Published in final edited form as:

J Mol Biol. 2011 July 29; 410(5): 984–996. doi:10.1016/j.jmb.2011.03.039.

A small molecule probe induces a conformation in HIV TAR RNA capable of binding drug-like fragments

Amy Davidson^{*}, Darren W. Begley[†], Carmen Lau^{*}, and Gabriele Varani^{*,§,¶}

^{*} Department of Chemistry - University of Washington, Box 351700, Seattle, WA USA 98195

[†] Emerald BioStructures, 7869 NE Day Road West, Bainbridge Island, WA USA 98110

[§] Department of Biochemistry - University of Washington, Box 357350, Seattle, USA 98195

Abstract

The HIV-1 TAR/Tat interaction is a potentially valuable target for treating HIV infection, but efforts to develop TAR-binding antiviral drugs have not yet yielded a successful candidate for clinical development. In this work, we describe a novel approach towards screening fragments against RNA that uses a chemical probe to target the Tat binding region of TAR. This probe fulfills two critical roles in the screen: by locking the RNA into a conformation capable of binding other fragments, it simultaneously allows the identification of proximal binding fragments by ligand-based NMR. Using this approach, we have discovered six novel TAR-binding fragments, three of which were docked relative to the probe/RNA structure using experimental NMR restraints. The consistent orientations of functional groups in our data-driven docked structures and common electrostatic properties across all fragment leads reveal a surprising level of selectivity by our fragment-sized screening hits. These models further suggest linking strategies for the development of higher affinity lead compounds for the inhibition of the TAR/Tat interaction.

Introduction

The development of highly active antiretroviral therapies over the last two decades has led to remarkable progress in HIV treatment, but drug-resistant strains remain an obstacle to controlling viral infections. Novel therapeutics are therefore needed to target vulnerable steps in the viral replication cycle which are less prone to mutagenesis. One such target is the activation of RNA transcription through binding of the viral protein Tat to the trans-activation response (TAR) element. The HIV-1 TAR element is a stem-loop located at the 5' end of nascent viral transcripts which has been recognized as a potentially valuable drug target for some time^{1–3}. The cooperative binding of Tat and host cell co-factor cyclin T₁ to TAR recruits the kinase CDK9 to hyperphosphorylate RNA polymerase II, leading to enhanced transcriptional processivity. Without trans-activation, the polymerase produces primarily short transcripts and viral replication is disrupted^{4,5}. Extensive structural studies of TAR RNA have paved the way for the design of molecules to compete with the Tat

© 2011 Elsevier Ltd. All rights reserved.

[¶]Address correspondence to: GV varani@chem.washington.edu, 1 206 543 7113 (Tel), 1-206 685 8665 (Fax).

Accession Numbers: Coordinates for the TAR-MV2003 structure have been deposited in the Protein Data Bank with accession number 2L8H.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

protein in binding to the TAR element, thus inhibiting viral replication^{3, 6–19}. However, to date none of these efforts have yielded leads with the potency and specificity required for successful clinical candidacy. Thus, new approaches may be necessary for the development of effective small molecule inhibitors of this critical interaction.

Fragment-based lead discovery could be a valuable avenue to discovering new leads, since this technique is quickly becoming an established strategy for the discovery and subsequent development of novel drug leads for protein targets^{20–23}. However, to date few studies have reported success in utilizing RNA as target of fragment screens, with one important exception^{24–26}. Compared with proteins, the flexibility, ionic character, and reduced chemical diversity all make targeting RNA with weakly binding fragments particularly challenging. Previous studies have shown that RNA can be screened by nuclear magnetic resonance (NMR) spectroscopy with fragment libraries^{9, 27–29}. However, the majority of hits were charged molecules unlikely to possess binding site specificity. One possible solution lies in designing more focused libraries. Since most screening collections have been developed for protein targets, it has been suggested that design of a RNA-specific fragment library could improve hit rates^{28, 30}. However, it remains to be demonstrated whether such a library will generate hits with sufficient specificity for development. Unfortunately, aminoglycosides, intercalating agents, and other known RNA-binding entities from which such libraries could be developed possess low intrinsic target specificity, a difficult property to reverse.

In pursuit of novel lead molecules that bind to HIV-1 TAR, we have undertaken fragment-based screening by nuclear magnetic resonance (NMR). However, to overcome the limitations described above, we used a target-specific chemical probe to enhance the likelihood of finding new lead molecules. In this approach, a first small molecule ligand known to bind the target is used to identify novel, weakly-binding fragments by detection of inter-ligand nuclear Overhauser effect (ILOE)³¹. We selected a chemical probe consisting of an arginine derivative which mimics the crucial arginine residue in Tat recognition of TAR. We hypothesized that such a probe would bind the RNA at a specific site, and, once bound, its presence would allow detection of interligand NOE signals to identify proximal fragments. We were surprised to observe that the small molecule probe fulfills an additional role by locking TAR into a conformation conducive to fragment recognition³². By detecting NOEs to the probe, we were able to identify a set of novel, non-charged, drug-like small molecule fragments which bind to HIV-1 TAR RNA only when the probe is present. The structure of the TAR-probe complex was then determined using standard solution-state NMR methods, and used to generate ILOE-driven binding mode predictions for the most promising fragment hits. We have thus identified a set of fragment molecules which serve as starting points for lead development through rational, structure-based elaboration.

Results

Probe selection

To obtain a suitable probe molecule for screening HIV-1 TAR RNA with a fragment library, a small panel of arginine derivatives was selected from commercial sources. Sixteen compounds were examined for adequate solubility and for binding to TAR by one-dimensional saturation transfer difference (STD)-NMR and two-dimensional proton-proton nuclear Overhauser effect (NOE)-NMR spectroscopy^{33, 34} (Table 1). The strength of STD signal attenuation and the presence of negative NOEs (positive proton-proton NOE cross-peaks) for these compounds in solution with sub-stoichiometric concentrations of TAR RNA were used as criteria for probe selection. Compounds were eliminated due to poor solubility, weak STD signals (associated with weak binding), or proton resonances occurring at the saturation frequency used to irradiate the target for STD-NMR (see Methods). Based on

strong STD signal attenuation and positive NOE cross-peaks observed in the presence of TAR (Figure 1), the arginine derivative ultimately selected for fragment screening was MV2003 (Table 1). This molecule combines an arginine group to bind the recognition pocket of TAR with a proton-rich methoxynaphthalene group ideal for detection of ILOEs with screening compounds.

Fragment screening

The fragment library used for screening consisted of 250 molecules from a Maybridge “Rule of 3” collection sorted into mixtures of 5 to 8 ligands dissolved in deuterated dimethyl sulfoxide (see Methods). Despite previous reports of poor success using STD-NMR for screening RNA targets^{27, 28}, we obtained clear binding signatures from this small diversity set when the screen was executed in the presence of MV2003. Using a cutoff of 10% STD peak difference as our initial criteria, over 100 of these fragments were observed to bind to the RNA, with clear non-binding fragments generating < 3% STD peak intensity. Fragments which generated ≥ 10% STD and observable ILOEs to MV2003 were classified as putative hits, then sorted into different groups based on scaffold similarity, each rank-ordered by STD signal attenuation. A total of 20 fragments with the most intense STD signals were selected from 13 different scaffold groups for individual follow-up experiments to confirm and characterize their modes of binding (Table 2).

The 20 candidate fragments were analyzed by ligand-observe NOE spectroscopy alone, with TAR, with MV2003, and with both MV2003 and TAR in solution. Of these 20 candidate fragments, 6 were validated as hits. These generated negative NOE cross-peaks and little to no STD-NMR binding signals when studied alone, pairwise with MV2003, or pairwise with TAR RNA; these results are indicative of non-binding. However, when MV2003 and TAR RNA were both added to the solution, each of these 6 individual fragments generated positive NOE cross-peaks and gave clear, unambiguous STD peak differences. As shown for fragment hit MV1480, these binding signatures are undetectable when individual fragments are studied with MV2003 in the absence of TAR, or when analyzed with TAR in the absence of MV2003 (Figure 2). The NOE and STD data thus identify these 6 fragments to effectively participate in binding to TAR RNA only when MV2003 is also present in solution. When mixed with TAR in the absence of MV2003, these 6 fragments generate null or negative intra-molecular NOE signals, indicative of fast-tumbling, non-binding small molecules. Due to these binding characteristics, these 6 fragments were carried forward for further investigation.

Structure of the binary complex

Structure determination of the binary MV2003-TAR complex was made possible by the observation of several inter-molecular NOEs in ¹H-¹H NOESY spectra collected with equimolar samples. The guanidinium group of the probe molecule binds near the U23-C24-U25 bulge region, likely stabilizing partial base triple formation between bulge residue U23 and base pair A27-U38 (Figure 3a) through a cation- π interaction, as seen in previous structures of HIV TAR RNA with peptidic ligands^{15–17, 35}. Precise positioning of the guanidinium group in the structure was not possible due to lack of NOE data for its exchangeable protons. However, protection of the U23 imino hydrogen from exchange was observed in ¹H-¹H NOESY data from a sample prepared in 90% H₂O/10% D₂O solution, indicating partial formation of the base triple. Initial structure calculations were performed without any hydrogen bonding restraints between U23 and A27. The results verified that other RNA intramolecular NOEs were consistent with partial triple formation, and a single hydrogen bond restraint was added in subsequent calculations for the U23 imino proton to the N7 atom of A27.

The methoxynaphthalene group of MV2003 binds in a pseudo-stacking arrangement at the top of the upper RNA helix and below the base of the apical loop, between the C29-G36 base pair and loop residue A35 (Figure 3a). Compared to aromatic groups of other known TAR-binding arginine derivatives, this ligand binds with a unique orientation (Figure 3b). Most other TAR-binding small molecules tend to be positioned further down the helix, participating in aromatic base-stacking interactions with nucleobases of the major groove^{8, 10}. By contrast, MV2003 contacts the apical loop, where cyclin T₁ binds, while also stabilizing partial formation of the U23/A27/U38 base triple^{8, 15}. The dual interactions of the methoxynaphthalene ring with the base of the apical loop and the guanidinium group with the arginine-binding pocket of TAR cause an overall tightening of the RNA major groove around MV2003. This creates a small pocket capable of binding small molecule fragments. This cavity was not observed with previous structures of TAR bound to argininamide or any other arginine derivatives (Figure 3b). In this manner, MV2003 locks TAR into a conformation capable of binding small molecule fragments, while simultaneously providing a chemical probe for fragment-based screening by ligand-observe NMR methods. The bound state of MV2003 provides a promising foundation for the development of inhibitors capable of abrogating both Tat and Cyclin T₁ binding to the viral RNA.

Models of the ternary complexes

NMR data were then collected to assess the feasibility of structure determination for the top fragment hits in complex with the probe-TAR complex. Unfortunately, NOESY data did not yield any intermolecular NOEs between fragment molecules and the RNA at stoichiometric concentrations. This is likely due to relatively weak fragment binding interactions, which result in poor restraint of the fragments to any single bound conformation. However, measurable differences in ILOE cross-peak intensities between protons within the fragments and MV2003 bound to TAR suggested their use as experimental restraints for docking fragments into the binary probe-RNA structure. We used the program HADDOCK (High Ambiguity Driven DOCKing) for modeling, due to its success in analyzing ligand-protein complexes^{36, 37}. Of the 6 fragment hits validated by control experiments, 3 were selected for docking experiments, based on uniqueness of structure and adequate molecular size and complexity for docking. We used ILOE cross-peak intensities collected on individual fragments in solution with MV2003 and TAR to generate ambiguous interaction restraints between the probe and fragment hits, as well as unambiguous restraints for specific nuclei (see Methods).

Structural differences between the MV2003 binary complex and other ligand-bound structures of TAR, plus strong ILOE cross-peaks between fragment hits and several aromatic protons of MV2003, suggest that the fragment-binding pocket is located in the RNA major groove. Consistent with this hypothesis, modeling studies placed all three fragments in the same region of the complex, approximately between nucleotides A27 and C29 (Figure 4). This result positions the fragments on the major groove face of the RNA, near the methoxy-naphthalene ring system of MV2003. Similar orientations in the binding pocket were obtained for all three structurally divergent fragment hits, suggesting a common binding epitope, and demonstrating binding site specificity for these fragments with the MV2003-TAR complex. The HADDOCK models position all fragments with hydrogen-bond acceptors toward the hydrogen-donating guanidinium group of MV2003, an orientation which is largely determined by specific burial of the fragment scaffold against the MV2003 naphthalene group (Figure 4). Thus, the fragment ring systems align with the naphthalene group of MV2003, while hydrophilic functional groups point toward the guanidinium group and the RNA phosphate backbone in the TAR bulge region. These preliminary structures immediately suggest linking strategies to create a single hydrophobic

core to be buried in the TAR major groove, while presenting two polar functional groups for interaction with RNA bulge residues and the U23-A27-U38 base triple.

Discussion

It has been known for some time that an arginine side chain within Tat protein is particularly important for binding TAR, and argininamide itself can capture at least some of the characteristics of the Tat-TAR interaction^{3, 38–40}. In this study, we observe that binding of an arginine derivative to TAR induces a pocket in the major groove of the upper helix capable of burying other small molecules. In this manner, our 'probe' plays a similar role to the crucial arginine of Tat protein for TAR recognition, by reorganizing the RNA structure; the formation of a new binding pocket allows binding of other fragments and suggests that more powerful ligands can be generated by linking the fragments together. Thus, NMR-based fragment screening by detection of ILOEs to our chemical probe has allowed the discovery of drug-like molecules that bind a new pocket within TAR. Such fragments can now be used to design and elaborate lead compounds which contain novel drug-like motifs not previously applied to RNA-binding molecules.

The arginine derivative selected as our chemical probe contains two functional regions: a polar arginine side chain for interaction with the TAR UCU bulge, and a non-polar, hydrogen-rich region for detection of ¹H-¹H ILOE to binding fragments. It proved ideal for fragment-based screening since it locked the RNA into a conformation capable of generating NMR-detectable interactions that the free RNA does not possess. The combination of the methoxy-naphthalene group and arginine side-chain mimics other known TAR ligands and contacts two of the crucial TAR interaction sites identified in previous studies of Tat mimetics^{3, 13, 15, 16, 41}. By using this TAR binding motif to direct our fragment search, we can now efficiently design anti-TAR lead compounds based on the models of the ternary complex (Figure 4).

Despite the use of the probe, follow-up studies were essential to identify and validate TAR-binding fragments from primary screening. Of the 20 fragments selected for further analysis, 5 failed to generate binding signals when separated from the screening mixture, even with MV2003 and TAR RNA present in solution (Table 2). These compounds likely generated false positive signals due to solubility limits or aggregation in the presence of other fragments at relatively high (i.e. mM) concentration during screening. An additional 6 hits generated positive NOE cross-peaks in the absence of RNA, either alone or in the presence of MV2003, and were therefore excluded from further study. Among the remaining candidates, 3 generated negative NOE cross-peaks alone or with MV2003 present, but positive NOE cross-peaks when alone with TAR in solution. These fragments (MV1062, MV1159 and MV1425) appear to bind to TAR with or without MV2003 present, suggestive of non-specific binding to RNA (Table 2). The remaining 6 fragment hits only generated binding signals when both MV2003 and TAR RNA were present, indicative of affinity for a pocket within TAR insufficiently populated without the chemical probe. Further demonstration of binding site specificity is apparent when these fragments are compared to non-hits sharing the same scaffold (Figure 5). This comparison shows that small changes to functional groups and their placement can disrupt recognition of the fragments by the MV2003-TAR complex. This level of specificity for small molecule fragments was surprising, but very encouraging for lead development.

Conclusion

We have undertaken a fragment-based screen to discover new hits for HIV-1 TAR RNA, but we have done so by first targeting the Tat-binding site with an arginine-derived probe. The

arginine derivative we have chosen induces a new conformation of TAR RNA, and its use has enabled us to discover non-charged, drug-like fragments with weak but specific affinity for this RNA. The chemical probe induces the formation of a small cavity in the major groove of the RNA for which the fragment molecules display a high degree of binding-site specificity. Remarkably, small changes in fragment structure, specifically to polar functional groups, can result in total loss of binding, further supporting the binding specificity, most likely due to stable formation of the binding pocket. Based on HADDOCK-generated models, all true positive fragment hits share the characteristic of hydrogen bond acceptors oriented toward the guanidinium group of MV2003, and hydrogen bond donors or neutral atoms oriented toward the RNA backbone. This functional group orientation is apparently determined by the burial of the fragment scaffold against the probe naphthalene group in the RNA major groove. The probe and fragment hits which we have identified in this work will now guide the discovery of new HIV-1 Tat-TAR inhibitors.

The approach presented here represents one of few published attempts to experimentally screen and successfully discover novel RNA binding motifs by ligand-based NMR^{30, 42, 43}. It is also an uncommon example of utilizing an all-purpose, drug-like library of molecules, biased toward protein targets, to successfully screen and identify novel chemical entities which bind to RNA. With proper selection of chemical probe, the screening methods used in this study could be expanded to other RNA drug targets with potential for medicinal intervention through small molecule inhibition.

Experimental Methods

RNA preparation—HIV-1 TAR RNA was prepared *in vitro* from commercially synthesized DNA templates (IDT)) using *in house* purified T7 RNA polymerase and commercially available nucleotides (Sigma) and purified as described⁴⁴. The construct used in all experiments was of the sequence 5' -

G₁₇GCAGAUUCUGAGCCUGGGAGCUCUCUGCC₄₅ -3'. Prior to each NMR experiment, the RNA samples were annealed by heating at 95 °C briefly, then quick cooling on ice. For screening, the RNA was dissolved in 90% H₂O/10% ²H₂O (v/v). For NMR experiments, the RNA was lyophilized from 99.9% ²H₂O twice, and then dissolved in 99.99% ²H₂O.

NMR experiments for fragment screening and binding characterization—NMR samples for fragment screening were prepared by diluting purified, snap-cooled HIV-1 TAR RNA into NMR screening buffer (10 mM potassium phosphate, 0.1 mM EDTA, 10% (v/v) ²H₂O, adjusted to pH = 6.6 and sterile-filtered). Samples contained pools of fragments at 500 μM each in the presence of 20 μM RNA with 500 μM of the chemical probe MV2003. Follow-up experiments were conducted with focused pools and individual fragments at 200 μM with MV2003 also at 200 μM and the RNA at 5 or 10 μM. All screening and follow-up experiments were conducted on a 750-MHz Bruker AV spectrometer with a TXI probe at 280 K. Screening was done using ligand-observe, proton-based one-dimensional saturation transfer difference nuclear magnetic resonance (STD-NMR)³³ and two-dimensional nuclear Overhauser effect spectroscopy (NOESY)⁴⁵, according to previously published methods⁴⁶.

For STD-NMR experiments, 32 scans and 32,000 points were acquired over a 14 ppm sweep width, with a total recycle delay of 4.0 seconds for each mixture. Reference and saturation FIDs were acquired in an interleaved fashion over the course of a single experiment for each NMR sample, and processed separately with identical phases and processing parameters. A low-power 30-ms spin-lock pulse was added to filter out low-level RNA peaks, and a WATERGATE sequence used to suppress bulk water signal⁴⁷. STD-NMR pre-saturation was done using a 3.0 second train of Gaussian-shaped pulses with a spectral width of approximately 750 Hz. The selective saturation pulse was focused at 5.5

ppm to excite ribose proton resonances of the target, and reference irradiation was focused at 30 ppm. The majority of fragment molecules, as well as MV2003, were free of proton resonances near this saturation frequency. Fragments with proton peaks close to 5.5 ppm were excluded from STD-NMR-based experiments, due to the likelihood of direct irradiation by the selective excitation pulse. STD peak intensities were measured by subtracting the integral area of a saturation spectrum peak from that of the corresponding reference spectrum. Previous studies suggest a cut-off value for fragment-based NMR screening hits to be a saturation transfer difference amplification factor (SAF) of 2.5 for one or more ligand protons⁴⁶. This SAF value corresponds to approximately 10% STD peak intensity for all fragments studied at the same ligand:target ratio, and was used as a preliminary hit criterion for subsequent screening.

For NOESY experiments, 2048×160 points were collected with a mixing time of 500 ms and a recycle delay of 2.0 s, with WATERGATE⁴⁷ solvent suppression for each mixture or sample. To derive restraints for HADDOCK, the same NOESY parameters were used to collect data for samples containing 200 μ M fragment, 200 μ M MV2003 and 5 μ M TAR RNA in the same buffer, at 50, 100, 150 and 200 ms mixing times.

NMR experiments for structure determination—NMR spectra for structure determination were collected on Bruker DRX 500 MHz and DMX 600 MHz spectrometers equipped with cryo-cooled HCN probes with triple-axis gradients. Data were processed with nmrPipe⁴⁸ and analyzed in Sparky⁴⁹. NOESY experiments in D₂O were recorded at two mixing times (100 ms and 200 ms). NOESY experiments recorded in water at 4 °C were collected with WATERGATE solvent suppression⁴⁷ and mixing times of 100 and 200 ms. 3D ¹³C NOESY-HMQC spectra were recorded at 600 MHz with a mixing time of 100 ms.

To determine optimal concentrations in solution for structural characterization of the complex, TAR imino chemical shifts were monitored by 1D proton NMR while titrating in MV2003. Significant changes in chemical shift for several imino peaks were observed during the titration up to a 4:1 stoichiometric ratio, after which changes were minimal. This result indicated that the primary binding site for MV2003⁵⁰ was fully titrated at a 4:1 ratio of MV2003 to TAR, and NMR data for structural studies were collected under these conditions.

Resonance assignments were already available for TAR bound to a single argininamide and several Tat-derived peptides^{3, 51}. Comparison of ¹H-¹H NOESY spectra for these previously determined complexes and the MV2003-bound RNA revealed several similar chemical shift values near the 5' - and 3' - ends of the RNA, i.e. for residues which are not involved in binding the peptides or MV2003. These observations, along with clearly identifiable adenine H2 proton resonances, were used to generate initial proton assignments⁵². HCCH-COSY and HMQC-NOESY experiments were then collected to complete proton assignments.

To measure residual dipolar couplings (RDC), in-phase-anti-phase (IPAP) HSQC data were collected for unaligned samples of the MV2003/RNA complex as reference⁵³. The sample was then partially aligned using Pf1 phage (Asla Ltd.) in a 99.9% D₂O buffer (10 mM potassium phosphate, pH 6.6) to a final concentration of 15 mg/mL of phage. The IPAP experiment was repeated for this partially aligned sample. Heteronuclear one-bond ¹H/¹³C couplings were measured from the IPAP data in the presence of phage ($J_{w/pf1}$) and in the absence of phage ($J_{w/outpf1}$). RDC values were determined using the equation $RDC = J_{w/pf1} - J_{w/outpf1}$.

Structure determination of TAR-MV2003 binary complex—The structure of the complex was calculated with Xplor-NIH software using a simulated annealing protocol⁵⁴.

Distance restraints were established from NOE intensities in 2D NOESY spectra at 100 ms mixing time and from 3D NOESY-HSQC spectra at 100 ms mixing time for overlapped resonances. Standard planarity and hydrogen bond restraints were applied to unambiguously base-paired nucleotides, as identified from NOESY spectra in water. Dihedral angle restraints for the bulge and loop nucleotides were experimentally established whenever possible: the strength of the intra-nucleotide H1'–H8 peak was used to restrain the glycosidic angle, while the sugar pucker was determined from the intensity of the H1'–H2' couplings in ^1H - ^1H TOCSY experiments.

Structures were first calculated without RDCs to test for convergence and adherence to the NOE data. RDC restraints were then applied as susceptibility anisotropy (*sani*) restraints with a harmonic potential well. The powder pattern-like distribution of RDCs⁵⁵ was used to calculate starting values for Da and R of approximately –21 and 0.3, respectively. Optimal values of Da and R were then established by a grid search over Da values ranging from –15 to –36 in increments of 2 Hz and R values ranging from 0.15 to 0.45 in 0.05 increments. The lowest energy structures were found for Da= –24 and R= 0.35. These parameters were used in subsequent refinement on calculations of the final 100 structures.

Twenty converged structures of the complex were selected as those with lowest total, NOE, and *sani* energies from the population of structures with no NOE restraint violations greater than 0.5 Å nor dihedral angle violations greater than 5 degrees. These structures were visualized with MolMol⁵⁶ and a final set of 10 structures were selected for convergence to the mean structure.

Fragment docking with MV2003-TAR binary complex—CNS style topology and parameter files were generated with the PRODRG online server, with use of the electrostatic option⁵⁷. Non-polar hydrogens were added from topology and parameters files using Phenix eLBOW software⁵⁸. Docking runs using HADDOCK were performed with primarily default settings, except for the addition of RNA restraints on base pair planarity, sugar pucker and backbone dihedral angles. ILOE signals from 2D NOESY experiments for individual fragments in the presence of probe and RNA were used to generate ambiguous interaction restraints (AIRs) between fragments and probe, with both entities characterized as “active residues”. The top 5–8 strongest ILOE peaks at 100 ms mixing time were translated into unambiguous distance restraints^{37, 59}. 500 initial structures were generated with rigid body docking, and the 100 lowest energy structures were subsequently used for semi-flexible simulated annealing and solvent explicit refinement. Solutions were ranked according to the RMSD from average structure and the total number of ambiguous and unambiguous NOE violations using HADDOCK analysis software. The 10 structures with lowest RMSD and NOE violations for fragments MV1303, MV1315 and MV1480 are shown in Figure 4; Table 3 provides statistics for these structures.

References

1. Jones KA, Peterlin BM. Control of RNA initiation and elongation at the HIV-1 promoter. *Annu Rev Biochem.* 1994; 63:717–43. [PubMed: 7979253]
2. Karn, J.; Gait, MJ.; Churcher, MJ.; Mann, DA.; Mikaelian, IPC. Control of human immunodeficiency virus gene expression by the RNA binding proteins Tat and Rev. In: Nagai, KMIW., editor. *RNA – Protein Interactions*. IRL Press; Oxford: 1994.
3. Aboul-ela F, Karn J, Varani G. The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J Mol Biol.* 1995; 253:313–32. [PubMed: 7563092]
4. Karn J. Tackling Tat. *J Mol Biol.* 1999; 293:235–254. [PubMed: 10550206]

5. Peterlin BM, Price DH. Controlling the elongation phase of transcription with P-TEFb. *Mol Cell*. 2006; 23:297–305. [PubMed: 16885020]
6. Lapidot A, Vijayabaskar V, Litovchick A, Yu J, James TL. Structure-activity relationships of aminoglycoside-arginine conjugates that bind HIV-1 RNAs as determined by fluorescence and NMR spectroscopy. *FEBS Letters*. 2004; 577:415–421. [PubMed: 15556620]
7. Lind KE, Du Z, Fujinaga K, Peterlin BM, James TL. Structure-based computational database screening, in vitro assay, and NMR assessment of compounds that target TAR RNA. *Chemistry & biology*. 2002; 9:185–93. [PubMed: 11880033]
8. Du Z, Lind KE, James TL. Structure of TAR RNA complexed with a Tat-TAR interaction nanomolar inhibitor that was identified by computational screening. *Chemistry & biology*. 2002; 9:707–12. [PubMed: 12079782]
9. Mayer M, James TL. Detecting ligand binding to a small RNA target via saturation transfer difference NMR experiments in D₂O and H₂O. *J Am Chem Soc*. 2002; 124:13376–7. [PubMed: 12418877]
10. Murchie AI, Davis B, Isel C, Afshar M, Drysdale MJ, Bower J, Potter AJ, Starkey ID, Swarbrick TM, Mirza S, Prescott CD, Vaglio P, Aboul-ela F, Karn J. Structure-based drug design targeting an inactive RNA conformation: exploiting the flexibility of HIV-1 TAR RNA. *J Mol Biol*. 2004; 336:625–38. [PubMed: 15095977]
11. Raghunathan D, Sanchez-Pedregal VM, Junker J, Schwiegk C, Kalesse M, Kirschning A, Carlomagno T. TAR-RNA recognition by a novel cyclic aminoglycoside analogue. *Nucl Acids Res*. 2006; 34:3599–3608. [PubMed: 16855296]
12. Leeper TC, Athanassiou Z, Dias RL, Robinson JA, Varani G. TAR RNA recognition by a cyclic peptidomimetic of Tat protein. *Biochemistry*. 2005; 44:12362–72. [PubMed: 16156649]
13. Davis B, Afshar M, Varani G, Murchie AI, Karn J, Lentzen G, Drysdale M, Bower J, Potter AJ, Starkey ID, Swarbrick T, Aboul-ela F. Rational design of inhibitors of HIV-1 TAR RNA through the stabilisation of electrostatic “hot spots”. *J Mol Biol*. 2004; 336:343–56. [PubMed: 14757049]
14. Brodsky AS, Williamson JR. Solution structure of the HIV-2 TAR-argininamide complex. *Journal of Molecular Biology*. 1997; 267:624–639. [PubMed: 9126842]
15. Davidson A, Leeper T, Athanassiou Z, Patora-Komisarska K, Karn J, Robinson J, Varani G. Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. *PNAS*. 2009
16. Davidson A, Patora-Komisarska K, Robinson JA, Varani G. Essential structural requirements for specific recognition of HIV TAR RNA by peptide mimetics of Tat protein. *Nucleic Acids Research*.
17. Ludwig V, Krebs A, Stoll M, Dietrich U, Ferner J, Schwalbe H, Scheffer U, Dürner G, Göbel MW. Tripeptides from Synthetic Amino Acids Block the Tat–TAR Association and Slow Down HIV Spread in Cell Cultures. *ChemBioChem*. 2007; 8:1850–1856. [PubMed: 17886825]
18. Mayer M, Lang PT, Gerber S, Madrid PB, Pinto IG, Guy RK, James TL. Synthesis and Testing of a Focused Phenothiazine Library for Binding to HIV-1 TAR RNA. 2006; 13:993–1000.
19. Hsieh MCE, Blomquist T, Lustig B. Flexibility of BIV TAR-Tat: models of peptide binding. *Journal of Biomolecular Structure and Dynamics*. 2002; 20:10.
20. Congreve M, Chessari G, Tisi D, Woodhead AJ. Recent developments in fragment-based drug discovery. *Journal of medicinal chemistry*. 2008; 51:3661–80. [PubMed: 18457385]
21. Erlanson DA, McDowell RS, O’Brien T. Fragment-based drug discovery. *Journal of medicinal chemistry*. 2004; 47:3463–82. [PubMed: 15214773]
22. Hajduk PJ, Greer J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat Rev Drug Discov*. 2007; 6:211–9. [PubMed: 17290284]
23. Rees DC, Congreve M, Murray CW, Carr R. Fragment-based lead discovery. *Nat Rev Drug Discov*. 2004; 3:660–72. [PubMed: 15286733]
24. Parsons J, Castaldi MP, Dutta S, Dibrov SM, Wyles DL, Hermann T. Conformational inhibition of the hepatitis C virus internal ribosome entry site RNA. *Nat Chem Biol*. 2009; 5:823–5. [PubMed: 19767736]
25. Seth PP, Miyaji A, Jefferson EA, Sannes-Lowery KA, Osgood SA, Propp SS, Ranken R, Massire C, Sampath R, Ecker DJ, Swayze EE, Griffey RH. SAR by MS: discovery of a new class of RNA-

- binding small molecules for the hepatitis C virus: internal ribosome entry site IIA subdomain. *Journal of medicinal chemistry*. 2005; 48:7099–102. [PubMed: 16279767]
26. Swayze EE, Jefferson EA, Sannes-Lowery KA, Blyn LB, Risen LM, Arakawa S, Osgood SA, Hofstadler SA, Griffey RH. SAR by MS: a ligand based technique for drug lead discovery against structured RNA targets. *Journal of medicinal chemistry*. 2002; 45:3816–9. [PubMed: 12190303]
 27. Lepre CA, Peng J, Fejzo J, Abdul-Manan N, Pocas J, Jacobs M, Xie X, Moore JM. Applications of SHAPES screening in drug discovery. *Comb Chem High Throughput Screen*. 2002; 5:583–90. [PubMed: 12470255]
 28. Johnson EC, Feher VA, Peng JW, Moore JM, Williamson JR. Application of NMR SHAPES screening to an RNA target. *J Am Chem Soc*. 2003; 125:15724–5. [PubMed: 14677945]
 29. Mayer M, James TL. NMR-based characterization of phenothiazines as a RNA binding scaffold. *J Am Chem Soc*. 2004; 126:4453–60. [PubMed: 15053636]
 30. Bodoor K, Boyapati V, Gopu V, Boisdore M, Allam K, Miller J, Treleaven WD, Weldeghiorghis T, Aboul-ela F. Design and implementation of an ribonucleic acid (RNA) directed fragment library. *J Med Chem*. 2009; 52:3753–61. [PubMed: 19445516]
 31. Li D, DeRose EF, London RE. The inter-ligand Overhauser effect: a powerful new NMR approach for mapping structural relationships of macromolecular ligands. *J Biomol NMR*. 1999; 15:71–6. [PubMed: 10549135]
 32. Zhang Q, Al-Hashimi HM. Domain-elongation NMR spectroscopy yields new insights into RNA dynamics and adaptive recognition. *RNA*. 2009; 15:1941–1948. [PubMed: 19776156]
 33. Mayer M, Meyer B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angewandte Chemie International Edition*. 1999; 38:1784–1788.
 34. Noggle, J.; Schirmer, R. *The Nuclear Overhauser Effect*. Academic Press Inc; New York: 1971.
 35. Pitt SW, Majumdar A, Serganov A, Patel DJ, Al-Hashimi HM. Argininamide Binding Arrests Global Motions in HIV-1 TAR RNA: Comparison with Mg²⁺-induced Conformational Stabilization. *J Mol Biol*. 2004; 338:7–16. [PubMed: 15050819]
 36. Dominguez C, Boelens R, Bonvin AM. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc*. 2003; 125:1731–7. [PubMed: 12580598]
 37. van Dijk ADJ, Boelens R, Bonvin AMJJ. Data-driven docking for the study of biomolecular complexes. *FEBS Journal*. 2005; 272:20.
 38. Zacharias M, Hagerman PJ. The bend in RNA created by the trans-activation response element bulge of human immunodeficiency virus is straightened by arginine and by Tat-derived peptide. *Proc Natl Acad Sci U S A*. 1995; 92:6052–6056. [PubMed: 7597079]
 39. Puglisi JD, Tan R, Calnan BJ, Frankel AD, Williamson JR. Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science*. 1992; 257:76–80. [PubMed: 1621097]
 40. Puglisi JD, Chen L, Frankel AD, Williamson JR. Role of RNA structure in arginine recognition of TAR RNA. *PNAS*. 1993; 90:3680–3684. [PubMed: 7682716]
 41. Aboul-ela F, Varani G. Recognition of HIV-1 TAR RNA by Tat protein and Tat-derived peptides. *J Mol Struct*. 1998; 423:29–39.
 42. Yu L, Oost TK, Schkeryantz JM, Yang J, Janowick D, Fesik SW. Discovery of Aminoglycoside Mimetics by NMR-Based Screening of Escherichia coli A-site RNA. *Journal of the American Chemical Society*. 2003; 125:4444–4450. [PubMed: 12683814]
 43. Johnson EC, Feher VA, Peng JW, Moore JM, Williamson JR. Application of NMR SHAPES Screening to an RNA Target. *Journal of the American Chemical Society*. 2003; 125:15724–15725. [PubMed: 14677945]
 44. Varani G, Aboul-ela F, Allain FHT. NMR Investigation of RNA Structure. *Progress in Nuclear Magnetic Spectroscopy*. 1996; 29:51–127.
 45. Becattini B, Pellecchia M. SAR by ILOEs: an NMR-based approach to reverse chemical genetics. *Chemistry*. 2006; 12:2658–62. [PubMed: 16121405]
 46. Begley DW, Zheng S, Varani G. Fragment-based discovery of novel thymidylate synthase leads by NMR screening and group epitope mapping. *Chem Biol Drug Des*. 2010; 76:218–33. [PubMed: 20626411]

47. Piotto M, Saudek V, Sklenar V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR*. 1992; 2:661–5. [PubMed: 1490109]
48. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*. 1995; 6:277–93. [PubMed: 8520220]
49. Goddard TD, Kneller DG. Sparky. :3. unpublished.
50. Ferner J, Suhartono M, Breitung S, Jonker HRA, Hennig M, Wöhnert J, Göbel M, Schwalbe H. Structures of HIV TAR RNA–Ligand Complexes Reveal Higher Binding Stoichiometries. *ChemBioChem*. 2009; 10:1490–1494. [PubMed: 19444830]
51. Aboul-ela F, Karn J, Varani G. Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. *Nucleic Acids Res*. 1996; 24:3974–81. [PubMed: 8918800]
52. Wuthrich, K. *NMR of Proteins and Nucleic Acids*. Wiley; New York: 1986.
53. Andersson P, Nordstrand K, Sunnerhagen M, Liepinsh E, Turovskis I, Otting G. Heteronuclear correlation experiments for the determination of one-bond coupling constants. *Journal of Biomolecular NMR*. 1998; 11:445–450. [PubMed: 9691285]
54. Schwieters CD, Kuszewski JJ, Tjandra N, Marius Clore G. The Xplor-NIH NMR molecular structure determination package. *Journal of Magnetic Resonance*. 2003; 160:65–73. [PubMed: 12565051]
55. Clore GM, Gronenborn AM, Bax A. A Robust Method for Determining the Magnitude of the Fully Asymmetric Alignment Tensor of Oriented Macromolecules in the Absence of Structural Information. *Journal of Magnetic Resonance*. 1998; 133:216–221. [PubMed: 9654491]
56. Koradi R, Billeter M, Wuthrich K. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph*. 1996; 14:51–5. 29–32. [PubMed: 8744573]
57. Schuettelkopf AW, van Aalten DMF. PRODRG - a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica*. 2004; 60:1355–1363.
58. Moriarty N, Grosse-Kunstleve R, Adams P. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallographica*. 2009; D65:1075–1080.
59. Meyer B, Peters T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew Chem Int Ed Engl*. 2003; 42:864–90. [PubMed: 12596167]

Figure 1.

NMR spectra for the small molecule MV2003 (a) in the absence of TAR RNA and (b) in the presence of TAR RNA. A two-dimensional, proton-proton nuclear Overhauser effect (NOE)-NMR spectrum appears in the middle, with positive (dark blue) and negative (cyan) contour peaks. One-dimensional reference (blue) and saturation (red) transfer difference (STD) spectra are shown on the perimeter. MV2003 was selected for fragment screening based on strong STD signal attenuation and positive NOE cross-peaks observed in the presence of sub-stoichiometric amounts of HIV-1 TAR.

Figure 2.

Characterization of binding for fragment hit MV1480. Left: Reference (red), saturation (blue) and difference (black) 1D STD-NMR spectra for (a) MV1480 with sub-stoichiometric HIV TAR RNA present; (b) MV1480 with equimolar MV2003 present; (c) MV1480, MV2003 and TAR at 40:40:1 molar ratios. Right: NOE spectra with peak-aligned STD-NMR data for (d) MV1480 with sub-stoichiometric HIV TAR RNA present; (e) MV1480 with equimolar MV2003 present; (f) MV1480, MV2003 and TAR at 40:40:1 molar ratios. Spectral binding signatures which appear in (c) and (f) indicate that MV1480 only binds to HIV TAR RNA when MV2003 is present in solution

Figure 3.

Structure of the binary MV2003/TAR complex. (a) On the far left, 10 representative structures are overlaid, with the RNA presented as blue lines and MV2003 as red lines. Left center is a surface representation of the RNA in gray, with MV2003 in red. Interactions of the small molecule MV2003 groups with the RNA are displayed on the right. The guanidinium group is located near the U23-C24-U25 bulge and likely stabilizes partial formation of the U23-A27-U38 base triple (right center), while the naphthylene ring system stacks at the top of the upper helix (far right). (b) The structure of the MV2003/TAR complex (far right) reveals a cavity, highlighted in orange, not observed with previous structures of TAR bound to other small molecules (surface images are shown in stereo). PDB accession numbers for these structures, left to right, are 1uts, 1aju, 1lvj, and 2l8h.

Figure 4. HADDOCK docking results for MV1303, MV1315 and MV1480, showing orientations for each fragment against the TAR-MV2003 complex. a.) The dashed lines represent the ILOE data used as experimental restraints for modeling. b.) The 10 structures with the fewest number of NOE violations and lowest pair-wise RMSD from the average structure are overlaid for each fragment.

Figure 5. Three fragment hits (left) and two non-hits (right) from scaffold Group 11 screened against HIV TAR RNA in the presence of MV2003. Even small changes in the fragment lead to complete loss of binding activity.

Table 1

Commercially available small arginine mimics tested for binding to HIV TAR RNA by ligand-observe NMR spectroscopy.

ID	Chemical Name	MW	Structure	STD with TAR	NOE with TAR
MV 2001	L-arginine methyl ester	261.2		W	W
MV 2002	L-arginine ethyl ester	275.2		W	W
MV 2003	Arginine 4-methoxy-β-naphthylamide	365.9		S	S
MV 2004	PTH-arginine	327.8		W	N
MV 2005	N, N-Dimethyl butylamine	101.2		N	N
MV 2008	Nα-carbo Benzoyloxy -L-arginine	308.2		N	N
MV 2009	Nα-Benzoyl -L-arginine	278.3		NS	NS
MV 2010	[1-(3-amino propyl) piperidin-3-yl]methanol	172		W	N
MV 2013	Benzoyl arginine-2 -naphthyl amide	439.9		NS	NS
MV 2014	Arginine-7-amido-4 -methyl coumarine	331.37		NS	NS
MV 2015	Arginine β-naphthyl amide	335.83		S	S

ID	Chemical Name	MW	Structure	STD with TAR	NOE with TAR
MV 2016	α -Dansyl -arginine	443.95		W	N
MV 2017	Guanfacine	282.55		N	N
MV 2018	Clonidine	266.55		N	N
MV 2019	H-Gly-Arg-4MbNA	459.38		S	S
MV 2020	H-Arg-pNA	367.24		M	M

S = strong binding; M = moderate binding; W = weak binding; N = no binding; NS = not soluble.

Top 20 primary screening fragment hits from 13 different scaffold groups and follow-up results from NMR experiments on individual compounds

Table 2

MV ID	structure	Group	% STD	NOE crosspeaks			
				Alone	+ MV 2003	+ TAR RNA	+MV 2003 + TAR RNA
MV 1429		1	59	-	-	-	+
MV 1348		1	57	-	+		
MV 1315		2	53	-	-	-	+
MV 1459		3	59	-	-	-	+
MV 1346		3	56	-	-	-	-
MV 1425		4	34	-	-	+	
MV 1292		5	60	-	+		

MV ID	structure	Group	% STD	NOE crosspeaks			
				Alone	+ MV 2003	+ TAR RNA	+MV 2003 + TAR RNA
MV 1437		6	51	-	-	-	-
MV 1312		7	67	-	+	-	-
MV 1072		7	67	-	-	-	+
MV 1491		7	62	+	-	-	-
MV 1303		8	68	-	-	-	+
MV 1074		8	68	-	-	-	-
MV 1497		8	61	+	-	-	-
MV 1206		8	66	+	-	-	-
MV 1062		9	64	-	-	+	-
MV 1159		10	49	-	-	+	-

MV ID	structure	Group	% STD	NOE crosspeaks			
				Alone	+ MV 2003	+ TAR RNA	+MV 2003 + TAR RNA
MV 1480		11	85	—	—	—	+
MV 1099		12	28	—	—	—	—
MV 1027		13	35	—	—	—	—

Table 3

Statistics for HADDOCK-driven modeling of the ternary complexes, generated by docking individual fragments to the TAR-MV2003 binary complex structure

Fragment	Statistics for 10 reported docking results	
MV1303	# ILOEs used for docking	5
	Avg. # ILOES violated (>0.5 Å)	0
	Avg. # ILOES violated (>0.3 Å)	1
	Avg. RMSD from mean (interface)	1.0
	Avg. RMSD from mean (all)	2.1
MV1315	# ILOEs used for docking	8
	Avg. # ILOES violated (>0.5 Å)	0
	Avg. # ILOES violated (>0.3 Å)	0
	Avg. RMSD from mean (interface)	1.4
	Avg. RMSD from mean (all)	2.1
MV1480	# ILOEs used for docking	8
	Avg. # ILOES violated (>0.5 Å)	0
	Avg. # ILOES violated (>0.3 Å)	2
	Avg. RMSD from mean (interface)	1.4
	Avg. RMSD from mean (all)	2.9