ChemPubSoc Europe

DOI: 10.1002/chem.201303664

CHEMISTRY A European Journal Full Paper

Cooperative Binding

Artificial Nucleobase–Amino Acid Conjugates: A New Class of TAR RNA Binding Agents**

Jean-Patrick Joly,[a] Guillaume Mata,[a] Patrick Eldin,[b] Laurence Briant,[b] Fabien Fontaine-Vive,[a] Maria Duca,*[a] and Rachid Benhida*[a]

Abstract: The human immunodeficiency virus type-1 (HIV-1) Tat protein stimulates transcriptional elongation. Tat is involved in the transcription machinery by binding to the transactivation response region (TAR) RNA stem-loop structure, which is encoded by the 5′ leader sequence found in all HIV-1 mRNAs. Herein, we report the rational design, synthesis, and in vitro evaluation of new RNA binding agents that were conceived in order to bind strongly and selectively to the stem-loop structure of TAR RNA and, thus, inhibit the Tat/TAR interaction. We have demonstrated that the conjugation of modified nucleobases, able to interact specifically with an RNA base pair, and various amino acids allows these motifs to bind the target RNA selectively and in a cooperative manner that leads to the inhibition of viral replication in HIV-infected cells.

Introduction

In recent years, RNA has been discovered to play important roles in many cellular processes and it is now recognized as a valid target for therapeutic intervention by drug-like small molecules.[1] Furthermore, many RNAs are related to diseases, either due to mutations that affect their normal function or because they are part of pathogenic organisms that invade human cells. Hence, RNA is an important target for the design of inhibitors directed at preventing pathological processes.[2] One such target is transactivation response region (TAR) RNA, a 59-nucleotide fragment from the human immunodeficiency virus type-1 (HIV-1) genome that forms a highly conserved stem-loop structure containing an internal bulge[3] and that is found at the 5′ end of all nascent HIV-1 transcripts. HIV-1 TAR RNA has a key role in viral replication because its interaction with the Tat protein allows the formation of a complex involving cellular cofactors such as cyclin T1 and its cognate kinase CDK9, which thus stimulates efficient transcription from the retroviral promoter (LTR).[4] Therefore, in the present study, we selected HIV-1 TAR RNA as a biologically relevant model to quantify RNA–small-molecule binding, to decipher the mode of interaction, and to examine the in vitro activity.

Given the global problem of HIV-1 infection and the high frequency of drug resistance to the current HIV therapies, the discovery of antiviral drugs that exhibit novel modes of action are of the utmost importance. The majority of studies on RNA targeting have focused on oligonucleotide-based antisense strategies, in which high-binding-affinity hybridization between the oligonucleotide and the target strand serves to alter gene expression.[5] However, even if these RNA binders exhibit high binding affinity and specificity, their intrinsic instability and their poor ability to penetrate the cell membrane greatly hamper their therapeutic application.[6] Peptides have also been studied as potent ligands of hairpin RNAs on the basis of their similarity with natural RNA ligands. Although a great number of short peptides and analogues have already been reported in the literature, none of them has shown sufficient specificity in order to be used in therapy.[7] Finally, aminoglycosides constitute a particularly well-studied class of RNA-binding molecules, which are essentially pseudo-oligosaccharides containing numerous amine groups within their structure. These molecules are well known for their antibiotic properties, which are related to their capacity to interact with bacterial ribosomal RNA, a property that impairs protein biosynthesis.[8] However, because their binding mechanism mainly involves nonspecific electrostatic interactions, aminoglycosides suffer from a lack of selectivity, which is responsible for their high toxicity.

Recently, based on a rational design, we discovered a new class of modified nucleosides able to bind specifically to TAR RNA at the stem-bulge junction.[9] These nucleosides contain a modified nucleobase (**S**, Figure 1 A) that is able to initiate RNA recognition in a sequence-selective manner by forming a triplet with an AU base pair.[10] The introduction of basic

[a] J.-P. Joly, G. Mata, Dr. F. Fontaine-Vive, Dr. M. Duca, Dr. R. Benhida
Institut de Chimie de Nice UMR7272 CNRS
Université de Nice Sophia Antipolis
Parc Valrose, 06108 Nice cedex 1 (France)
Fax: (+33) 4-92076151
E-mail: duca@unice.fr
benhida@unice.fr

[b] Dr. P. Eldin, Dr. L. Briant
Centre d'Etudes d'Agents Pathogènes et Biotechnologies pour la Santé
UMR5236 CNRS
Université de Montpellier 1 and 2
1919 route de Mende, 34293 Montpellier cedex 5 (France)

[**] TAR: Transactivation response region.

Supporting information for this article (containing the synthetic procedures) is available on the WWW under http://dx.doi.org/10.1002/chem.201303664.

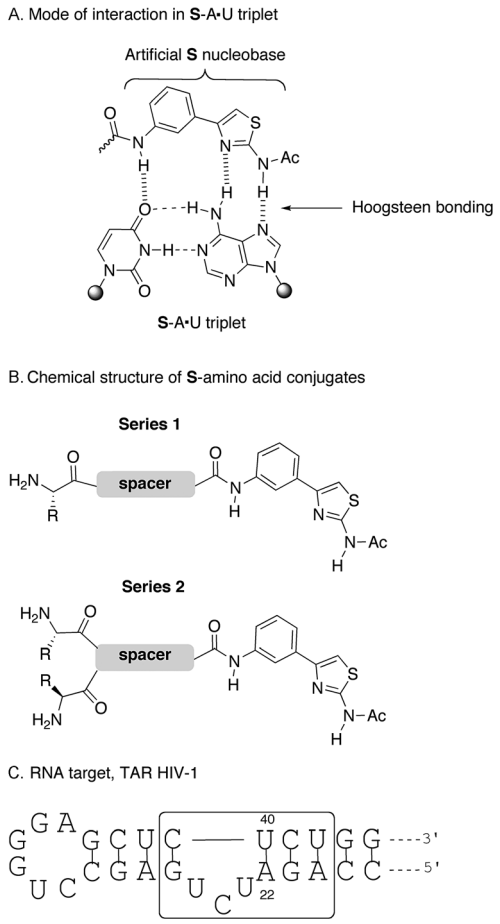


Figure 1. A) Mode of interaction in an **S**–A•U triplet. B) General structure of the newly synthesized RNA ligands (**S**–amino acid conjugates): series 1 and 2. C) The RNA target (TAR HIV-1).

amino acids, such as lysine and arginine, at the 3′ and 5′ positions of the nucleoside strengthened the interaction with targeted RNA in a cooperative way. Unfortunately, we found that this class of compounds has no activity in cellular HIV assays. A careful study clearly pointed out their rapid degradation in biological media through cleavage of the labile ester bond (see the Supporting Information).

In line with these observations and taking advantage of the previously modified nucleobase **S**, we report herein the design, synthesis, and biological evaluation of **S**–amino acid conjugates as potent, cell-permeable, and stable TAR RNA ligands. In order to increase the in vitro stability of this novel class of RNA binders and to decipher the mode of action, we modified the three-dimensional distribution of the substituents by replacing the 2′-deoxyribose scaffold with chemically stable spacers between **S** and the amino acids (series 1 and 2, Figure 1 B). We also investigated the effect of the amino acid and the linker length on the binding affinity (series 1, Figure 1 B B), as well as the effect of introducing two amino acids in the same ligand in order to further improve their affinity and to maintain their specificity (series 2, Figure 1 B).

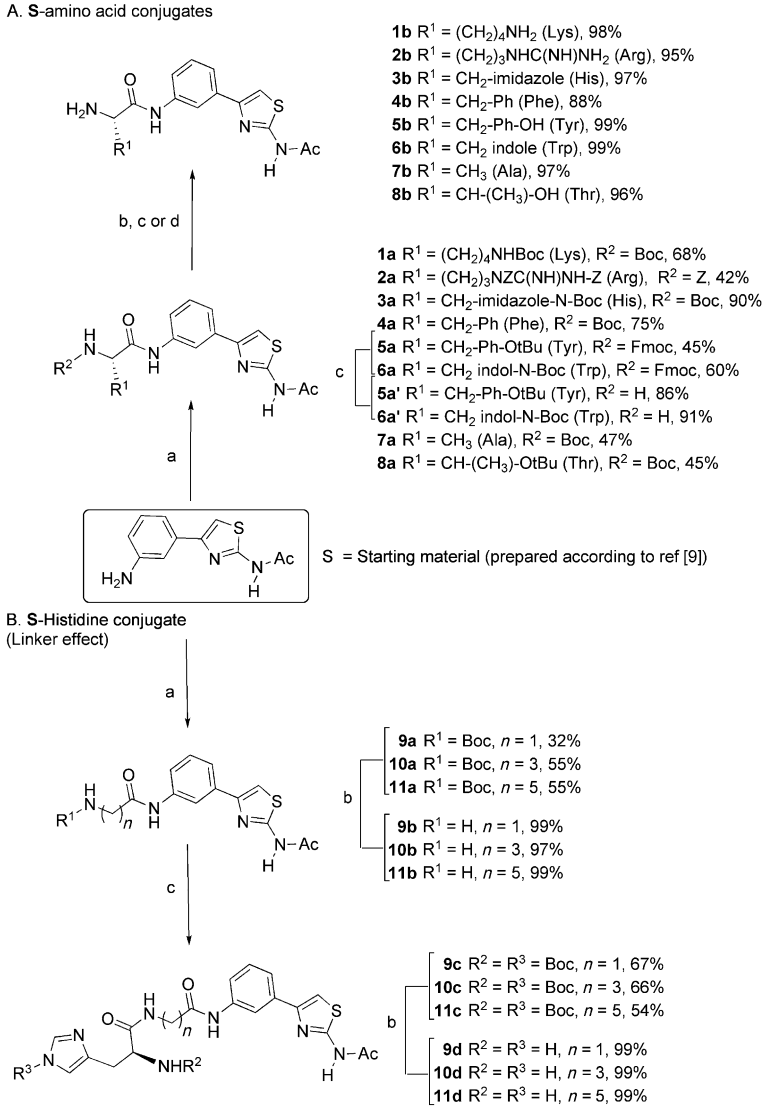
Results and Discussion

Synthesis of new modified nucleobase–amino acid conjugates

In a first approach, we envisioned the synthesis of a class of conjugates in which only one amino acid is coupled to the artificial nucleobase **S** through an amide bond (series 1, Figure 1 B). We thus synthesized the first series of conjugates (**1 b**–**8 b**) by coupling **S** and various amino acids, as depicted in Scheme 1. For this first series, we selected the basic amino acids lysine (Lys), arginine (Arg), and histidine (His), the aromatic amino acids phenylalanine (Phe), tyrosine (Tyr), and tryptophane (Trp), and the aliphatic amino acids alanine (Ala) and threonine (Thr). The basic amino acids should allow for strong electrostatic interactions with the bulge bases of the TAR sequence, as previously demonstrated for glycoconjugates containing Lys and Arg.[9] The aromatic amino acids should account for more hydrophobic interactions that could be established with the surrounding free RNA nucleobases or the base pairs at the bulge. The aliphatic amino acids should involve hydrophobic interactions inside the RNA helix. In Ala, only the α-NH_2 group is available for interaction, whereas the hydroxy group of the side chain in Thr could also participate in hydrogen bonding.

The synthesis of compounds **1 b**–**8 b** comprises only two steps. The first step involved the coupling of **S** with protected amino acids, in the presence of 2-chloro-1-methylpyridinium iodide in CH_2Cl_2, which yielded the protected compounds **1 a**–**8 a** in moderate to good yields (42–91 %, Scheme 1 A). The second step consisted of the removal of the Boc and *t*Bu protecting groups, as appropriate, in the presence of TFA in CH_2Cl_2, which led to compounds **1 b**, **3 b**, **4 b**, **7 b**, and **8 b**. An intermediate step for the removal of the Fmoc group (20 % piperidine in CH_2Cl_2) is necessary with compounds **5 a** and **6 a** and led compounds **5 a**′ and **6 a**′, respectively. The final cleavage of the *t*Bu and Boc groups, respectively, led to desired compounds **5 b** and **6 b**. The arginine derivative **2 a** was converted by hydrogenolysis (H_2, Pd/C) into compound **2 b**. All of the deprotected compounds, **1 b**–**8 b**, were obtained in high yields (88–99 %) and were fully characterized.

After the synthesis of this first set of compounds, we decided to introduce a spacer between the two binding modules (**S** and amino acids) by using aliphatic linkers of different length: glycine in compound **9 d**, 4-aminobutyric acid in compound **10 d**, and 6-aminocaproic acid in compound **11 d** (Scheme 1 B). These compounds all contain the histidine amino acid and will allow us to study the effect of linker size on the overall efficiency of the ligands in terms of affinity and selectivity. Indeed, the best linker should orient **S** and the amino acid moieties for optimal interaction with TAR RNA. The preparation of this second set of compounds is depicted in Scheme 1 B. First, **S** was coupled with Boc-protected linkers (*N*-Boc-glycine, *N*-Boc-4-aminobutyric acid, and *N*-Boc-6-aminocaproic acid) to afford compounds **9 a**–**11 a**, respectively, in moderate yields (32–55 %). The subsequent removal of the Boc group in the presence of TFA in CH_2Cl_2 led to compounds **9 b**–**11 b** in almost



Scheme 1. Synthesis of ligands with one amino acid (series 1). A) Synthesis of compounds **1b**–**8b**. Reagents: a) Boc-Lys-OH for **1a**, Z-Arg-OH for **2a**, Boc-His-OH for **3a**, Boc-Phe-OH for **4a**, Fmoc-Tyr(*t*Bu)-OH for **5a**, Fmoc-Trp(Boc)-OH for **6a**, Boc-Ala-OH for **7a**, Boc-Thr(*t*Bu)-OH for **8a**, 2-chloro-1-methylpyridinium iodide, Et_3N, CH_2Cl_2, reflux, 3 h; b) 50% TFA, CH_2Cl_2, RT, 24 h, for **1b**, **3b**, **4b**, **6b**–**8b**; c) 20% piperidine, CH_2Cl_2, RT, 24 h, for **5b** and **6b**; d) H_2, Pd/C, 24 h, for **2b**. B) Synthesis of histidine conjugates with different linker lengths (**9d**–**11d**). Reagents: a) *N*-Boc-L-glycine for compound **9a**, *N*-Boc-4-aminobutyric acid for **10a**, *N*-Boc-6-aminocaproic acid for **11a**, 2-chloro-1-methylpyridinium iodide, Et_3N, CH_2Cl_2, reflux, 3 h; b) 50% TFA, CH_2Cl_2, RT, 24 h; c) HOBt, DIC, DMAP, Et_3N, CH_2Cl_2, RT, 4 h. Boc: *tert*-butoxycarbonyl; Z: benzyloxycarbonyl; Fmoc:9-fluorenylmethoxycarbonyl; TFA: trifluoroacetic acid; HOBt: 1-hydroxy-1*H*-benzotriazole; DIC: diisopropylcarbodiimide; DMAP: 4-dimethylaminopyridine.

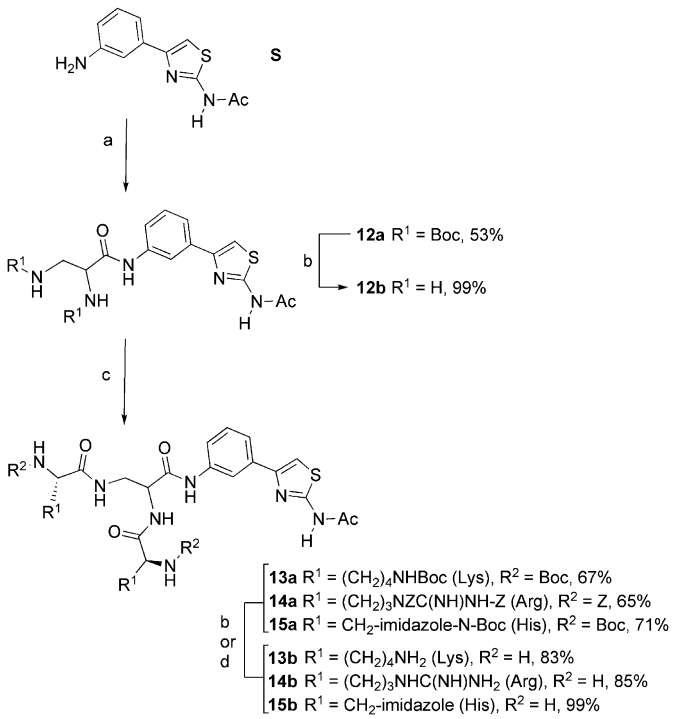
quantitative yields (97–99%). The spacer's free amino group was coupled with Boc-protected histidine in the presence of HOBt, DIC, DMAP, and Et_3N in CH_2Cl_2, which led to compounds **9c**–**11c** in 54–67% yields. The final deprotection step was achieved by using TFA and led to compounds **9d**–**11d** in quantitative yields (99%).

Finally, we decided to investigate the effect of introducing a second amino acid in the same molecule to increase the affinity. To this end, we selected a spacer bearing two reactive amines (2,3-diaminopropanoic acid), and the synthesis of this third set of compounds is illustrated in Scheme 2. For these compounds, we decided to couple the lysine, arginine, and histidine amino acids (in **13b**–**15b**, respectively), which, as mentioned above, were expected to bind strongly to RNA, thanks to their basic nature. In the first step, nucleobase **S** was coupled with *N*,*N*′-di-Boc-2,3-diaminopropanoic acid by using a similar coupling procedure to that described above to afford the desired compound **12a** in 53% yield. The second step consisted of Boc cleavage in the presence of TFA and led to compound **12b** in 99% yield. The spacer's free amino groups were then coupled with protected amino acids (Boc-Lys for **13c**, Z-Arg for **14c**, and Boc-His for **15c**) in the presence of HOBt, DIC, DMAP, and Et_3N in CH_2Cl_2. which led to compounds **13a**–**15a** in 65–71% yields. The final deprotection step was achieved by using TFA for compounds **13a** and **15a** or by catalytic hydrogenation for compound **14a** to afford high yields (83–99%) of the corresponding products **13b**–**15b**, as a mixture of two diastereomers.

Ligand binding to HIV-1 TAR RNA: affinity and specificity

After the synthesis of the first series of compounds (**1b**–**8b** and **9d**–**11d**, Scheme 1), we evaluated their ability to bind HIV-1 TAR RNA. These binding studies were performed in 20 mM 2-[4-(2-hydroxyethyl)-1-piperazinyl]ethanesulfonic acid (HEPES) buffer (pH 7.2) containing 20 mM NaCl, 140 mM KCl, and 3 mM $MgCl_2$ at 20 °C by measuring the fluorescence change of a fluorescently labeled TAR fragment (5′-Alexa^{488}) with increasing concentrations of ligands. This method, broadly used to study the interactions between RNA and small ligands,[11] allowed us to determine the dissociation constants (K_d) for all of the compounds (Table 1) after analysis of the binding isotherm diagram. Neomycin was used as a positive control. This compound belongs to the class of aminoglycoside antibiotics and, besides its ability to bind bacterial ribosomal RNA, it is also reported as a very efficient and general RNA ligand.[12]

As illustrated by the K_d values reported in Table 1, compounds **1b**, **2b** and **6b**, in which **S** was coupled with lysine, arginine, and tryptophan, respectively, exhibited increased affinity for TAR RNA (K_d = 7.5, 2.4, and 3.1 μM, respectively; Table 1, entries 2, 3, and 7) relative to that of neomycin (K_d = 17.2 μM; Table 1, entry 1). Compound **3b**, which featured a histidine residue, showed a binding affinity very close to that of neomycin (K_d = 17.5 μM; Table 1, entry 4). The aromatic amino acids Phe and Tyr seemed to be less favorable for the interaction than neomycin because compound **4b** had a moderate affinity (K_d = 30.3 μM) and **5b** was not able to bind the target (Table 1, entries 5 and 6, respectively). In similar way, aliphatic amino acids **7b** and **8b** bound TAR RNA with low or no affinity (Table 1, entries 8 and 9, respective-



Scheme 2. Synthesis of ligands featuring two amino acids (**13b**–**15b**). Reagents: *N*,*N*′-di-Boc-2,3-diaminopropanoic acid, 2-chloro-1-methylpyridinium iodide. a) Et_3N, CH_2Cl_2, reflux, 3 h; b) 50% TFA, CH_2Cl_2, RT, 24 h; c) Boc-Lys-OH for **13a**, Z-Arg-OH for **14a**, Boc-His-OH for **15a**, HOBt, DIC, DMAP, Et_3N, CH_2Cl_2, RT, 4 h; d) H_2, Pd/C, 24 h for **14b**.

Table 1. Dissociation constants for the ligand–TAR RNA interaction.[a]

Entry	Ligand	K_d (TAR) [μM]
1	neomycin	17.2
2	**1b**	7.5
3	**2b**	2.44
4	**3b**	17.5
5	**4b**	30.3
6	**5b**	n.b.[b]
7	**6b**	3.10
8	**7b**	47.4
9	**8b**	n.b.[b]
10	**S**	n.b.[b]
11	Lys	n.b.[b]
12	Arg	n.b.[b]
13	**9d**	15.3
14	**10d**	30.5
15	**11d**	n.b.[b]
16	**13b**	0.24
17	**14b**	0.11
18	**15b**	1.82

[a] Fluorescence measurements were performed in buffer A (20 mM HEPES, pH 7.4, 20 mM NaCl, 140 mM KCl, and 3 mM $MgCl_2$). K_d values are given with an uncertainty of ±10%. [b] n.b.: no binding.

ly). Interestingly, we observed that the artificial nucleobase **S** alone and the free lysine and arginine amino acid residues were not able to bind to TAR RNA (Table 1, entries 10–12). These results clearly demonstrate that these ligands bear an original mode of binding, in which both **S** and the amino acid interact with TAR RNA in a cooperative way to increase the overall affinity.

The fact that **1b** (Lys), **2b** (Arg), **3b** (His), and **6b** (Trp) bind with high affinity compared with **4b** (Phe), **5b** (Tyr), **7b** (Ala), and **8b** (Thr) is probably due to additional interactions between the functional groups of the amino acid side chains (Lys, Arg, His, and Trp) and TAR RNA. In a similar way, the increased affinity observed with conjugates **3b** (His) and particularly **6b** (Trp) could be ascribed to the possibility of additional H-bonding involving the NH amino groups of the imidazole (His) and indole moieties (Trp). The increased affinity of **6b** relative to that of **3b** (3.10 μM versus 17.5 μM, respectively) could be ascribed to the known proptotropic tautomerism inherent to the imidazole ring (N1–N3 H shift) compared with the indole system.[13]

In the case of ligands **9d**–**11d**, each bearing a histidine residue linked to the nucleobase with a spacer of 1, 3, or 5 carbon atoms, respectively, we clearly observed that short linkers give rise to a more favorable interaction (Table 1, entries 13–15). In fact, a study of the K_d values demonstrated that the C1 linker arm (glycine) of compound **9d** had the best affinity (15.3 μM) compared with those of **10d** (C3, 30.5 μM) and **11d** (C5, no affinity). The use of a longer and more flexible linker probably exposed the imidazole domain out of the helix and led to disfavored interactions. By contrast, a short linker length gives rise to a ligand with a more rigid conformation and favorable interactions.

Encouraged by the results obtained with the first series of compounds, we synthesized ligands **13b** (di-Lys), **14b** (di-Arg), and **15b** (di-His) bearing two amino acid residues in order to further increase the affinity. These conjugates were obtained by coupling **S** and the appropriate amino acids with diaminopropanoic acid as a linker (Scheme 2). Interestingly, **13b**–**15b** showed the best K_d values of 0.24, 0.11, and 1.82 μM, respectively (Table 1, entries 16–18). These results are of great importance because the affinities of these small molecules, particularly **13b** and **14b**, are approximately three- to eightfold higher than that of the previously reported high-molecular-weight Tat peptide (15-mer, $K_d = 0.71$–0.81 μM).[14] These high observed affinities could be ascribed to the additional cooperative interactions involving the second amino acid.

To evaluate the specificity of this class of compounds, the binding affinities of the best ligands were determined in the presence and absence of a 100-fold excess of tRNA (K'_d) and a 100-fold excess of double-stranded DNA (dsDNA; K''_d), respectively (Table 2). For ligands with one amino acid, we observed that **1b** (Lys) and **3b** (His) maintained their affinity toward TAR RNA and were highly specific ($1.1 < K'_d/K_d < 1.4$ and $1.2 < K''_d/K_d < 1.5$; Table 2, entries 2 and 4), whereas **2b** (Arg) showed very low specificity ($K'_d/K_d = 10.7$ and $K''_d/K_d = 7.4$; Table 2, entry 3). For ligands featuring two amino acids, we observed that **13b** (di-Lys) and **14b** (di-Arg) exibited only a minimal decrease of the specificity compared to **1b** and **2b** ($K'_d/K_d = 2.9$ and 3.2, $K''_d/K_d = 3.0$ and 3.6, respectively; Table 2, entries 5 and 6). In contrast, **15b** (di-His) showed a significant decrease of the specificity in the presence of tRNA ($K'_d/K_d = 25$; Table 2, entry 7) and DNA competitors ($K''_d/K_d = 6.5$) relative to

Table 2. Competition assays in the presence of tRNA and dsDNA.[a]

Entry	Ligand	K_d (TAR) [μM]	K'_d (TAR) with tRNA[b] [μM]	K'_d/K_d	K''_d (TAR) with dsDNA[c] [μM]	K''_d/K_d
1	neomycin	17.2	78.5	4.6	n.b.	–
2	**1b**	7.5	10.4	1.4	11.5	1.5
3	**2b**	2.44	26.2	10.7	18.20	7.4
4	**3b**	17.5	18.10	1.1	20.5	1.2
5	**13b**	0.24	0.69	2.9	0.72	3.0
6	**14b**	0.11	0.35	3.2	0.40	3.6
7	**15b**	1.82	45.5	25	11.89	6.5

[a] Fluorescence measurements were performed in buffer A (20 mM HEPES, pH 7.4, 20 mM NaCl, 140 mM KCl, and 3 mM $MgCl_2$). K_d values are given with an uncertainty of ±10%. [b] Measured in the presence of a 100-fold excess of a mixture of natural tRNAs (tRNA mix). [c] Measured in the presence of a 100-fold excess of a 15-mer DNA. n.b.: no binding.

the specificity of its analogue, **3b**, bearing one amino acid ($K'_d/K_d = 1.1$ and $K''_d/K_d = 1.2$; Table 2, entry 4).

From these results, we can conclude that: 1) compared to neomycin, the **S**-based ligands have higher TAR affinities and better specificities and 2) **S** conjugates featuring one amino acid have slightly lower TAR affinities but better specificities than their **S**–diamino acid analogues. These results also illustrate how the binding affinity and specificity of **S**–amino acid conjugates can be tuned by optimizing the linker length and the nature and number of amino acid residues.

To get further insights on the ligand–TAR binding mode, the thermodynamic parameters associated with the formation of the complexes were determined. Nonelectrostatic (ΔG°_{nel}, ΔH°_{nel}, and $T\Delta S^\circ_{nel}$) and electrostatic (ΔG°_{el}, ΔH°_{el}, and $T\Delta S^\circ_{el}$) parameters were obtained by plotting the Gibbs free energy (ΔG°) versus the temperature and by examining the dependency of the dissociation constants on the ionic strength of the solution (Table 3). The enthalpy of binding (ΔH°) is independent of the salt concentration, so only the contribution of the Gibbs energy to the overall binding is discussed in this section. Indeed, the ΔG° value, which represents the total energy, can be divided into two components: 1) the ΔG°_{nel} value, which reflects the contribution of nonelectrostatic interactions to the total free energy, such as nonionic hydrophobic effects driven by entropy, and specific interactions, including H-bonds, van der Waals interactions, and π stacking, and 2) the pure electrostatic (polyelectrolyte) contribution, ΔG°_{el}, which reflects the ionic interactions occurring between two groups of opposite charge and is highly dependent on the salt concentration. As expected, and in line with the K_d values, we found that the interactions between all of the **S**–amino acid conjugates and TAR RNA mainly involve specific nonelectrostatic interactions because the ΔG°_{nel} component represents 81–96% of the overall free energy (ΔG°; Table 3), whereas the neomycin–TAR interaction is clearly dominated by the electrostatic component ($\Delta G^\circ_{nel} = 33\%$ and $\Delta G^\circ_{el} = 67\%$). Interestingly, we also observed that an increase in ammonium groups has no significant effect on the ΔG°_{nel} component because ligands **13b**–**15b**, which contain twice as many ammonium groups as **1b**–**6b**, have very similar contributions from the ΔG°_{nel} values to those obtained for **1b**–**6b** (**1b**/**13b**, **2b**/**14b**,and **3b**/**15b**; Table 3). This original mode of binding could be ascribed to: 1) the high specific H-bonding contribution of the **S** system and 2) the optimal position of the amino acid, which allows additional specific interactions.[10b,15] These results also showed that the overall free energy (ΔG°) is driven by the enthalpy of the binding (ΔH°) and by the high contribution of its nonelectrostatic component (ΔG°_{nel}; Table 3).

Finally, CD spectra were recorded in the absence and presence of ligands to show whether the involved interactions affect the RNA structure. A typical example is shown for **13b**, which was selected for its high affinity (Figure 2). The CD spectrum of the TAR hairpin alone shows strong positive and negative peaks at 265 and 210 nm, respectively, and a weak negative signal at 240 nm, in accordance with the A form of RNA. Addition of one equivalent of ligand **13b** slightly affects the signal at 210 nm. This indicates that the involved interactions do not produce a significant change in the RNA structure, and the contribution of the **S** system (H-bonding) does not abolish the overall base stacking (kissing interaction). In a similar way, when the ligand concentration was increased to 5 or 10 equivalents (high excess), a similar decrease of signal intensity at 210 nm was observed, which is mainly due to the phosphate (TAR)–ammonium (residual excess of ligand) interactions. Moreover, as attested by NMR, fluorescence, and CD spectra of **13b**, no self-structural organization of this ligand was observed at high concentrations, which completely excluded the contribution of **13b** alone to the observed CD variations at 210 nm. The fact that the CD spectra only indicated minimal alteration of the base stacking clearly indi-

Table 3. Thermodynamic parameters for ligand–TAR interactions.

Ligand	ΔG° [kJ mol^{-1}]	ΔH°[a] [kJ mol^{-1}]	$T\Delta S^{\circ}$[a] [kJ mol^{-1}]	$T\Delta S^\circ/\Delta H^\circ$	ΔG°_{nel}[b] [kJ mol^{-1}]	$T\Delta S^\circ_{nel}$ [kJ mol^{-1}]	ΔG°_{el}[c] [kJ mol^{-1}]
neomycin	−26.7	−26.7 ± 1.3	−0.025 ± 0.01	0.00093	−17.8 (67%)	−8.9 ± 0.9	8.9 ± 0.8
1b	−28.8	−38.1 ± 1.5	−9.26 ± 1.5	0.24	−25.7 (89%)	−12.4 ± 1.1	−3.1 ± 1.0
2b	−31.5	−38.4 ± 0.62	−6.84 ± 0.63	0.18	−28.6 (91%)	−9.74 ± 0.9	−2.9 ± 1.1
3b	−25.6	−49.0 ± 3.1	−23.4 ± 3.2	0.61	−24.7 (96%)	−0.9 ± 0.3	−0.9 ± 0.4
13b	−37.1	−46.7 ± 2.4	−9.84 ± 2.9	0.21	−30.1 (81%)	−16.8 ± 2.2	−7.0 ± 1.1
14b	−39.1	−38.7 ± 1.9	−0.317 ± 0.1	0.0082	−35.9 (92%)	−3.52 ± 1.2	−3.2 ± 1.0
15b	−32.2	−58.6 ± 5.1	−26.8 ± 4.9	0.46	−30.2 (94%)	−28.8 ± 1.7	−2.0 ± 0.7

[a] Determined by temperature effect experiments by using the equation $\Delta G^\circ_T = \Delta H^\circ_{Tr} + \Delta C_p(T-Tr) - T\Delta S^\circ_{Tr} - T\Delta C_p \ln(T/Tr)$. See the Supporting Information for definitions and further details. [b] Determined by salt effect experiments by using the equation $\log(K_d) = \log(K_{nel}) - Z\Psi\log[KCl]$. See the Supporting Information for definitions and further details. The percentage of nonelectrostatic interactions ($\Delta G^\circ_{nel}/\Delta G^\circ$) is given in parentheses. [c] $\Delta G^\circ_{el} = \Delta G^\circ - \Delta G^\circ_{nel} = T\Delta S^\circ_{el}$.

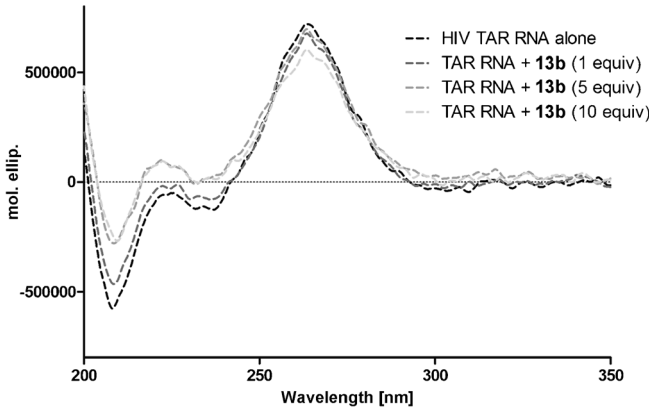


Figure 2. CD spectra of TAR RNA in the absence (black dotted line) and in the presence of 1 equivalent (gark gray dotted line), 5 equivalents (gray dotted line), and 10 equivalents (light gray dotted line) of compound **13b**.

cates that the overall structure of the targeted RNA is maintained, even at high ligand concentrations.

Furthermore, the results obtained by monitoring the fluorescence variation as a function of the concentration of ligand (K_d experiments performed in triplicate) are reproducible and perfectly fit with 1:1 stoichiometry (ligand/RNA), even in the presence of a high excess of ligand. If nonspecific interactions took place, this would have been visible in these fluorescence studies, which would have shown nonconstant values for the stoichiometry ratio, other than 1:1.

Antiviral activity of TAR RNA ligands in cells

The **S**–amino acid conjugates were evaluated for their capacity to inhibit HIV-1 replication. To this end, we determined the ability of the laboratory-adapted HIV-1 NL4.3 strain to replicate in MAGIC-5B cells maintained in the presence of increasing concentrations of the newly synthesized compounds. This indicator cell line expresses the CD4 receptor, the CXCR4 coreceptor, and an integrated copy of the β-galactosidase gene under control of the HIV-1 LTR promoter. The capacity of incoming and newly synthesized Tat to transactivate the retroviral promoter was therefore evaluated by quantification of the β-galactosidase gene expression in the cells. The efficiency of the TAR ligands was compared with that of azidothymidine (AZT), a reference HIV-1 inhibitor, used at a final concentration of 20 μM (Table 4 and Figure 3A).

We found that compounds **1b** (Lys) and **3b** (His) inhibited HIV-1 replication with interesting IC_{50} values of 0.6 and 0.4 μM, respectively (Table 4). Compounds **13b** (di-Lys) and **14b** (di-Arg) exhibited similar activities (1.7 μM), which were approximately fourfold less efficient then **3b**. Most importantly, ligands **2b** (Arg) and **15b** (di-His), which exhibited good affinity but low TAR specificity, were found to be almost ineffective under our experimental conditions. It is noteworthy that neomycin has no anti-HIV activity, even at high concentrations.

Careful analysis of these data clearly showed a strong correlation between the specificity of a ligand for TAR RNA and its anti-HIV activity. For example, the most active ligands (**1b** and **3b**) were those exhibiting the highest specificity ($1.1 < K'_d/K_d < 1.4$ and $1.2 < K''_d/K_d < 1.5$). In a similar way, compounds **13b** and **14b**, which are approximately threefold less specific than **1b** and **3b**, retain the activity, whereas ligands **2b** and **15b**, which have the lowest degree of specificity, are completely inactive.

Table 4. Cellular inhibition of viral proliferation.

Ligand	IC_{50}[a] [μM]
neomycin	n.a.
1b	0.63
2b	n.a.
3b	0.41
13b	1.78
14b	1.79

[a] IC_{50}: 50% inhibitory concentration. The obtained values are given with an uncertainty of ±5%. n.a: no activity up to 50 μM.

Finally, cytotoxicity assays performed by incubating increasing concentrations of drugs with MAGIC-5B cells for 48 h revealed no modification of cell viability relative to cells incubated in the presence of culture medium (Figure 3B). These results clearly indicate that this class of ligands is not toxic in the given cells. Moreover, the fact that these ligands inhibited HIV-1 replication in vitro clearly attested for their capacity to penetrate into the cell and to interfere with the retroviral replication cycle. Hence, this class of ligands clearly shows an interesting profile and could be used for further investigations in RNA targeting.

Conclusion

A new series of artificial nucleobase–amino acid conjugates has been synthesized and the binding affinity of the compounds for a labeled HIV-1 TAR RNA has been evaluated by fluorescence spectroscopy. In addition to the obtained K_d values, competitive assays with tRNA and dsDNA were performed. We observed that: 1) **S** conjugates comprising one basic amino acid displayed higher affinity than neomycin and strong specificity for TAR RNA and 2) **S**–diamino acid conjugates exhibited enhanced ligand–RNA binding relative to that of their monosubstituted analogues while retaining significant degrees of specificity. These properties can be attributed to the exceptional cooperative mode of binding.

Furthermore, we have unambiguously established that this class of ligand is highly stable in biological media and active in cell-based assays in the submicromolar range, without apparent toxicity, and that the specificity of the **S**-based ligands for the TAR RNA model is strongly correlated with the in vitro activity.

Altogether, these results support the idea that amino acids and Hoogsteen based binding can be combined to create a new class of high-affinity, high-specificity RNA-binding ligands. This approach is highly promising and could be applied for the rational targeting of different RNA structures (structure-

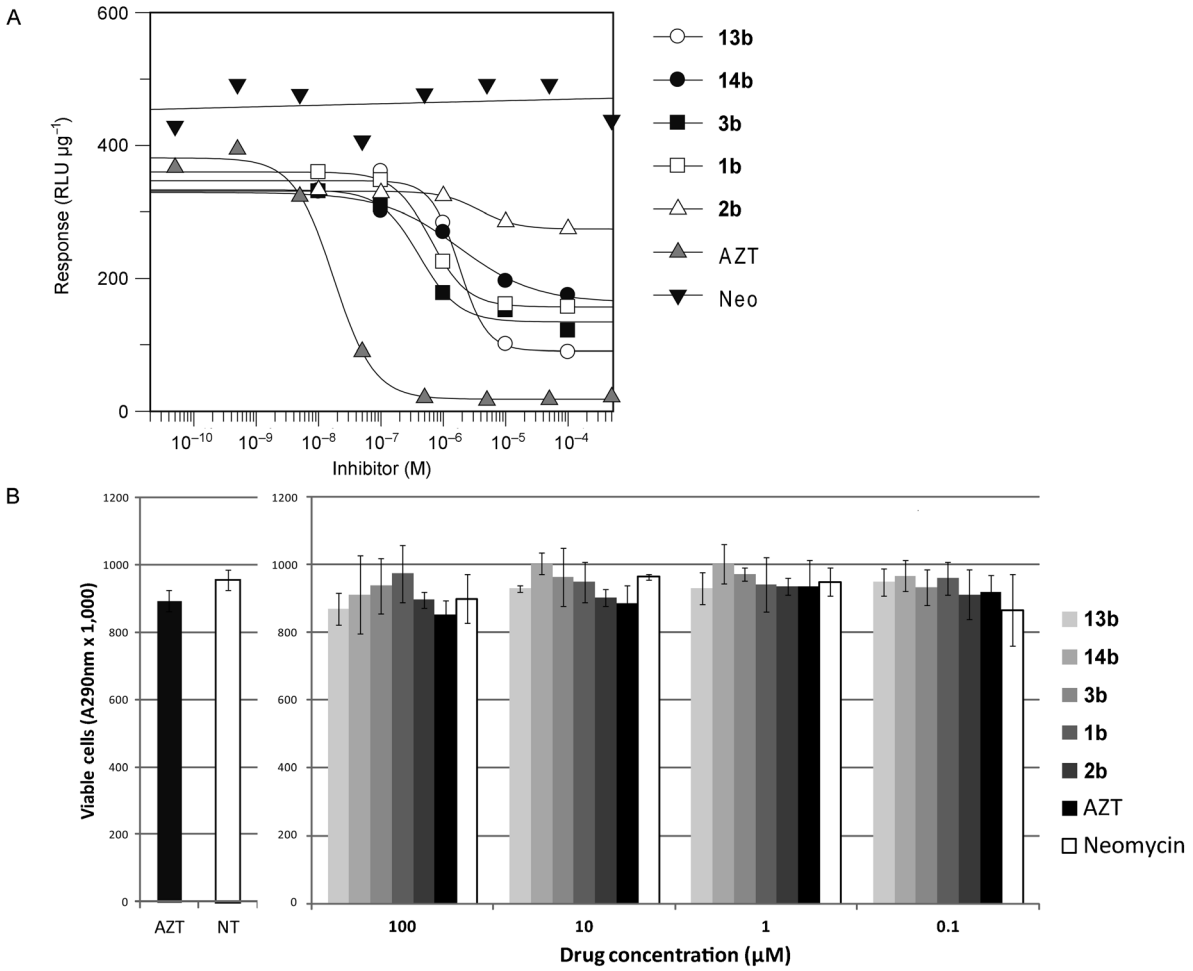


Figure 3. In vitro activity of RNA ligands. A) The capacity of various concentrations of TAR RNA ligands to inhibit replication of the HIV-1 NL4.3 strain in cultures of MAGIC-5B cells was determined by quantification of the β-galactosidase reporter gene activity in the cell culture. B) The toxicity of the TAR ligands was compared with that of AZT. MAGIC-5B cells were exposed to increasing concentrations of ligands for 48 h and the cell viability was measured as described in the Experimental Section. The cell viability was similar to the control conditions for any ligand concentration tested. The controls consisted of mock infected cells or HIV-1-infected cells maintained in the absence of drug (not treated, NT) or in medium supplemented with the nucleoside reverse transcriptase inhibitor AZT and the results are presented in the left-hand panel.

based ligand design), including viral and oncogenic RNAs in tumor cells.

Experimental Section

Materials

Solvents and most of the starting materials were purchased from Aldrich and Alfa Aesar. All reactions involving air- or moisture-sensitive reagents or intermediates were performed under an argon atmosphere. Flash column chromatography was carried out on silica gel (Merck, SDS 60A, 40–63 μm). Analytical thin-layer chromatography (TLC) was conducted on Merck precoated silica gel 60F254 plates and compounds were visualized with the ninhydrin test and/or under ultraviolet light (254 nm). ^{1}H and ^{13}C NMR spectra were recorded on a Bruker AC spectrometer (200 MHz or 500 MHz). Chemical shifts (δ) are reported in parts per million (ppm) referenced to the residual ^{1}H resonance of the solvent ($CDCl_3$, $\delta = 7.26$ ppm; CD_3OD, $\delta = 3.31$ ppm; D_2O, $\delta = 4.79$ ppm; [D_6]DMSO, $\delta = 2.50$ ppm). ^{13}C spectra were referenced to the residual ^{13}C resonance of the solvent ($CDCl_3$, $\delta = 77.3$ ppm; CD_3OD, $\delta = 49.0$ ppm; [D_6]DMSO, $\delta = 39.5$ ppm). Splitting patterns have been designated as follows: s: singlet; d: doublet; dt: doublet of triplets; t: triplet; q: quartet; m: multiplet; br: broad. Coupling constants (*J* values) are listed in hertz (Hz). High-resolution mass spectra were obtained with a LTQ Orbitrap hybrid mass spectrometer with an electrospray ionization probe (Thermo Scientific, San Jose, CA) by direct infusion from a pump syringe to confirm the correct molar mass and high purity of the compounds. The final products were analyzed by HPLC on a Waters Alliance 2695 pump coupled with a Waters 996 photodiode array detector and a Thermo Scientific Betasil RP C18 column (250×4.6 mm, 5 μm). Solvent A (0.1% TFA in water and solvent B (0.1% TFA in acetonitrile) were used for HPLC studies. A gradient of A/B (100:0 to 40:60 for 30 min) was employed at a flow rate of 1 mL min^{-1}.

Unless otherwise stated, all reagents and solvents were of analytical grade and were obtained from Sigma. HEPES and all inorganic salts for buffers were purchased from Aldrich (molecular biology grade). RNA, Tat peptide, and DNA oligonucleotides were purchased from IBA GmbH and used without further purification.

Buffers

All buffers were filtered through 0.22 μm Millipore filters (GP ExpressPLUS membrane). A small aliquot (100 mL) was first filtered and then discarded to avoid any contaminants that might be leached from the filter. The solutions to be used in the fluorescence experiments were prepared by diluting the concentrated stocks in Milli-Q water and filtering again as described above. All standard fluorescence measurements were performed in buffer A (20 mM HEPES, pH 7.4, containing 20 mM NaCl, 140 mM KCl, and 3 mM $MgCl_2$, at 25 °C). For competitive experiments in the presence of tRNAs, a mixture of pre- and mature yeast tRNAs (containing over 30 different species from baker's yeast (*Saccharomyces cerevisiae*, Sigma, type X-SA)) was added to buffer A to obtain a 100-fold nucleotide excess with regard to TAR RNA. Stock solutions of tRNAs were prepared in water and quantified by using an extinction coefficient of 9640 $cm^{-1}M^{-1}$ per base. For competitive experiments in the presence of a dsDNA, a 15-mer sequence (5′-CGTTTTTATTTTTGC-3′) and its complement, annealed beforehand, were added to buffer A to obtain a 100-fold nucleotide excess with regard to TAR RNA (900 nM duplex; 5 nM RNA).

Binding studies and K_d determination

Ligand solutions were prepared as serial dilutions in buffer A at a concentration four times higher than the desired final concentration to allow for the subsequent dilution before the addition of the RNA solution. An automated pipetting system (epMotion 5075, Eppendorf) was used in order to perform these binding studies on 384-well plates (Greiner bio-one). Refolding of the RNA was performed by using a thermocycler (ThermoStat Plus, Eppendorf) as follows: the 5′-Alexa^{488} TAR RNA (0.2 nmol) was diluted in buffer A (1 mL), denatured by heating to 90 °C for 2 min, cooled to 4 °C for 10 min, then incubated at 25 °C for 15 min. After refolding, the RNA was diluted to a working concentration of 20 nM through addition of the appropriate amount of buffer A. After addition of each ligand (30 μL) on the 384-well plates in 15 dilutions (from 1 mM to 61 nM as final concentrations) and in duplicate, the RNA solution (30 μL) was added to each well containing ligand to give a final concentration of 10 nM. The fluorescence was measured on a GeniosPro apparatus (Tecan) with an excitation filter of (485 ± 10) nm and an emission filter of (535 ± 15) nm. Each point was measured five times with a 500 μs integration time and averaged. Binding was allowed to proceed overnight at 5 °C to achieve equilibrium. To study the temperature dependence, the plates were incubated after overnight equilibrium at different temperatures ranging from 5 °C to 35 °C. Neomycin was used as a control because its binding to TAR has already been studied by using several methods.[12] Its K_d value of (12.0 ± 3.6) μM is in good agreement with previously reported values.

Data analysis

Binding data were analyzed by using Graphpad Prism 5 software. Unless otherwise stated, binding profiles were well modeled by using a simple model and assuming one-to-one stoichiometry. A higher initial fluorescence value is observed in the presence of tRNA, which is consistent with the modification of the polarity of the solvent and a small amount of fluorescence from the tRNA mixture.

Antiviral activity

The $CD4^+$ $CXCR4^+$ MAGIC-5B indicator cell line, which stably expresses the β-galactosidase reporter gene cloned downstream of the HIV-1 LTR, was cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum, 1% glutamax, and 1% penicillin–streptomycin antibiotic mixture (Life Technologies, Inc.) at 37 °C in a 5% CO_2 atmosphere. Each drug was prepared in four serial tenfold dilutions with culture medium. MAGIC-5 cells were seeded (12500 per well) and cultured in a 96-well tissue culture plate for 24 h. The cells were preincubated with the diluted compounds for 2 h at 37 °C and then challenged with the HIV-1 NL4.3 strain at a multiplicity of infection (m.o.i.) of 1. After 48 h, the cells were washed and lysed, and infection of the cell culture was monitored by measurement of the β-galactosidase activity from the total cell lysate by using the Galacto-Star chemiluminescent assay kit according to the manufacturer instructions (Applied Biosystem, USA). Luminescence was recorded by using a Centro XS3 LB960 luminometer (Berthold, France). Values were normalized with respect to the exact protein content of each cell lysate, as determined with the BCA protein assay (Thermo Scientific). The susceptibility of HIV-1 to the various compounds was determined by using triplicate samples of infected MAGIC-5B cells in each assay. The IC_{50} value for each drug was estimated from plots of luciferase per μg of protein reduction versus drug concentration.

Cell viability assays

The viability of cells exposed to the drugs was assessed by incubating MAGIC-5B cells (12500 per well in a 96-well tissue culture plate) in the presence of increasing concentrations of the drugs. After 48 h of culture, the cell viability was determined by using the CellTiter 96 AQueous One Solution cell proliferation assay (Promega). Determinations were performed in quadruplicate.

Acknowledgements

This work was supported by CNRS, ANR, Université de Nice Sophia Antipolis, Conseil Régional PACA, MAE, ANRT, and Oribase Pharma (grant to J.-P.J.).

Keywords: amino acids · antiviral agents · cooperative binding · nucleobases · RNA recognition

[1] For recent reviews, see: a) L. Guan, M. D. Disney, *ACS Chem. Biol.* **2012**, *7*, 73–86; b) R. Moumne, M. Catala, V. Larue, L. Micouin, C. Tisne, *Biochimie* **2012**, *94*, 1607–1619; c) Y. Tor, *Pure Appl. Chem.* **2009**, *81*, 263–272; d) J. R. Thomas, P. J. Hergenrother, *Chem. Rev.* **2008**, *108*, 1171–1224; e) N. Foloppe, N. Matassova, F. Aboul-Ela, *Drug Discovery Today* **2006**, *11*, 1019–1027.

[2] For recent examples, see: a) K. Limmer, D. Aschenbrenner, H. E. Gaub, *Nucleic Acids Res.* **2013**, *41*, e69; b) T. Lombès, R. Moumné, V. Larue, E. Prost, M. Catala, M. T. Lecourt, F. Dardel, L. Micouin, T. Carine, *Angew. Chem.* **2012**, *124*, 9668–9672; *Angew. Chem. Int. Ed.* **2012**, *51*, 9530–9534; c) S. Umemoto, S. Im, J. Zhang, M. Hagihara, A. Murata, Y. Harada, T. Fukuzumi, T. Wazaki, S. Sasaoka, K. Nakatani, *Chem. Eur. J.* **2012**, *18*, 9999–10008; d) R. Parkesh, J. L. Childs-Disney, M. Nakamori, A. Kumar, E. Wang, T. Wang, J. Hoskins, T. Tran, D. Housman, C. A. Thornton, M. D. Disney, *J. Am. Chem. Soc.* **2012**, *134*, 4731–4742; e) C. L. Beisel, Y. Y. Chen, S. J. Culler, K. G. Hoff, C. D. Smolke, *Nucleic Acids Res.* **2011**, *39*, 2981–2994; f) P. Lopez-Senin, I. Gomez-Pinto, A. Grandas, V. Marchan, *Chem. Eur. J.* **2011**, *17*, 1946–1953.

[3] S. Feng, E. C. Holland, *Nature* **1988**, *334*, 165–167.

[4] S. Bannwarth, A. Gatignol, *Curr. HIV Res.* **2005**, *3*, 61 – 71.
[5] L. M. Alvarez-Salas, *Curr. Top. Med. Chem.* **2008**, *8*, 1379 – 1404.
[6] a) R. S. Geary, *Expert Opin. Drug Metab. Toxicol.* **2009**, *5*, 381 – 391; b) R. Juliano, J. Bauman, H. Kang, X. Ming, *Mol. Pharm.* **2009**, *6*, 686 – 695; c) X. Zhao, F. Pan, C. M. Holt, A. L. Lewis, J. R. Lu, *Expert Opin. Drug Delivery* **2009**, *6*, 673 – 686.
[7] a) A. Davidson, K. Patora-Komisarska, J. A. Robinson, G. Varani, *Nucleic Acids Res.* **2011**, *39*, 248 – 256; b) Z. Athanassiou, R. L. Dias, K. Moehle, N. Dobson, G. Varani, J. A. Robinson, *J. Am. Chem. Soc.* **2004**, *126*, 6906 – 6913; c) S. Chirayil, R. Chirayil, K. J. Luebke, *Nucleic Acids Res.* **2009**, *37*, 5486 – 5497.
[8] a) S. Kumar, P. Kellish, W. E. Robinson, Jr., D. Wang, D. H. Appella, D. P. Arya, *Biochemistry* **2012**, *51*, 2331 – 2347; b) S. Magnet, J. S. Blanchard, *Chem. Rev.* **2005**, *105*, 477 – 498; c) M. J. Belousoff, B. Graham, L. Spiccia, Y. Tor, *Org. Biomol. Chem.* **2009**, *7*, 30 – 33; d) M. D. Disney, L. P. Labuda, D. J. Paul, S. G. Poplawski, A. Pushechnikov, T. Tran, S. P. Velagapudi, M. Wu, J. L. Childs-Disney, *J. Am. Chem. Soc.* **2008**, *130*, 11185 – 11194.
[9] M. Duca, V. Malnuit, F. Barbault, R. Benhida, *Chem. Commun.* **2010**, *46*, 6162 – 6164.
[10] a) V. Malnuit, M. Duca, R. Benhida, *Org. Biomol. Chem.* **2010**, *8*, 326 – 336; b) D. Guianvarc'h, R. Benhida, J-L. Fourrey, R. Maurisse, R. J.-S. Sun, *Chem. Commun.* **2001**, 1814 – 1815; c) D. Guianvarch, J.-L. Fourrey, R. Maurisse, J.-S. Sun, R. Benhida, *Bioorg. Med. Chem.* **2003**, *11*, 2751 – 2759; d) Y. Wang, D. A. C. Powers, S. O. Lack, D. Osborne, K. R. Fox, T. Brown, *Biochemistry* **2005**, *44*, 5884 – 592.
[11] a) V. Bonnard, S. Azoulay, A. Di Giorgio, N. Patino, *Chem. Commun.* **2009**, 2302 – 2304; b) N. W. Luedtke, Q. Liu, Y. Tor, *Biochemistry* **2003**, *42*, 11391 – 11403; c) J. R. Thomas, X. Liu, P. J. Hergenrother, *J. Am. Chem. Soc.* **2005**, *127*, 12434 – 12435.
[12] a) T. Tran, M. D. Disney, *Biochemistry* **2010**, *49*, 1833 – 1842; b) Y. Xie, A. V. Dix, Y. Tor, *J. Am. Chem. Soc.* **2009**, *131*, 17605 – 17614; c) S. Freisz, K. Lang, R. Micura, P. Dumas, E. Ennifar, *Angew. Chem.* **2008**, *120*, 4178 – 4181; *Angew. Chem. Int. Ed.* **2008**, *47*, 4110 – 4113; d) J. Kondo, B. Francois, A. Urzhumtsev, E. Westhof, *Angew. Chem.* **2006**, *118*, 3388 – 3392; *Angew. Chem. Int. Ed.* **2006**, *45*, 3310 – 3314; e) F. Zhao, Q. Zhao, K. F. Blount, Q. Han, Y. Tor, T. Hermann, *Angew. Chem.* **2005**, *117*, 5463 – 5468; *Angew. Chem. Int. Ed.* **2005**, *44*, 5329 – 5334.
[13] S. Li, M. Hong, *J. Am. Chem. Soc.* **2011**, *133*, 1534 – 1544.
[14] a) D. I. Bryson, W. Zhang, P. M. McLendon, T. M. Reineke, W. L. Santos, *ACS Chem. Biol.* **2012**, *7*, 210 – 217; b) H. Suryawanshi, H. Sabharwal, S. Maiti, *J. Phys. Chem. B* **2010**, *114*, 11155 – 11163.
[15] A. L. Smith, J. Kassman, K. J. Srour, A. M. Soto, *Biochemistry* **2011**, *50*, 9434 – 944.

Received: September 17, 2013
Published online on January 15, 2014

