zaluski, M. C., Roques, B., and Darlix, J. L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6472−6476.
9. Sakaguchi, K., Zambrano, N., Baldwin, E. T., Shapiro, B. A., Erickson, J. W., Omichinski, J. G., Clore, G. M., Gronenborn, A. M., and Appella, E. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5219−5223.
10. Laughrea, M., Shen, N., Jette, L., Darlix, J. L., Kleiman, L., and Wainberg, M. A. (2001) *Virology 281*, 109−116.
11. Tanchou, V., Gabus, C., Rogemond, V., and Darlix, J. L. (1995) *J. Mol. Biol. 252*, 563−571.
12. Barat, C., Schatz, O., Le Grice, S., and Darlix, J. L. (1993) *J. Mol. Biol. 231*, 185−190.
13. Berkhout, B. (1996) *Prog. Nucleic Acid Res. Mol. Biol. 54*, 1−34.
14. Berkhout, B., and van Wamel, J. L. (2000) *RNA 6*, 282−295.
15. Clever, J., Sassetti, C., and Parslow, T. G. (1995) *J. Virol. 69*, 2101−2109.
16. Pappalardo, L., Kerwood, D. J., Pelczer, I., and Borer, P. N. (1998) *J. Mol. Biol. 282*, 801−818.
17. Zeffman, A., Hassard, S., Varani, G., and Lever, A. (2000) *J. Mol. Biol. 297*, 877−893.
18. Hoglund, S., Ohagen, A., Goncalves, J., Panganiban, A. T., and Gabuzda, D. (1997) *Virology 233*, 271−279.
19. Vogt, V. M., and Simon, M. N. (1999) *J. Virol. 73*, 7050−7055.
20. Ennifar, E., Yusupov, M., Walter, P., Marquet, R., Ehresmann, B., Ehresmann, C., and Dumas, P. (1999) *Struct. Fold. Des. 7*, 1439−1449.
21. Mujeeb, A., Clever, J. L., Billeci, T. M., James, T. L., and Parslow, T. G. (1998) *Nat. Struct. Biol. 5*, 432−436.
22. Mujeeb, A., Parslow, T. G., Zarrinpar, A., Das, C., and James, T. L. (1999) *FEBS Lett. 458*, 387−392.
23. Takahashi, K., Baba, S., Hayashi, Y., Koyanagi, Y., Yamamoto, N., Takaku, H., and Kawai, G. (2000) *J. Biochem. (Tokyo) 127*, 681−686.

24. Girard, F., Barbault, F., Gouyette, C., Huynh-Dinh, T., Paoletti, J., and Lancelot, G. (1999) *J. Biomol. Struct. Dyn. 16*, 1145−1157.
25. Amarasinghe, G. K., De Guzman, R. N., Turner, R. B., and Summers, M. F. (2000) *J. Mol. Biol. 299*, 145−156.
26. Kerwood, D. J., Cavaluzzi, M. J., and Borer, P. N. (2001) *Biochemistry 40*, 14518−14529.
27. Clever, J. L., Eckstein, D. A., and Parslow, T. G. (1999) *J. Virol. 73*, 101−109.
28. Clever, J. L., and Parslow, T. G. (1997) *J. Virol. 71*, 3407−3414.
29. McBride, M. S., and Panganiban, A. T. (1996) *J. Virol. 70*, 2963−2973.
30. McBride, M. S., and Panganiban, A. T. (1997) *J. Virol. 71*, 2050−2058.
31. McBride, M. S., Schwartz, M. D., and Panganiban, A. T. (1997) *J. Virol. 71*, 4544−4554.
32. Hayashi, T., Shioda, T., Iwakura, Y., and Shibuta, H. (1992) *Virology 188*, 590−599.
33. Hayashi, T., and Iwakura, Y. (1993) *Tanpakushitsu Kakusan Koso 38*, 779−783.
34. Damgaard, C. K., Dyhr-Mikkelsen, H., and Kjems, J. (1998) *Nucleic Acids Res. 26*, 3667−3676.
35. Amarasinghe, G. K., De Guzman, R. N., Turner, R. B., Chancellor, K. J., Wu, Z. R., and Summers, M. F. (2000) *J. Mol. Biol. 301*, 491−511.
36. De Guzman, R. N., Wu, Z. R., Stalling, C. C., Pappalardo, L., Borer, P. N., and Summers, M. F. (1998) *Science 279*, 384−388.
37. South, T. L., and Summers, M. F. (1993) *Protein Sci. 2*, 3−19.
38. Morellet, N., Demene, H., Teilleux, V., Huynh-Dinh, T., de Rocquigny, H., Fournie-Zaluski, M. C., and Roques, B. P. (1998) *J. Mol. Biol. 283*, 419−434.
39. Shubsda, M. F., Kirk, C. A., Goodisman, J., and Dabrowiak, J. C. (2000) *Biophys. Chem. 87*, 149−165.
40. Amarasinghe, G. K., Zhou, J., Miskimon, M., Chancellor, K. J., McDonald, J. A., Matthews, A. G., Miller, R. R., Rouse, M. D., and Summers, M. F. (2001) *J. Mol. Biol. 314*, 961−970.
41. Fisher, R. J., Rein, A., Fivash, M., Urbaneja, M. A., Casas-Finet, J. R., Medaglia, M., and Henderson, L. E. (1998) *J. Virol. 72*, 1902−1909.
42. Druillennec, S., Meudal, H., Roques, B. P., and Fournie-Zaluski, M. C. (1999) *Bioorg. Med. Chem. Lett. 9*, 627−632.
43. Vuilleumier, C., Bombarda, E., Morellet, N., Gerard, D., Roques, B. P., and Mely, Y. (1999) *Biochemistry 38*, 16816−16825.
44. Maki, A. H., Ozarowski, A., Misra, A., Urbaneja, M. A., and Casas-Finet, J. R. (2001) *Biochemistry 40*, 1403−1412.
45. Lee, B. M., De Guzman, R. N., Turner, B. G., Tjandra, N., and Summers, M. F. (1998) *J. Mol. Biol. 279*, 633−649.
46. Tummino, P. J., Scholten, J. D., Harvey, P. J., Holler, T. P., Maloney, L., Gogliotti, R., Domagala, J., and Hupe, D. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 969−973.
47. Dharmacon. http://www.dharmacon.com/chem.html.
48. Molinaro, M., and Tinoco, I. (1995) *Nucleic Acids Res. 23*, 3056−3063.
49. Allain, F. H., and Varani, G. (1995) *J. Mol. Biol. 250*, 333−353.
50. Record, M. T., Jr., Lohman, M. L., and De Haseth, P. (1976) *J. Mol. Biol. 107*, 145−158.
51. Cann, J. R. (1996) *Anal. Biochem. 237*, 1−16.
52. Shubsda, M., Goodisman, J., and Dabrowiak, J. C. (1999) *Biophys. Chem. 76*, 95−115.
53. Berglund, J. A., Charpentier, B., and Rosbash, M. (1997) *Nucleic Acids Res. 25*, 1042−1049.
54. Schulz, G. E., and Schirmer, R. H. (1979) *Principles of Protein Structure*, Springer-Verlag, New York.
55. Gurney, R. W. (1962) *Ionic Processes in Solution*, Dover, New York.
56. Ramanathan, P. S., and Friedman, H. L. (1971) *J. Chem. Phys. 54*, 1086−1099.
57. Friedman, H. L., and Krishnan, C. V. (1973) in *Water: A Comprehensive Treatise* (Franks, F., Ed.) pp 1−118, Plenum, New York.
58. Haq, I., Ladbury, J. E., Chowdhry, B. Z., Jenkins, T. C., and Chaires, J. B. (1997) *J. Mol. Biol. 271*, 244−257.
59. Chaires, J. B. (1998) *Curr. Opin. Struct. Biol. 8*, 314−320.
60. Hamilton, T. B., Borel, F., and Romaniuk, P. J. (1998) *Biochemistry 37*, 2051−2058.

BI016045+