Noncoding (nc)RNAs exhibit remarkable architectural diversity that contributes to function in multiple gene-regulatory settings (1). Although the human proteome is derived from only a small fraction of the genome (0.05%), the preponderance of the DNA blueprint is transcribed into ncRNA (2, 3). As a result, ncRNA transcripts provide important opportunities to intervene in a range of biological processes and diseases (4–6). Such pursuits are especially meaningful in light of the fact that only 3.5–10% of the proteome is likely to be druggable (7, 8). Substantial evidence demonstrates that a handful of ncRNAs adopt elegant three-dimensional folds with distinct topologies and recurrent architectural motifs (9–14), including cavities and deep grooves predisposed to ligand binding (7, 15, 16). These properties are suited for shape-specific recognition of small molecules or peptides and provide a basis to manipulate conformation or dynamics to alter downstream function. Several notable achievements accentuate such efforts, including the identification of inhibitors that target the following: cancer-associated miR-21; CUG repeats of myotonic dystrophy; riboswitches in pathogenic bacteria; and exon splicing in spinal muscular atrophy (17–28). These successes underscore the feasibility of sequence-specific targeting of RNAs to create research tools or as a means to treat human disease. Accordingly, delineating principles of molecular recognition represents a cornerstone for therapeutic design, especially as part of a combination-drug strategy to circumvent drug resistance by pathogens that undergo multiple genomic mutations per generation (26, 29).

New mandates in HIV eradication and cure research (https://grants.nih.gov/grants/guide/notice-files/NCT-OD-15-137.html) (30, 31) have led to a resurgence in efforts to target the transactivation response (TAR) element. This 59-nucleotide RNA is located in the 5′-LTR of all viral transcripts and features a conserved hairpin that harbors an apical loop and pyrimidine-rich bulge that are each indispensable for transactivation (33–40). TAR interacts with the viral Tat protein, which recruits the host pTEFb complex away from inactivating HEXIM–7SK RNA complexes (Fig. 1A) (41–47). When localized to TAR, host kinase CDK9 within a super-elongation complex (SEC) phosphorylates RNA polymerase II, releasing it from a paused state to produce full-length viral mRNA (48–50). Structural and biochemical analysis of the Tat–pTEFb complex revealed Tat-specific conformational changes (51, 52). Because of the essential role of Tat in securing pTEFb for processive viral transcription, efforts are focused on the development of inhibitors that block this key host–virus protein interaction (53–55). Sustained inhibition of Tat could lock HIV into a state of deep latency and represents one strategy to produce a functional cure (56, 57).
Targeting TAR RNA represents a fundamentally different antiviral approach (58). TAR is one of the most conserved RNA sequences in the viral genome (Fig. 1B). In addition to SEC binding, TAR functions as a pre-miRNA whose Dicer cleavage products block host–cell apoptosis, prolonging the viral life span in infected cells (59–61). For these reasons, TAR is a high-value drug target whose inhibition could potentially disrupt viral transcription in chronic as well as latent infections. However, no such inhibitors are clinically available, and TAR has resisted the development of therapeutics, despite success in the identification of compounds that target the RNA with specificity and affinity (62–64).

As with other RNA molecules, TAR is dynamic and adopts multiple conformations (65–67), undermining efforts to obtain high-resolution crystal structures (68). NMR has bridged many gaps in our understanding of TAR with nearly 20 distinct structures determined of the isolated (apo) RNA and in complex with peptides or small molecules (61, 62, 67–81). Even so, TAR has been historically challenging for NMR due to significant line broadening of resonances, which has hindered the acquisition of restraints needed to generate high-quality models (69–71). For this reason, many efforts have focused on closely related HIV-2 or BIV TAR (Fig. 1C), which are structurally better defined (69, 72, 73). Additional structural improvements have been attained through engineered RNA constructs to promote crystal contacts or by exploiting structurally well-characterized proteins, such as U1A, as a starting platform for lab-based evolution and structural studies. These developments have led to a series of new structures including: exciting TAR–Tat and TAR–Tat–SEC complexes (74–76), a co-crystal structure of TAR bound to a lab-evolved protein from the Wedekind lab (77), and an ultrahigh-affinity cyclic peptide bound to TAR (63). Here, we put these novel discoveries into perspective by considering prior characterization of TAR apo- and bound-state conformations. We then consider molecular recognition by representative small molecules, which are then contrasted with recent high-quality ncRNA–inhibitor complexes. A major take-home message is that peptide-mediated TAR recognition utilizes some common molecular-recognition principles, such as the arginine-sandwich motif (ASM)—a primary determinant of affinity and specificity observed in both natural TAR–Tat complexes, as well as TAR binding by a lab-evolved protein. In contrast, no consistent rules of recognition could be discerned for existing TAR–small molecule complexes, despite the use of common guanidinium groups. As the reader will see, new TAR–peptide and TAR–protein complexes offer the most cogent details to address challenges and opportunities associated with effective TAR targeting. In this respect, the best days of RNA drug discovery appear to lie ahead.

TAR adopts two major conformations that depend on ligand binding

The discovery of TAR–Tat-mediated gene regulation in HIV-1 (78) started a race to elucidate the underlying molecular determinants that give rise to this unique viral RNA–protein interaction. Major steps were made by NMR analyses of TAR in complex with the arginine analogue argininamide and in a ligand-free (apo) state. This work revealed TAR’s overall hairpin architecture as well as substantial backbone rearrangements at the central bulge resulting from ligand binding (79, 80). Indeed, when specific effectors interact with the major groove, the RNA adopts a slightly bent (~165°) helical axis...
formed by coaxial stacking of stem 1a and stem 1b (s1a and s1b) (Fig. 2A), wherein the bases of the central bulge jut outward. This conformation exhibits a high degree of concave surface suited to ligand binding (Fig. 2, A and B). The TAR major groove is characterized by a narrow width (3.9 ± 0.5 Å) and substantial depth (10.3 ± 0.3 Å) reminiscent of an ideal A-form duplex (i.e., 2.7 Å wide by 13.5 Å deep (81)). In contrast, the minor-groove width (9.9 ± 0.6 Å) and depth (1.0 ± 0.6 Å) are substantially wider and shallower than a typical A-form helix (5.7 Å wide by 7.5 Å deep (81)). A hallmark of the ligand-bound conformation is that Uri23 interacts with the Hoogsteen edge of a nearby adenine to form a Uri23–Ade27–Uri38 base triple (Fig. 2C)—a feature observed in most peptide- and protein-bound TAR structures (70, 75, 77, 79, 82). Cyt24 and Uri25 extrude from the helical core with bases pointing into solvent. Molecular dynamics simulations of TAR in complex with a lab-evolved protein revealed that this long-range triple is preserved over 16 μs but disintegrates rapidly when the protein is omitted from the simulation (77).

In the absence of interacting ligands, the TAR helical axis is bent more acutely to 121° (Fig. 2D). The major groove is extraordinarily wide (13.1 ± 4.2 Å) and shallow (4.4 ± 3.2 Å) compared with an A-form helix. These features are accompanied by a relative reduction of concave surface in the major groove (Fig. 2, D and E versus A and B), a property that is less conducive to binding by small molecules or peptides (83). Because Uri23 and Cyt24 reside inside the duplex without base pairing to the adjacent strand, the s1a–s1b coaxial stack is wedged apart yielding an underpacked core (Fig. 2, D and F) (80). Whereas Uri23 and Cyt24 adopt stacked and inclined base orientations relative to underlying base pairs, Uri25 loops out of the duplex. These features prohibit formation of the hallmark base triple, leaving only the canonical Ade27–Uri38 pair (Fig. 2F). As a result, the central bulge exhibits significantly more conformational flexibility in the apo-state compared with the ligand-bound state, as observed by NMR analysis and molecular dynamics simulations (77, 80).

Overall, the propensity of TAR to adopt two major bulge conformations is well-suited to ligand binding. Solution studies showed that the major groove is narrow and deep in the presence of ligand (73, 79, 84). As such, the RNA is capable of folding around a ligand—such as the unstructured peptide of Tat—giving rise to a complementary interface with a substantial buried surface area (76). Understanding the details of such RNA–peptide interactions provides insight into the basis for affinity and specificity, while revealing stereochemical features that are unique to the respective apo- and ligand-bound states. Such information is of high value for the design of novel antivirals that target the HIV TAR element.

### Targeting TAR with small molecules

During the past 2 decades, multiple labs have worked to identify small molecules that bind HIV-1 TAR (64, 85–91). To gain perspective about the successes and ongoing challenges, it is instructive to examine the handful of structurally characterized TAR–small molecule complexes to assess compound localization and commonalities in their modes of molecular recognition. A survey of such complexes (Table 1) reveals common
chemical features, including positively charged alkylamine or guanidinium groups and planar heteroaromatic groups, such as naphthyl, indole, phenyl, or phenothiazine moieties. Although neomycin and derivatives thereof are known to bind TAR (92), we will not consider aminoglycosides here, due to their promiscuous RNA binding resulting from many positively charged amines (Fig. 3A), their toxicity (86, 93–95), and the recent focus on compounds with “drug-like” properties in terms of potency, solubility, selectivity, and distribution, as well as RNA targeting by use of specific modes of molecular recognition (96).

Historically, arginine is one of the first small molecules shown to bind TAR, leading to a conformational change (97, 98). This amino acid—and derivatives thereof—has served as a proxy for selective binding of the Tat protein to TAR. This key interaction arises from specific contacts at Uri23, Gua26, and Ade27 (97, 99–101). Because more structural restraints were discernible for the HIV-2 TAR–argininamide complex compared with HIV-1 TAR–argininamide, the former analysis is considered to provide a definitive basis to evaluate this RNA–ligand interaction (69). Indeed, HIV-2 TAR differs from HIV-1 by deletion of Cyt24 in the UCU bulge (Fig. 1C). Both HIV TAR variants have similar $K_I$ values of $\sim 2 \text{nM}$ for argininamide (Fig. 3A). The HIV-2 TAR–argininamide NMR ensemble reveals that the ligand localizes to the major groove near the central bulge, where the guanidinium moiety engages in cation–π stacking between Ade22 and Uri23 (Fig. 3, B and C). Although NMR spectra did not provide direct evidence for hydrogen bonding between argininamide and Gua26 (77), the guanidinium position and orientation are consistent with co-planar hydrogen bonding predicted by theoretical calculations (69). One metric of surface complementarity at the receptor–ligand interface is the shape correlation statistic ($S_S$) (102). The calculated $S_S$ of 0.67 for the HIV-2 TAR–argininamide interaction (69) suggests substantial surface complementarity. The complex buries 229 Å² of the argininamide solvent-accessible surface, which is 65% of the total ligand surface area. The observation that this RNA–ligand complex shares NOEs with the HIV-1 TAR–argininamide complex suggested similar modes of effector binding (79, 103). Although the hallmark Uri23–Ade27–Uri38 triple of the bound state was absent in NOE assignments in one study (103), another study revealed that an isomorphic C⁺⁻G⁻⁻C mutant interacts with argininamide in a pH-dependent manner, suggesting base-triple formation is needed for amino acid binding (84). Argininamide binding to TAR provided several insights in terms of ligand localization and the determinants of binding (Fig. 3, B and C). As we will see, this mode of binding—known as an arginine sandwich motif (ASM)—was observed next in the context of TAR–Tat interactions (described below). Of course, high-affinity Tat-mediated recognition requires multiple arginines (101, 104, 105). This knowledge and the application of electrostatic analysis to the TAR–argininamide complex prompted high-throughput screening of bis-guanidine compounds designed to mimic argininamide binding. Based on Tat–peptide-displacement assays, a top hit, RBT-203 (Fig. 3A), showed a $K_I$ of 1.5 μM by FRET displacement, evidence of binding by surface plasmon resonance (SPR), as well as inhibition of Tat-mediated transcription in cell-free extracts at levels of 5–15 μM (106, 107). NMR analysis revealed the compound induces a conformation similar to that of the TAR–argininamide complex, although neither guanidinium group of RBT-203 interacts with a guanine base (107). Addition of an indole ring into the RBT-203 benzyl scaffold and replacement of the guanidinium groups by piperazine and a primary amine improved the $K_I$ to 39 nM, although antiviral activity was not assessed (106). This new compound, RBT-550 (Fig. 3A), was shown by NMR to bind TAR in a fundamentally different manner compared with argininamide. The indole ring appears to intercalate adjacent to the UCU bulge between the Gua26–Cyt39 and Ade22–Uri40 base pairs (Fig. 3D). Uri23 does not form the hallmark base triple, and the primary amine of RBT-550 interacts with the Gua26 backbone; the piperazine moiety protrudes into solvent but appears to restrict the propylamine conformation in some members of the structural ensemble. Intercalation produces a high degree of shape complementarity (average $S_S$ of 0.67), and 290 Å² of the ligand is buried in the interface, representing 44% of its solvent-accessible surface. The observation that RBT-550 promotes a TAR conformation that differs from the argininamide-bound complex lends support to the idea that some small molecules can shift the RNA conformational equilibrium to an “inactive” state, which is a generally accepted drug strategy (66, 92, 106, 108).

Computational screening of a small molecule library identified acetylpromazine (Fig. 3A) as a TAR binder, providing an early example of how this approach could be used to target RNA. Electrophoretic mobility shift assays suggested that the compound blocks formation of a TAR–Tat–CycT1 complex at $\sim 100 \text{nM}$ (64). NMR analysis indicated that acetylpromazine localizes within the bulged loop of TAR in a manner analogous to RBT-550 (109). Binding appears to be conferred primarily by stacking between the Gua26–Cyt39 and Ade22–Uri40 base pairs like RBT-550 and is accompanied by dissolution of the Uri23–Ade27–Uri38 base-triple. Like RBT-203 and RBT-550, there are no base-specific interactions comparable with Hoogsteen-edge readout by argininamide (Fig. 3C). The RNA–ligand

### Table 1

| TAR type | Ligand | Ligand scaffold | Method | PBD codes | Refs. |
|----------|--------|----------------|--------|-----------|-------|
| HIV-1    | Arginine | Amino acid | NMR    | 1arj      | 103   |
| HIV-2    | Argininamide | Amino acid derivative | NMR    | 1aks and 1aju | 69    |
| HIV-1 (Δ apical loop) | Calcium ion | Divalent ion | X-ray | 397d | 112 |
| HIV-1    | Neomycin | Aminoglycoside | NMR    | 1q43      | 92    |
| HIV-1    | Acetylpromazine | Phenothiazine drug | NMR    | 1lvj      | 109   |
| HIV-1    | RBT-550 | Benzyl indole | NMR    | 1ats      | 106   |
| HIV-1    | RBT-203 | Methoxy phenoxy | NMR    | 1uui      | 107   |
| HIV-1    | RBT-158 | Benzylpiperazine | NMR    | 1uui      | 107   |
| HIV-1    | Chemical probe | 4-Methoxynaphthalene-2-amine | NMR    | 288h      | 95    |
interface buries 428 Å² of solvent-accessible surface area (83% of the total), and the average $S_c$ is 0.62, suggesting a modest degree of shape complementarity.

**Model ncRNA–inhibitor interactions: base pairing and shape complementary**

Small molecules that strongly target a specific RNA are uncommon, and these are likely to engage multiple unintended partners (91). At the outset of screening, effective approaches strive to limit off-target recognition by conducting binding assays in the presence of a molar excess of tRNA (85), or by gauging nonspecific binding by use of decoy RNAs (85, 86), RNase footprinting (110), or whole transcriptome analysis (111). Even after the identification of a tight-binding RNA inhibitor, the structuredetermination of such a complex is even more extraordinary. As we noted, many technical obstacles were overcome to obtain reliable experimental structures of TAR (69, 112). To improve such outcomes, the analysis of TAR binding to various small-molecule ligands would have benefitted from complementary biophysical approaches to rigorously and reproducibly assess the binding determinants of hit compounds (113–115). Methods that provide thermodynamic parameters ($\Delta G$, $\Delta H$, and $-T\Delta S$) and $K_D$ values have proven especially useful to relate structural observations to specific
modes of binding (116). SPR is also considered a rigorous secondary screen to validate high-throughput approaches, while providing kinetic constants $k_{\text{on}}$ and $k_{\text{off}}$ for lead optimization (117). A future challenge for inhibitor studies of TAR will be to relate quantitatively vetted molecular-recognition attributes of ligand binding to the drug-discovery process. Accordingly, we now consider examples of well-defined ncRNA–effector complexes with distinct RNA recognition features, supporting equilibrium binding constants, and analyses of downstream inhibitor effects on antiviral or antibacterial function.

Benzimidazole derivatives have been identified by MS-based screening that target the internal ribosome entry site (IRES) of the hepatitis C virus (HCV) genomic RNA (118). The IRES features a series of folded domains, including conserved domain II. This region comprises a bent, bulged loop that is key for positioning the viral mRNA initiation codon and activation of the host ribosome (119). Biophysical analysis demonstrated benzimidazole compounds straighten domain IIa and reorder the bulge (110, 120), which has parallels to the ligand-bound and apo-states of TAR (Fig. 2, A and D). Significantly, the restructured S-shape of domain IIa produces a cavity suited to small-molecule binding.

Lead optimization led to compound 12 that binds HCV IRES domain IIa with a $K_D$ of 860 nM (Fig. 4A) (118). A 2.2 Å resolution co-crystal structure reveals the mode of RNA recognition by 12 (Fig. 4B). Specifically, the 2-aminoimidazole moiety donates hydrogen bonds to the Hoogsteen edge of Gua110, like argininamide (Fig. 3C). The dimethylamino-propyl group makes an electrostatic interaction with a nonbridging oxygen of Ade109, whereas the dimethylamino-methyl group forms a water-mediated contact to a nonbridging oxygen of Ade53. The benzimidazole moiety engages in $\pi$-stacking between purines Gua52 and Ade53. These features are corroborated strongly by structure–activity relationships (118). Compound 12 sequesters 384 Å² of its surface in the RNA pocket or 71% of the ligand’s solvent-accessible area. As expected from the structure, the shape correlation between the RNA and ligand surfaces is high with an $S_C$ value of 0.82. Importantly, compound 12 was also active in HCV-replicon assays. The inhibitor reduced HCV RNA levels in cells with an EC50 of 3.9 μM (118), similar to levels needed to inhibit translation from subgenomic replicons (2.8 ± 0.4 μM) and full-length virus (3.4 ± 0.5 μM) (28). Although a related compound 11 showed slightly poorer binding affinity ($K_D$ 1.7 μM), it performed better in the replicon assay (EC50 of 1.5 μM). Compound 11 replaces the tetrahydropyran ring with a smaller tetrahydrofuran. This subtle difference has been attributed to differences in cellular penetration (118), which is a major consideration beyond a tight-binding $K_D$.

Riboswitches represent another class of structured ncRNAs that change their conformations in response to the binding of cognate ions or small molecules (121–123). Such sensing results in mRNA regulatory feedback that controls downstream genes (124–126). The flavin mononucleotide (FMN) riboswitch is notable because it senses the cofactor FMN as well as the natural product roseoflavin (23, 127). This vulnerability has fostered efforts to target the
FMN riboswitch with novel antibiotics (19, 21, 22, 25, 26). Recent structure-guided design led to the discovery of compound BRX1555—an FMN analogue that binds with a $K_D$ of 39.0 ± 0.7 nM based on in-line probing (Fig. 4C) (25). The ligand has an $\text{EC}_{50}$ of $1.70 \pm 0.18 \mu M$ in single-turnover transcription termination assays and an $\text{IC}_{50}$ of 0.49 ± 0.09 μM in bacterial growth inhibition assays (25).

The 2.80 Å resolution co-crystal structure of the FMN riboswitch in complex with BRX1555 reveals key details about its mode of molecular recognition. As expected, the inhibitor overlaps with the binding site of the natural ligand, which resides at the center of a six-way helical junction comprising two pairs of stabilizing loops (128). The isoalloxazine ring of the inhibitor stacks centrally between Ade48 (junction 3-4) and Ade85 (pairing region P6) (Fig. 4D). The face of Ade99 (J6-1) hydrogen bonds to the uracil-like edge of BRX1555 in a manner similar to FMN. Gua62 (J4-5) stacks against the phenyl group of the inhibitor, reminiscent of the 2-methylaminopurinidine moiety of ribocil—the synthetic FMN analogue discovered by Merck (21). In terms of binding and localization, the similarities of BRX1555 and ribocil are remarkable, especially because the former molecule was developed by structure-based design and the latter was identified by phenotypic screens that yielded a novel chemical scaffold distinct from FMN (21, 25). Like FMN and ribocil, the riboswitch–BRX1555 complex buries a large amount of the inhibitor’s solvent-accessible surface in the interface (468 Å² or 88%). The riboswitch–BRX1555 complex also shows significant shape complementarity, as indicated by an $S_c$ value of 0.72. Interestingly, significant commonalities exist in the interactions used by HCV IRES domain IIIa and the FMN riboswitch in terms of ligand recognition; these likenesses include hydrogen bonding that imparts base-specific readout, co-axial base stacking, solvent exclusion, and high shape complementarity (Fig. 4, B and D). These features also represent key molecular recognition determinants in peptide binding to TAR, which we will now explore.

**Molecular recognition of TAR by Tat peptides**

The HIV-1 Tat protein comprises multiple functional domains that are needed to complete the viral life cycle (Fig. 5A). TAR binding requires a basic ARM (37, 99) harboring nine arginines. Of these, Arg-52 is most essential because its mutation to lysine yields a drastic loss of transactivation (99). Mutagenesis of flanking residues supports the presence of a supplemental electrostatic interaction network that modulates RNA binding as well as transactivation (100, 129). Thus far, elucidation of the intact TAR–Tat–SEC complex (Fig. 1A) has remained elusive, although divide-and-conquer efforts have led to core SEC complexes in the presence of Tat’s transactivation domain. Nevertheless, these co-crystal structures currently lack the Tat ARM domain (51, 74), providing an incomplete picture of RNA recognition. Accordingly, we will now focus on recent structures of Tat-derived ARM peptides in complex with TAR that have led to a new understanding of this key RNA–protein interaction and how it provides a foundation for HIV inhibitor design. A structural survey of known peptides and proteins bound to TAR is presented in Table 2.

To provide perspective on the recent structure of the HIV-2 TAR–Tat complex, it is important to recognize that initial high-resolution insights came from NMR studies of the BIV TAR–Tat complex (73, 75). Like HIV-1 Tat, the ARM domain of BIV Tat is also arginine-rich (Fig. 5B). The peptide binds BIV TAR in the major groove near the central UU bulge, where it forms a short antiparallel strand capped by a distorted type V $\beta$-turn (Fig. 5C) (130). Like many $\beta$-turns, the $\iota$th to $\iota$ + 3rd hydrogen bond is absent, but the carbonyl oxygen of the $\iota$th residue (Arg-73) receives a hydrogen bond from the $\iota$ + 4th side chain (Arg-77) (Fig. 5D). The net result is a $\beta$-hairpin spanning the width of the major groove. Base-specific readout is mediated by guanidinium groups from Arg-70, Arg-73, and Arg-77, which hydrogen bonds to the Hoogsteen edges of Gua14, Gua11, and Gua9. Cation–$\pi$ stacking is observed between Arg-70 and Ade13 of the central base triple and between Arg-73 and Gua9. A handful of salt-bridge and hydrogen-bond interactions occur, including Lys-75 Ne to the pro-($R_p$)-oxygen of Uri24 and the backbone amide of Gly-71 to N7 of Gua22. The complex buries 62% of the total Tat peptide (Ser-65 to Arg-81) solvent-accessible surface or 1187 Å². The interface exhibits a substantial amount of shape complementarity, as indicated by an $S_c$ value of 0.70. These molecular recognition properties are consistent with the $K_D$ of 1.3 ± 0.1 nM measured for this strong peptide–RNA binding interaction (131).

More recently, the solution–NMR structure of the HIV-1 Tat peptide (amino acids 44–60) was determined in complex with HIV-2 TAR (76). This exciting new complex reveals unprecedented chemical details about the mode of TAR–Tat molecular recognition (Fig. 5B). Remarkably, the Tat ARM spans the length of the TAR major groove, starting with the N terminus abutting the well-ordered apical loop (Fig. 5E). This tight RNA turn is fortified by a canonical Cyt30–Gua34 base pair first observed in the HIV-1 TAR complex with the lab-evolved protein TBP6.7 (77) (discussed below). The C terminus of the peptide extends through stem s1a (Fig. 5E) and protrudes into solvent past Arg-57. Consistent with CD spectra of the isolated peptide in solution (129), the ensemble of Tat conformers in the bound state lacks regular secondary-structure features in contrast to the $\beta$-turn in BIV Tat (Fig. 5, E versus C).

As anticipated, the determinants of TAR–Tat binding specificity include key arginines that read the Hoogsteen edges of conserved guanine bases in the TAR sequence (Fig. 5, B and F). The indispensable nature of Arg-52 (97, 100) is consistent with its recognition of Gua26 (Fig. 5F)—the site of argininamide binding (Fig. 3C). Arg-52 is sequestered by cation–$\pi$ stacking of its guanidinium group between bases from Ade22 and Uri23. The latter base engages in the hallmark bound-state base triple. The constellation of bases and mode of amino acid recognition compose the specialized ASM protein–RNA interaction module (Fig. 5F) (132)—first observed for argininamide (above). Although the ASM appears only once in the TAR–Tat complex, it is utilized four times in the 7SK–Tat complex (data not shown) (76). Arg-73 of BIV Tat uses comparable ASM-like readout, although the Arg-73 guanidinium group does not stack beneath the Uri10 base (Fig. 5D).

A different mode of TAR recognition is used by the Arg-49 group of Tat, which also hydrogen bonds to the Hoogsteen edge
of a conserved guanine (i.e. Gua28), while making contacts to the \( \Phi^2 \cdot \Omega \) of Uri23 (Fig. 5, B and F). Although the latter nucleobase stacks upon the Arg-49 side chain, this binding mode does not constitute an ASM because the guanidinium is not flanked by bases on both sides (i.e. it is an "open-faced" arginine sandwich). Beyond arginine, Tat uses additional stabilizing hydrogen bonds to recognize TAR in the upper and lower stems. These include the following: the e-amino groups of Lys-50 and Lys-51, which interact with backbone oxygens from Gua36 and Cyt37; the Gly-48 carbonyl oxygen, which interacts with the exocyclic amine of Cyt29 (Fig. 5, B and F); and Arg-53 and Arg-55 from the flexible C-terminal tail of Tat, which interact with the backbone at Cyt39 and Uri40, whereas Gln-54 recognizes the TAR major groove.

The HIV-1 Tat peptide recognizes TAR with a \( K_D \) of 22.5 ± 15.2 nM based on ITC (76). This slightly reduced affinity compared to BIV TAR–Tat represents a change in free energy.
(ΔΔG) of only +1.7 kcal mol⁻¹, e.g. the difference of 2–3 hydrogen bonds. Like the BIV TAR–Tat complex, the lowest energy peptide of the HIV TAR–Tat ensemble is significantly sequestered in the major groove with 41% of the peptide (1185 Å²) buried from solvent. This degree of similarity is striking, considering that the HIV-1 Tat peptide adopts an extended conformation compared with the BIV U-shaped polypeptide path (Fig. 5, E versus C). As expected, the HIV TAR–Tat interface exhibits substantial shape complementarity in its core, indicated by an S₁ value of 0.66—comparable with antibody–antigen interfaces and peptides designed to inhibit β-amyloid aggregation (102, 133).

### Lab-evolved proteins for HIV-1 TAR recognition

Advances in protein engineering have facilitated the design of novel RNA-binding proteins with distinct functions (134). In this regard, the TAR-binding protein (TBP) is a model system that was “evolved” from RR1M of the U1A spliceosomal protein by combining saturation mutagenesis, yeast display, and cell sorting (77, 85). The unique mode of HIV-1 TAR recognition by variant TBP6.7 was visualized recently by a co-crystal structure determined to 1.80 Å resolution (77). Unexpectedly, TAR recognition by TBP6.7 entails doubled-stranded RNA recognition of s1b and the UCU bulge (Fig. 5G). This mode of binding differs entirely from the parental U1A protein, which binds to a single-stranded loop within the U1 small nuclear RNA (156). The major determinants of TAR RNA recognition by TBP6.7 are attributable to residues in the evolved β2–β3 loop. For rigor, every amino acid in the loop was mutated and analyzed for TAR binding by ITC, thereby relating structure and recognition in terms of free-energy changes. Arg-47, Arg-49, and Arg-52 are the most energetically significant residues as reflected by their ΔΔG values of +3.8, +3.2, and +2.8 kcal mol⁻¹ for Arg→to-Ala mutations. These observations agree well with the structure wherein each residue penetrates deeply into the major groove to recognize a conserved guanine. Like the HIV-1 TAR–Tat interaction, Arg-47 utilizes the ASM in which its guanidinium group stacks between Ade22 and Uri23, while forming hydrogen bonds to the Hoogsteen edge of Gua26 (Fig. 5, B and H). Unlike the modes of TAR RNA recognition by BIV and HIV Tat peptides (Fig. 5, D and F), Arg-47 simultaneously makes two electrostatic contacts to Uri23 phosphate. The collective interactions appear to be a variation of a hypothetical “arginine fork” interaction, wherein both edges of the Tat-derived guanidinium group were hypothesized to bind TAR’s phosphate backbone (99). Otherwise, the TAR–TBP6.7 complex typifies the TAR-bound conformation featuring the hallmark Uri23·Ade27–Uri38 base triple and canonical Cyt30–Gua34 pair in the apical loop (77).

Other similarities exist between the modes of HIV-1 Tat and TBP6.7 recognition of TAR. Specifically, Arg-49 of TBP6.7 stacks upon Ade27 while hydrogen bonding and engaging in electrostatic interactions with the Gua28 Hoogsteen edge and phosphate group (Fig. 5H). Arg-49 of HIV-1 Tat forms similar stacking and base-pairing interactions but hydrogen bonds to the 2’-OH of Uri23 (Fig. 5F). Unlike the HIV-1 TAR–Tat complex, TBP6.7 uses a third arginine for guanine recognition. Arg-52 of TBP6.7 reads the Hoogsteen edge of Gua36 while stacking beneath Gua34. Beyond TBP6.7, BIV Tat is the only other example of major-groove guanine recognition by three peptide arginines (Fig. 5, B and D). Despite similarities in TAR recognition among TBP6.7, HIV-1 Tat, and BIV Tat, the commonalities are entirely local and do not reflect common poly-peptide folds (Fig. 5, C, E, and G). In terms of the buried surface area and shape complementarity of the TAR–TBP6.7 interface, a total of 718 Å² of TAR is sequestered, wherein 384 Å² is attributable to the β2–β3 loop. Recognition of TAR by TBP6.7 gives an S₂ value of 0.79 (77), which is comparable with S₂ values of peptides selected by phage display to bind the insulin receptor ectodomain (136). Overall, these properties closely resemble comparable metrics for the BIV and HIV TAR–Tat complexes (Fig. 5, C, E, and G).

The observation that the major determinants of TAR recognition by TBP6.7 are localized mostly to the lab-evolved β2–β3 loop has ramifications for inhibitor design using a short peptide that comprises the isolated β2–β3 loop. Indeed, a series of complementary experiments demonstrated that the β2–β3-loop sequence could be removed from the context of TBP6.7 and was still capable of TAR binding. When synthesized as a stapled peptide, the restrained β2–β3 loop still exhibited affinity for TAR (K_D of 1.8 ± 0.5 μM) and was capable of inhibiting TAR–Tat-dependent transcription in HeLa nuclear lysate (77). At present, it is unknown whether stapled β2–β3-loop peptides enter cells or whether they possess antiviral activity. Nevertheless, this work provides proof–of–principle that small peptides can be derived from proteins evolved in the lab to recognize TAR.

### TAR recognition by structure-based design of cyclic peptides

A more traditional approach to disrupt the SEC–TAR interaction (Fig. 1A) is to exploit existing knowledge of TAR–Tat molecular recognition to guide design of restrained, inhibitory peptides (62, 70). Past studies leveraged structural information from the BIV TAR–Tat interaction (Fig. 5, C and D) to produce a number of cyclic peptides (82, 137, 138) that culminated recently in an “ultrahigh affinity” cyclic peptidomimetic. This inhibitor, JB181, binds HIV-1 TAR with an unprecedented K_D...
of 28 ± 4 pm (63). NMR solution analysis revealed that JB181 recognizes TAR in the s1b major groove and bulge (Fig. 5I). However, rather than adopting an elongated peptide as observed for the HIV-2 TAR–Tat complex, the designed peptide forms a β-hairpin comprising 14 residues (Fig. 5B). To reduce conformational flexibility, the peptide termini are linked by an innovative ε- and δ-proline turn that covalently cyclizes the inhibitor (Fig. 5J). The RNA recognition-end of the peptide adopts a distorted type II β-turn wherein the carbonyl oxygen of Arg-5 (ith amino acid) accepts a hydrogen bond from the backbone amide of Arg-8 (i + 3rd) (data not shown). Overall, cyclization stabilizes the antiparallel β-strand structure and positions the ith and i + 1st amino acids to interact with the major groove and UCU bulge.

Combining natural and unnatural amino acids in the cyclic peptide offers advantages to elicit desired RNA–peptide interactions. Placement of L-2,4-diaminobutyric acid (B) at position 1—as opposed to Arg-1 used in precursor peptide L-22 (70)—induces favorable salt bridges between the B1 amino group and phosphates at Gua21 and Ade22 (Fig. 5J). This pairing serves to anchor the peptide in the major groove and promotes electrostatic binding by other basic groups introduced to recognize both bulge and major-groove features. For example, the guanidinium groups of Arg-3 and Arg-5 interact with Gua26 and Gua28, and the Lys-6 Nζ group hydrogen bonds to the carbonyl oxygen of Uri125.

In some respects, the determinants of TAR molecular recognition by JB181 are comparable with naturally occurring modes of TAR recognition by the Tat ARM domains from BIV and HIV. JB181 buries 920 Å² or 55% of its solvent-accessible surface area in the RNA–inhibitor interface. This level of sequestration is comparable with Tat binding to BIV or HIV-2 TAR (~1200 Å²). Recognition of BIV TAR Gua9 and Gua11 by Arg-77 and Arg-73 of BIV Tat are analogous to JB181’s use of Arg-5 to recognize Gua28 because both sets of interactions involve favorable co-planar positioning of a guanidinium group to donate two hydrogen bonds to O6 and N7 of the base Hoogsteen edge (Fig. 5, D and H). HIV-1 Tat similarly employs a single imino group of Arg-52 and Arg-49 to recognize the Hoogsteen edges of Gua26 and Gua28 within HIV-2 TAR, akin to JB181’s use of Arg-3 and Lys-6 to recognize O6 and O4 of Gua26 and Uri25—albeit JB181 does not utilize the ASM. Although JB181 binds TAR with 100–1000-fold greater affinity than HIV and BIV Tat, it uses fewer specific interactions to recognize TAR (Fig. 5B). Whereas BIV and HIV-1 Tat peptides use every arginine of the ARM sequence for RNA binding, JB181 utilizes half of its complement. This attribute may be indicative of a greater role for JB181’s charged residues in general electrostatic recognition of the RNA. Notably, the S value of 0.59 for the TAR–JB181 complex agrees well with that of a similar antiviral cyclic peptide, L-22, whose shape complementarity score is 0.60 in the context of the TAR complex (70, 77, 88). L-22 likewise relies on electrostatic features to bind TAR RNA (62).

Model of the SEC–core complex bound to HIV-1 Tat(1–60)

In addition to the new structures showing TAR recognition, a recent co-crystal structure of the TAR–Tat–SEC–core complex was determined recently to 3.5 Å resolution. This exhilarating complex comprises Tat(1–48)–CycT1–AFF4–CDK9 and the apical loop attached to stem s1b of HIV-1 TAR (74). Although the Tat ARM domain is absent in electron density maps, the complex shows how the transactivation domain of Tat is interwoven into CycT1 and that both viral and host proteins contact the TAR apical loop (Fig. 6A), burying 350 Å² of the RNA’s solvent-accessible surface. To visualize a more complete model of HIV-1 Tat binding to TAR—including the Tat ARM domain and TAR bulged loop—we superimposed the recent TAR–SEC–core complex (74) upon the HIV-2 TAR–Tat(44–60) structure (i.e. Fig. 5E) (76) based on the position of the common RNA elements. This model provides an integrated view of HIV-1 TAR–Tat–SEC–core binding (Fig. 6A). With this new perspective, the mode of TAR recognition by Tat encompasses the following: (i) the Tat ARM domain and (ii) the Tat transactivation domain—including contributions from CycT1. This bipartite mode of Tat recognition buries an estimated 1550 Å² of TAR’s solvent-accessible surface. In this manner, Tat binding facilitates SEC–core recruitment to TAR by shifting the conformational equilibrium of the unbound RNA to the bound-state conformation (Fig. 2, A and D) (74).

Implications for drug discovery and design

Based on the model of the TAR–Tat(1–60)–SEC–core complex, we can evaluate prior studies of antiviral molecules to assess progress and future challenges. As shown above, a comparison of the newly determined HIV-1 TAR–JB181 complex to the recent HIV-2 TAR–Tat complex reveals that the former cyclic peptide inhibitor induces conformational changes in the RNA apical loop and UCU bulge (i.e. a canonical Cyt30–Gua34 pair and the hallmark base triple) that are comparable with folding features elicited by HIV-1 Tat (Fig. 5, I and J versus E and F) (63). Despite this similarity and the remarkably high affinity of JB181 for TAR, the antiviral properties of this inhibitor are limited. Indeed, both JB181 (Kd 28.4 ± 4 ps) and its precursor L-22 (Kd ~30 nm) (63, 70, 88) can reduce HIV-1 replication and viral spreading in cell culture, but only to a similar extent (Kd ~40 μm) (63, 88). Hence, the 1000-fold tighter binding to TAR by JB181 compared with L-22 does not appear sufficient to overcome SEC–Tat binding to TAR. Poor intracellular delivery or low stability could account for the unexpectedly low Kd of JB181. However, a biochemical explanation from the authors of the JB181 study offers another possibility. Specifically, they noted that JB181 efficiently displaces peptide mimics of the Tat ARM domain from TAR, but the inhibitor fails to block recruitment of the SEC–core complex onto TAR (63).

A key implication from this work is that cyclic peptides—or lab-evolved proteins—designed to target the HIV-1 UCU-bulge and s1b may not be sufficient to function as potent antivirals (63). Instead, effective Tat inhibition likely requires overcoming its extensive ARM contacts to the entire TAR major groove, as well as apical loop recognition by both the Tat transactivation domain and host CycT1 (Fig. 6B). This observation also has implications for small-molecule inhibitors (63). As we have seen, the preponderance of structurally characterized drug-like molecules bind TAR at the s1a–s1b interface (Fig. 3B). Future drug-design efforts should consider approaches...
that target TAR’s apical loop interaction with Tat and CycT1 (Fig. 6, A and B). Alternatively, multivalent molecules can be envisioned that target TAR; such molecules would simultaneously displace the Tat ARM domain from TAR’s major groove, while blocking the Tat transactivation domain and CycT1 from binding the TAR apical loop. Another possibility is to create steric blocks at distal sites of the SEC that do not directly interact with TAR. For example, we superposed TAR from the TAR–TBP6.7 complex (Fig. 5, G and H) upon the TAR–Tat(1–60)–SEC–core model (Fig. 6C). The results not only reveal competition between the lab-evolved H9252–H9253 loop and Tat ARM but also a steric block arising from the lab-evolved protein where its H9252–H92511 and H92512–H92524 loops clash with loop 112–124 of CycT1. This observation provides a possible explanation for why TBP6.7 hinders TAR–Tat-dependent transcription at a concentration of 0.2 μM TBP6.7 (85), whereas the minimal stapled β2–β3-loop peptide from TBP6.7—which is missing the β1–α1 and α2–β4 loops—requires 20 μM concentrations (77).

In closing, the field is only beginning to understand the basis of molecular recognition of HIV-1 TAR by cognate host and virus proteins. New investigations and innovative approaches are needed to make progress on this complex and multifaceted problem. Our experience is that lab-evolved proteins offer a flexible strategy to cultivate the development of peptide inhibitors directed at specific regions of TAR (85). Proof–of–concept has been demonstrated by the recent TAR–TBP6.7 co-crystal structure (Fig. 5, G and H), which has been reduced to a small restrained peptide (77). By producing a series of peptide inhibitors that target multiple discrete TAR sites and reducing these to small molecules (e.g. employing HIV protease methods (29)), it may be feasible to create a multivalent drug by covalently tethering the disparate compounds together. A related approach was used to target nucleotide repeat tran-
scripts that give rise to myotonic dystrophies (140). The use of lab-evolved proteins to target RNA is broadly applicable to the development of new reagents and drugs that target a variety of functional ncRNAs.

Acknowledgments—We thank Ivan Belashov, Jermaine Jenkins, Chapin Cavender, and David Mathews for helpful discussions. We thank Matthew Shortridge and Gabriele Varani for providing unreleased coordinates. We acknowledge Martin Spiering and Catherine Goodman for helpful comments and editing. We are grateful for figure preparation assistance from Luciana Giono.

References
1. Cech, T. R., and Steitz, J. A. (2014) The noncoding RNA revolution—trashing old rules to forge new ones. Cell 157, 77–94 CrossRef Medline
2. Pennisi, E. (2012) Genomics. ENCODE project writes eulogy for junk DNA. Science 337, 1159–1161 CrossRef Medline
3. Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G. K., Khatun, J., Williams, B. A., Zaleski, C., et al. (2012) Landscape of transcription in human cells. Nature 489, 101–108 CrossRef Medline
4. Matsui, M., and Corey, D. R. (2017) Non-coding RNAs as drug targets. Nat. Rev. Drug Discov. 16, 167–170 CrossRef Medline
5. Li, Z., and Rana, T. M. (2014) Therapeutic targeting of microRNAs: cur-}

with myotonic dystrophy. Proc. Natl. Acad. Sci. U.S.A. 106, 18551–18556 CrossRef Medline
19. Blount, K. F., Megyola, C., Plummer, M., Osterman, D., O’Connell, T., Aristoff, P., Quinn, C., Chrusciel, R. A., Poel, T. J., Schostarek, H. I., Stewart, C. A., Walker, D. P., Wuts, P. G., and Breaker, R. R. (2015) Novel riboswitch-binding flavin analog that protects mice against Clostridium difficile infection without inhibiting cecal flora. Antimicrob. Agents Chemother. 59, 5736–5746 CrossRef Medline
20. Blount, K. F., Wang, J. X., Lim, J., Sudarsan, N., and Breaker, R. R. (2007) Antibacterial lysine analogs that target lysine riboswitches. Nat. Chem. Biol. 3, 44–49 CrossRef Medline
21. Howe, J. A., Wang, H., Fischmann, T. O., Balibar, C. J., Xiao, L., Galgoci,A. M., Malinverni, J. C., Mayhood, T., Villafana, A., Nahvi, A., Murgolo, N., Barbieri, C. M., Mann, P. A., Carr, D., Xia, E., et al. (2015) Selective small-molecule inhibition of an RNA structural element. Nature 526, 672–677 CrossRef Medline
22. Howe, J. A., Xiao, L., Fischmann, T. O., Wang, H., Tang, H., Villafana, A., Zhang, R., Barbieri, C. M., and Roemer, T. (2016) Atomic resolution mechanistic studies of ribocil: a highly selective unnatural ligand mimic of the e. coli FNM riboswitch. RNA Biol. 13, 946–954 CrossRef Medline
23. Lee, F. R., Blount, K. F., and Breaker, R. R. (2009) Roseoflavins is a natural antibacterial compound that binds to FMIN riboswitches and regulates gene expression. RNA Biol. 6, 187–194 CrossRef Medline
24. Sudarsan, N., Cohen-Chalamish, S., Nakamura, S., Eminsson, G. M., and Breaker, R. R. (2005) Thiamine pyrophosphate riboswitches are targets for the antimicrobial compound pyrithiamine. Chem. Biol. 12, 1325–1335 CrossRef Medline
25. Vicens, Q., Mondragón, E., Reyes, F. E., Coish, P., Aristoff, P., Berman, J., Kaur, H., Kells, K. W., Wickens, P., Wilson, J., Gadwood, R. C., Schostarek, H. J., Suto, R. K., Blount, K. F., and Batye, R. T. (2018) Structure–activity relationship of flavin analogues that target the flavin mononucleotide ribo-

26. Wang, H., Mann, P. A., Xiao, L., Gill, C., Galgoci, A. M., Howe, J. A., Villafana, A., Barbieri, C. M., Malinverni, J. C., Sher, X., Mayhood, T., McCurry, M. D., Murgolo, N., Flattery, A., Mack, M., and Roemer, T. (2017) Dual-targeting small-molecule inhibitors of the Staphylococcus aureus F MN riboswitch disrupt riboflavin homeostasis in an infectious setting. Cell. Chem. Biol. 24, 576–588.e6 CrossRef Medline
27. Palacin, J., Swalley, S. E., Song, C., Cheung, A. K., Shu, L., Zhang, X., Van Hoosier, M., Shin, Y., Chin, D. N., Keller, C. G., Beibel, M., Renaud, N. A., Smith, T. M., Salcius, M., Shi, X., et al. (2015) SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice. Nat. Chem. Biol. 11, 511–517 CrossRef Medline
28. Dibrov, S. M., Ding, K., Brunn, N. D., Parker, M. A., Bergdahl, B. M., Wyles, D. L., and Hermann, T. (2012) Structure of a hepatitis C virus RNA domain in complex with a translation inhibitor reveals a binding mode reminiscent of riboswitches. Proc. Natl. Acad. Sci. U.S.A. 109, 5223–5228 CrossRef Medline
29. Ghosh, A. K., Anderson, D. D., Weber, I. T., and Mitsuya, H. (2012) Enhancing protein backbone binding—a fruitful concept for combating drug-resistant HIV. Angew. Chem. Int. Ed. Engl. 51, 1778–1802 CrossRef Medline
30. Dieffenbach, C. W., and Fauci, A. S. (2011) Thirty years of HIV and AIDS: future challenges and opportunities. Ann. Intern. Med. 154, 766–771 CrossRef Medline
31. Dahabieh, M. S., Battivelli, E., and Verdin, E. (2015) Understanding HIV latency: the road to an HIV cure. Annu. Rev. Med. 66, 407–421 CrossRef Medline
32. Fener, I., Suhartono, M., Breitung, S., Jonker, H. R., Hennig, M., Wöhner, J., Göbel, M., and Schwab, H. (2009) Structures of HIV TAR RNA—ligand complexes reveal higher binding stoichiometries. Chembi- ochim. Acta. 10, 1490–1494 CrossRef Medline
33. Wei, P., Garber, M. E., Fang, S. M., Fischer, W. H., and Jones, K. A. (1998) A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. Cell 92, 451–462 CrossRef Medline
34. Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heaphy, S., Karn, J., Lowe, A. D., Singh, M., and Skinner, M. A. (1990) HIV-1 tat protein
stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. *EMBO J.* 9, 4145–4153 CrossRef MedLine

35. Dingwall, C., Ernberg, L., Gait, M. J., Green, S. M., Heaphy, S., Karn, J., Lowe, A. D., Singh, M., Skinner, M. A., and Valerio, R. (1989) Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 86, 6925–6929 CrossRef MedLine

36. Jakobovits, A., Smith, D. H., Jakobovits, E. B., and Capon, D. J. (1988) A discrete element 3′ of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activation by an HIV trans activator. *Mol. Cell. Biol.* 8, 2555–2561 CrossRef MedLine

37. Weeks, K. M., Ampe, C., Schultz, S. C., Steitz, T. A., and Crothers, D. M. (1990) Fragments of the HIV-1 Tat protein specifically bind TAR RNA. *Science* 249, 1281–1285 CrossRef MedLine

38. Cordingley, M. G., LaFemina, R. L., Callahan, P. L., Condra, J. H., Sardana, V. V., Graham, D. J., Nguyen, T. M., LeGro, K., Godlib, L., Schlabach, A. J., and et al. (1990) Sequence-specific interaction of Tat protein and Tat peptides with the transactivation-responsive sequence element of human immunodeficiency virus type 1 in vitro. *Proc. Natl. Acad. Sci. U.S.A.* 87, 8985–8989 CrossRef MedLine

39. Delling, U., Reid, L. S., Barnett, R. W., Ma, M. Y., Climie, S., Sumner-Smith, M., and Sonenberg, N. (1992) Conserved nucleotides in the TAR RNA stem of human immunodeficiency virus type 1 are critical for Tat binding and trans activation: model for TAR RNA tertiary structure. *J. Virol.* 66, 3018–3025 MedLine

40. Roy, S., Delling, U., Chen, C. H., Rosen, C. A., and Sonenberg, N. (1990) A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation. *Genes Dev.* 4, 1365–1373 CrossRef MedLine

41. Barboric, M., Yik, J. H., Czudnochowski, N., Yang, Z., Chen, R., Contreiras, X., Geyer, M., Matija Peterlin, B., and Zhou, Q. (2007) Tat competes with HEXIM1 to increase the active pool of P-TEFb for HIV-1 transcription. *Nucleic Acids Res.* 35, 2003–2012 CrossRef MedLine

42. D’Orso, I., Jang, G. M., Pastuszak, A. W., Faust, T. B., Quezada, E., Booth, D. S., and Frankel, A. D. (2012) Transition step during assembly of HIV Tat-P-TEFb transcription complexes and transfer to TAR RNA. *Mol. Cell. Biol.* 32, 4780–4793 CrossRef MedLine

43. Gu, J., Babayeva, N. D., Suwa, Y., Baranovskiy, A. G., Price, D. H., and Tahirov, T. H. (2010) Crystal structure of HIV-1 Tat complexed with human P-TEFb. *Nature* 465, 747–751 CrossRef MedLine

51. Tahirov, T. H., Babayeva, N. D., Varzavand, K., Cooper, J. J., Sedore, S. C., and Price, D. H. (2010) Crystal structure of HIV-1 Tat complexed with human P-TEFb. *Nature* 465, 747–751 CrossRef MedLine

52. Gu, J., Babayeva, N. D., Suwa, Y., Baranovskiy, A. G., Price, D. H., and Tahirov, T. H. (2014) Crystal structure of HIV-1 Tat complexed with human P-TEFb and AFF4. *Cell Cycle* 13, 1788–1797 CrossRef MedLine

53. Mousseau, G., Clementz, M. A., Bakeman, W. N., Nagarsheth, N., Cameron, M., Shi, J., Baran, P., Fromentin, R., Chomont, N., and Valente, S. T. (2012) An analog of the natural steroidal alkaloid cortistatin A potently suppresses Tat-dependent HIV transcription. *Cell Host Microbe* 12, 97–108 CrossRef MedLine

54. Wan, Z., and Chen, X. (2014) Triptolide inhibits human immunodeficiency virus type 1 replication by promoting proteasomal degradation of Tat protein. *Retrovirology* 11, 88 CrossRef MedLine

55. Narayan, V., Ravindra, K. C., Chiaro, C., Cary, D., Aggarwal, B. B., Henderson, A. J., and Prabhu, K. S. (2011) Celastrol inhibits Tat-mediated human immunodeficiency virus (HIV) transcription and replication. *J. Mol. Biol.* 410, 972–983 CrossRef MedLine

56. Mousseau, G., and Valente, S. (2012) Strategies to block HIV transcription: focus on small molecule Tat inhibitors. *Biolo* 1, 668–697 CrossRef MedLine

57. Kessing, C. F., Nixon, C. C., Li, C., Tsai, P., Takata, H., Mousseau, G., Ho, P. T., Honeycutt, J. B., Fallahi, M., Trautmann, L., Garcia, J. V., and Valente, S. T. (2017) *In vivo* suppression of HIV rebound by Didehydro-Cortistatin A, a “Block-and-Lock” strategy for HIV-1 treatment. *Cell Rep.* 21, 600–611 CrossRef MedLine

58. Mousseau, G., Mediouni, S., and Valente, S. T. (2015) Targeting HIV transcription: the quest for a functional cure. *Curr. Top. Microbiol. Immunol.* 389, 121–145 CrossRef MedLine

59. Klase, Z., Kale, P., Winograd, R., Gupta, M. V., Heydarian, M., Berro, R., McCaffrey, T., and Kashanchi, F. (2007) HIV-1 TAR element is processed by Dicer to yield a viral micro-RNA involved in chromatin remodeling of the viral LTR. *BMC Mol. Biol.* 8, 63 CrossRef MedLine

60. Klase, Z., Winograd, R., Davis, J., Carpio, L., Hildreth, R., Heydarian, M., Fu, S., McCaffrey, T., Meiri, E., Ayash-Rashkovsky, M., Gilad, S., Bentwich, Z., and Kashanchi, F. (2009) HIV-1 TAR miRNA protects against apoptosis by altering cellular gene expression. *Retrovirology* 6, 18 CrossRef MedLine

61. Ouellet, D. L., Plante, I., Landry, P., Barat, C., Janelle, M. E., Flaman, L., Tremblay, M. J., and Provost, P. (2008) Identification of functional microRNAs released through asymmetrical processing of HIV-1 TAR element. *Nucleic Acids Res.* 36, 248–256 CrossRef MedLine

62. Davidson, A., Patora-Komisarska, K., Robinson, J. A., and Varani, G. (2011) Essential structural requirements for specific recognition of HIV TAR RNA by peptide mimetics of Tat protein. *Nucleic Acids Res.* 39, 248–256 CrossRef MedLine

63. Shortridge, M. D., Wille, P. T., Jones, A. N., Davidson, A., Bogdanovic, J., Arts, E., Karn, J., Robinson, J. A., and Varani, G. (2018) An ultra-high affinity ligand of HIV-1 TAR reveals the RNA structure recognized by P-TEFb. *Nucleic Acids Res.* 47, 1523–1531 CrossRef MedLine

64. Lind, K. E., Du, Z., Fujinaga, K., Peterlin, B. M., and James, T. L. (2002) Structure-based computational database screening, *in vitro* assay, and NMR assessment of compounds that target TAR RNA. *Chem. Biol.* 9, 185–193 CrossRef MedLine

65. Dethoff, E. A., Hansen, A. L., Musseman, C., Watt, E. D., Andricioaei, I., and Al-Hashimi, H. M. (2008) Characterizing complex dynamics in the transcription response element apical loop and motion correlations with the bulge by NMR, molecular dynamics, and mutagenesis. *Biochem.* J.*95*, 3906–3915 CrossRef MedLine

66. Dethoff, E. A., Petzold, K., Chugh, J., Casiano-Negroni, A., and Al-Hashimi, H. M. (2012) Visualizing transient low-populated structures of RNA. *Nature* 491, 724–728 CrossRef MedLine

67. Borkar, A. N., Bardaro, M. F., Jr., Camilloni, C., Aprile, F. A., Varani, G., and Vendruscolo, M. (2016) Structure of a low-population binding intermediate in protein–RNA recognition. *Proc. Natl. Acad. Sci. U.S.A.* 113, 7171–7176 CrossRef MedLine

68. Schlueze-Gahmen, U., Echeverria, I., Stepanovac, G., Bai, Y., Lu, H., Schneideman-Duhovsky, D., Doudna, J. A., Zhou, Q., Sali, A., and Hurley, J. H.
(2016) Insights into HIV-1 proviral transcription from integrative structure and dynamics of the Tat:AFF4:T-TEFB:TAR complex. Elife 5, e15910 CrossRef Medline

Brodsky, A. S., and Williamson, J. R. (1997) Solution structure of the HIV-2 TAR–argininamide complex. J. Mol. Biol. 267, 624–639 CrossRef Medline

Davidson, A., Leeper, T. C., Athanassiou, Z., Patora-Komisarska, K., Karn, J., Robinson, J. A., and Varani, G. (2009) Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. Proc. Natl. Acad. Sci. U.S.A. 106, 11931–11936 CrossRef Medline

Merriman, D. K., Xue, Y., Yang, S., Kimsey, I. J., Shakya, A., Clay, M., and Principles of Nucleic Acid Structure

Leeper, T. C., Athanassiou, Z., Dias, R. L., Robinson, J. A., and Varani, G. (2004) Inhibition of HIV-1 reverse transcription and gene expression by a cyclic peptide that binds the Tat-transactivating response element (TAR) RNA. PLoS Pathog 7, e1000238 CrossRef Medline

Richter, S., Parlolin, C., Gatto, B., Del Vecchio, C., Brocca-Cofano, E., Fravolini, A., Palù, G., and Palumbo, M. (2004) Inhibition of human immunodeficiency virus type 1 tat-transactivation-responsive region interaction by an antiviral quinolone derivative. Antimicrob. Agents Chemother. 48, 1895–1899 CrossRef Medline

Mei, H. Y., Mack, D. P., Galan, A. A., Halim, N. S., Heldsinger, A., Loo, J. A., Moreland, D. W., Sannes-Rowdy, K. A., Sharmeen, L., Truong, H. N., and Caznik, A. W. (1997) Discovery of selective, small-molecule inhibitors of RNA complexes–I. The Tat protein/TAR RNA complexes required for HIV-1 transcription. Bioorg. Med. Chem. 5, 1173–1184 CrossRef Medline

Abulwerdi, F. A., Shartridge, M. D., Sztuba-Solinska, J., Wilson, R., Le Grice, S. F., Varani, G., and Schneekloth, J. S. Jr. (2016) Development of small molecules with a noncanonical binding mode to HIV-1 trans activation response (TAR) RNA. J. Med. Chem. 59, 11148–11160 CrossRef Medline

Connelly, C. M., Moon, M. H., and Schneekloth, J. S., Jr. (2016) The emerging role of RNA as a therapeutic target for small molecules. Cell Chem. Biol. 23, 1077–1090 CrossRef Medline

Colman, P. M. (1993) Shape complementarity at protein/protein interfaces. Perspect. Med. Chem. 3, 21–37 Medline

Aboul-ela, F., Karn, J., and Varani, G. (1995) The structure of the human immunodeficiency virus type-1 TAR trans-activation response (TAR) RNA. J. Mol. Biol. 240, 984–996 CrossRef Medline

D’Auria, B., Monti, A., and Brodie, I. M. (2007) Inhibition of human immunodeficiency virus type 1 Tat–TAR RNA binding by cyclic peptide mimetics. Biochemistry 46, 6458–6464 CrossRef Medline

Wang, Z., Shah, K., and Rana, T. M. (2001) Probing Tat peptide–TAR RNA interactions by psoralen photo-cross-linking. Biochemistry 40, 8885–8895 CrossRef Medline

Lawrence, M. C., and Colman, P. M. (1993) Shape complementarity at protein/protein interfaces. J. Mol. Biol. 234, 946–950 CrossRef Medline

Aboul-ela, F., Karn, J., and Varani, G. (1995) The structure of the human immunodeficiency virus type-1 TAR RNA. J. Mol. Biol. 253, 313–332 CrossRef Medline

Churcher, M. J., Lamont, C., Hamy, F., Dingwall, C., Green, S. M., Lowe, A. D., Butler, J. G., Gait, M. J., and Karn, J. (1995) High affinity binding of TAR RNA by the human immunodeficiency virus type-1 tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region. J. Mol. Biol. 230, 90–110 CrossRef Medline

Lawrence, M. C., and Colman, P. M. (1993) Shape complementarity at protein/protein interfaces. J. Mol. Biol. 234, 946–950 CrossRef Medline

Aboul-ela, F., Karn, J., and Varani, G. (1995) The structure of the human immunodeficiency virus type-1 TAR RNA. J. Mol. Biol. 253, 313–332 CrossRef Medline

Churcher, M. J., Lamont, C., Hamy, F., Dingwall, C., Green, S. M., Lowe, A. D., Butler, J. G., Gait, M. J., and Karn, J. (1995) High affinity binding of TAR RNA by the human immunodeficiency virus type-1 tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region. J. Mol. Biol. 230, 90–110 CrossRef Medline

Murchie, A. L., Davis, B., Isel, C., Afshar, M., Drysdale, M. J., Bower, J., Potter, A. J., Starkey, I. D., Searbrick, T. M., Mirza, S., Prescott, C. D., Vaglio, P., Aboul-ela, F., and Karn, J. (2004) Structure-based drug design
targeting an inactive RNA conformation: exploiting the flexibility of HIV-1 TAR RNA. J. Mol. Biol. 336, 625–638 CrossRef Medline

107. Davis, B., Afsar, M., Varani, G., Murchie, A. I., Karlin, J., Lentzen, G., Drysdale, M., Bower, J., Potter, A. J., Starkey, I. D., Swarbrick, T., and Aboul-ela, F. (2004) Rational design of inhibitors of HIV-1 TAR RNA through the stabilisation of electrostatic “hot spots”. J. Mol. Biol. 336, 343–356 CrossRef Medline

108. Mei, H. Y., Cui, M., Heldsinger, A., Lemrow, S. M., Loo, J. A., Sanses-Lowery, K. A., Sharmeen, L., and Czarnik, A. W. (1998) Inhibitors of protein–RNA complexation that target the RNA: specific recognition of human immunodeficiency virus type 1 TAR RNA by small organic molecules. Biochemistry 37, 14204–14212 CrossRef Medline

109. Du, Z., Lind, K. E., and James, T. L. (2002) Structure of TAR RNA complexed with a Tat–TAR interaction nanomolar inhibitor that was identified by computational screening. Chem. Biol. 9, 707–712 CrossRef Medline

110. Parsons, J., Castaldi, M. P., Dutta, S., Dibrov, S. M., Wyles, D. L., and Cer, R. Z., Mudunuri, U., Stephens, R., and Lebeda, F. J. (2009) IC50-to-IC90 conversion of drug candidates. JBC REVIEWS: Molecular recognition of HIV TAR RNA

111. Huber, W. (2005) A new strategy for improved secondary screening and lead optimization using high-resolution SPR characterization of compound-target interactions. J. Mol. Recognit. 18, 273–281 CrossRef Medline

112. Pollard, T. D. (2010) A guide to simple and informative binding assays. Mol. Biol. Cell 21, 4061–4067 CrossRef Medline

113. Cer, R. Z., Mudunuri, U., Stephens, R., and Lebeda, F. J. (2009) IC50-to-Ki: a web-based tool for converting IC50 to Ki values for inhibitors of protein–nucleic acid interactions. J. Mol. Recognit. 23, 273–281 CrossRef Medline

114. Davis, B., Afshar, M., Varani, G., Murchie, A. I., Karn, J., Lentzen, G., Drysdale, M., Bower, J., Potter, A. J., Starkey, I. D., Swarbrick, T., and Aboul-ela, F. (2004) Rational design of inhibitors of HIV-1 TAR RNA through the stabilisation of electrostatic “hot spots”. J. Mol. Biol. 336, 343–356 CrossRef Medline

115. Ippolito, J. A., and Steitz, T. A. (1998) A 1.3-A resolution crystal structure of the HIV-1 trans-activation response region RNA stem reveals a metal ion-dependent bulge conformation. Proc. Natl. Acad. Sci. U.S.A. 95, 9819–9824 CrossRef Medline

116. Hellman, L. M., and Fried, M. G. (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. Nat. Protoc. 2, 1849–1861 CrossRef Medline

117. Freire, E. (2009) A thermodynamic approach to the affinity optimization of drug candidates. Chem. Biol. Drug Des. 74, 468 – 472 CrossRef Medline

118. Huber, W. (2005) A new strategy for improved secondary screening and lead optimization using high-resolution SPR characterization of compound-target interactions. J. Mol. Recognit. 18, 273–281 CrossRef Medline

119. Seth, P. P., Miyaji, A., Jefferson, E. A., Sanses-Lowery, K. A., Osgood, S. A., Propp, S. S., Ranken, R., Massire, C., Sampath, R., Ecker, D. J., Swayne, E. E., and Griffey, R. H. (2005) SAR by MS: discovery of a new class of RNA-binding small molecules for the hepatitis C virus: internal ribosome entry site IIA subdomain. J. Med. Chem. 48, 7099–7102 CrossRef Medline

120. Davis, D. R., and Seth, P. P. (2011) Therapeutic targeting of HCV internal ribosom al entry site RNA. Antivir. Chem. Chemother. 21, 117–128 CrossRef Medline

121. Paulsen, R. B., Seth, P. P., Swayne, E. E., Griffey, R. H., Skalicky, J. I., Cheatham, T. E., 3rd., and Davis, D. R. (2010) Inhibitor-induced structural change in the HCV IRES domain IIA RNA. Proc. Natl. Acad. Sci. U.S.A. 107, 7263–7268 CrossRef Medline

122. McCown, P. J., Corbino, K. A., Stav, S., Sherlock, M. E., and Breaker, R. R. (2017) Riboswitch diversity and distribution. RNA 23, 995–1011 CrossRef Medline

123. Hiller, D. A., and Strobel, S. A. (2011) The chemical versatility of RNA. Philos. Trans. R. Soc. Lond. B Biol. Sci. 366, 2929–2935 CrossRef Medline

124. Breaker, R. R. (2018) Riboswitches and translation control. Cold Spring Harb. Perspect. Biol. 10, a032797 CrossRef Medline

125. Sherwood, A. V., and Henkin, T. M. (2016) Riboswitch-mediated gene regulation: novel RNA architectures dictate gene expression responses. Annu. Rev. Microbiol. 70, 361–374 CrossRef Medline

126. Serganov, A., and Nudler, E. (2013) A decade of riboswitches. Cell 152, 17–24 CrossRef Medline

127. Otani, S., Takeda, M., Nakano, M., Kasa, S., and Miura, R. (1974) Letter: Roseoflavin, a new antimicrobial pigment from Streptomyces. J. Antibiot. 27, 86–87 CrossRef Medline

128. Serganov, A., Huang, L., and Patel, D. J. (2009) Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. Nature 458, 233–237 CrossRef Medline

129. Calnan, B. J., Biancalana, S., Hudson, D., and Frankel, A. D. (1991) Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA–protein recognition. Genes Dev. 5, 201–210 CrossRef Medline

130. Creighton, T. E. (1993) Proteins: Structures and Molecular Properties, 2nd Ed., pp. 1–507, W.H. Freeman and Co., New York

131. Anand, K., Schulte, A., Vogel-Bachmayr, K., Scheffzek, K., and Geyer, M. (2008) Structural insights into the cyclin T1–Tat–TAR RNA transcription activation complex from EIAV. Nat. Struct. Mol. Biol 15, 1287–1292 CrossRef Medline

132. Lawrence, C. F., Margetts, M. B., Menting, J. G., Smith, N. A., Smith, B. J., Ward, C. W., and Lawrence, M. C. (2016) Insulin mimetic peptide disrupts the primary binding site of the insulin receptor. J. Biol. Chem. 291, 15473–15481 CrossRef Medline

133. Athanassiou, Z., Dias, R. L., Moeleh, K., Dobson, N., Varani, G., and Robinson, J. A. (2004) Structural mimicry of retroviral Tat proteins by constrained, β-hairpin peptidomimetics: ligands with high affinity and selectivity for viral TAR RNA regulatory elements. J. Am. Chem. Soc. 126, 6906–6913 CrossRef Medline

134. Athanassiou, Z., Patora, K., Dias, R. L., Moeleh, K., Robinson, J. A., and Varani, G. (2007) Structure-guided peptidomimetic design leads to nanomolar β-hairpin inhibitors of the Tat–TAR interaction of bovine immunodeficiency virus. Biochemistry 46, 741–751 CrossRef Medline

135. Schulze-Gahmen, U., and Hurley, J. H. (2018) Structural mechanism for HIV-1 TAR loop recognition by Tat and the super-elongation complex. Proc. Natl. Acad. Sci. U.S.A. 115, 12973–12978 CrossRef Medline

136. Lee, M. C. (2016) Insulin mimetic peptide disassociation mechanisms. Wiley Interdiscip. Rev. Comput. Mol. Sci. 6, 233–237 CrossRef Medline

137. Peterlin, B. M., Brogie, J. E., and Price, D. H. (2012) 7SK snRNA: a non-coding RNA that regulates genes. Wiley Interdiscip. Rev. RNA 3, W441–W445 CrossRef Medline

138. Lawrence, M. C. (2016) Insulin mimetic peptide disassociation mechanisms. Wiley Interdiscip. Rev. Comput. Mol. Sci. 6, 233–237 CrossRef Medline

139. Salkhorst, E. J., and Lifson, J. D. (2014) AFF4 binding to its TAR RNA: a new mechanism for targeting an inactive RNA conformation: exploiting the flexibility of HIV-1 TAR RNA. J. Mol. Biol. 336, 625–638 CrossRef Medline

140. Lee, M. M., Childs-Disney, J. L., Pushechnikov, A., French, J. M., Sobczak, T. C., Czarnik, A. W., and Breaker, R. R. (1998) Inhibitors of protein–RNA complexation that target the RNA: specific recognition of HIV-1 TAR RNA. J. Am. Chem. Soc. 120, 9441–9450 CrossRef Medline

141. McCown, P. J., Corbino, K. A., Stav, S., Sherlock, M. E., and Breaker, R. R. (2017) Riboswitch diversity and distribution. RNA 23, 995–1011 CrossRef Medline

142. Smith, E., Lin, C., and Shilatifard, A. (2011) The super-elongation complex (SEC) and MLL in development and disease. Genes Dev. 25, 661–672 CrossRef Medline
145. Zhu, Y., Pe’ery, T., Peng, J., Ramanathan, Y., Marshall, N., Marshall, T., Amendt, B., Mathews, M. B., and Price, D. H. (1997) Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro. Genes Dev. 11, 2622–2632 CrossRef Medline

146. Mancebo, H. S., Lee, G., Flygare, J., Tomassini, J., Luu, P., Zhu, Y., Peng, J., Blau, C., Hazuda, D., Price, D., and Flores, O. (1997) P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. Genes Dev. 11, 2633–2644 CrossRef Medline

147. Yang, X., Gold, M. O., Tang, D. N., Lewis, D. E., Aguilar-Cordova, E., Rice, A. P., and Herrmann, C. H. (1997) TAK, an HIV Tat-associated kinase, is a member of the cyclin-dependent family of protein kinases and is induced by activation of peripheral blood lymphocytes and differentiation of promonocytic cell lines. Proc. Natl. Acad. Sci. U.S.A. 94, 12331–12336 CrossRef Medline

148. Kulinski, T., Olejniczak, M., Huthoff, H., Bielecki, L., Pachulska-Wieczorek, K., Das, A. T., Berkhout, B., and Adamiak, R. W. (2003) The apical loop of the HIV-1 TAR RNA hairpin is stabilized by a cross-loop base pair. J. Biol. Chem. 278, 38892–38901 CrossRef Medline

149. Richter, S., Cao, H., and Rana, T. M. (2002) Specific HIV-1 TAR RNA loop sequence and functional groups are required for human cyclin T1-Tat–TAR ternary complex formation. Biochemistry 41, 6391–6397 CrossRef Medline

150. Foley, B., Leitner, T., Apetrei, C., Hahn, B., Mizrahi, I., Mullins, J., Rambaut, A., Wolinsky, S., and Korber, B. (2018) HIV Sequence Compendium, pp. LA-UR-18–25673, Los Alamos National Laboratory, Los Alamos, NM

151. Chen, L., and Frankel, A. D. (1994) An RNA-binding peptide from bovine immunodeficiency virus Tat protein recognizes an unusual RNA structure. Biochemistry 33, 2708–2715 CrossRef Medline

152. Blanchet, C., Pasi, M., Zakrzewska, K., and Lavery, R. (2011) CURVES+ web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures. Nucleic Acids Res. 39, W68–W73 CrossRef Medline

153. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797 CrossRef Medline

154. Mousseau, G., and Valente, S. T. (2017) Role of host factors on the regulation of Tat-mediated HIV-1 transcription. Curr. Pharm. Des. 23, 4079–4090 CrossRef Medline

155. Calabro, V., Daugherty, M. D., and Frankel, A. D. (2005) A single intermolecular contact mediates intramolecular stabilization of both RNA and protein. Proc. Natl. Acad. Sci. U.S.A. 102, 6849–6854 CrossRef Medline

156. Oubridge, C., Ito, N., Evans, P. R., Teo, C. H., and Nagai, K. (1994) Crystal structure at 1.92 A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. Nature 372, 432–438 CrossRef Medline