Exploring the impact of the side-chain length on peptide/RNA binding events

Lola Sbicca,[a] Alejandro Lo´pez Gonza´lez,[b] Alexandra Gresika,[a] Audrey Di Giorgio,[a] Jordi Teixido Closa,[b] Roger Estrada Tejedor,[b] Marie-Line Andre´ola,[c] Ste´phane Azoulay[a] and Nadia Patino

The impact of the amino-acid side-chain length on peptide–RNA binding events has been investigated using HIV-1 Tat derived peptides as ligands and the HIV-1 TAR RNA element as an RNA model. Our studies demonstrate that increasing the length of all peptide side-chains improves unexpectedly the binding affinity (K_D) but reduces the degree of compactness of the peptide–RNA complex. Overall, the side-chain length appears to modulate in an unpredictable way the ability of the peptide to compete with the cognate TAR RNA partner. Beyond the establishment of non-intuitive fundamental relationships, our results open up new perspectives in the design of effective RNA ligand competitors, since a large number of them have already been identified but few studies report on the modulation of the biological activity by modifying in the same way the length of all chains connecting RNA recognition motives to the central scaffold of a ligand.

Introduction

Non-coding RNAs (ncRNAs) play essential roles in a wide range of biologically relevant systems, representing attractive targets for therapeutic intervention.[1–3] They fold into a multitude of irregular hairpin structures, giving rise to unique pockets, well-suited for the specific binding of molecules targeting them.[4] However, due to the high conformational dynamics of these ncRNAs it is difficult to predict and design specific ligands. Despite the huge efforts already made in this field and except antibiotics – that target bacterial RNAs,[5] no RNA ligand has led to therapeutics. Therefore, there is still fundamental interest in gaining deep insight into the mechanisms of RNA recognition by small ligands.

RNA hairpins may coexist as a set of conformations in a non-bound state and ligand recognition likely occurs through conformational capture as in the case of protein-RNA complexes.[6–10] In these processes, both partners may undergo marked conformational changes, the ligand flexibility being expected to have a crucial role in the formation of the complex. One way to measure the impact of the ligand flexibility on such adaptive processes is by comparing the thermodynamic binding mode of a series of ligands that have the same interaction potential for an RNA but display different flexibilities. Such studies may afford a quantification of the contribution of ligand flexibility to the thermodynamics of binding,[11–13] highlighting non-intuitive structure–activity relationships and validating *in silico* predictions.[14] RNA-binding proteins play crucial roles in RNA processing and function as regulators of gene expression.[15] Numerous peptidic derivatives mimicking RNA-binding protein regions have been developed to target RNAs of therapeutic interest. In particular, the HIV-1 TAR RNA (Trans Activation Responsive Region) element has been targeted by a myriad of molecules,[2,16–19] among them short peptides (or analogs) deriving from the HIV-1 Tat basic region. The HIV transcription step involves the formation of a Tat–TAR complex[20,21] and the inhibition or disruption of this interaction, by means of synthetic TAR ligands, represents an attractive strategy to inhibit viral replication.[22] Thereby, the TAR–Tat system constitutes an ideal model for fundamental studies of the RNA recognition mode. Consequently, we designed two synthetic peptides deriving from the Tat basic fragment ($Tat_{49–57}$: RKKRRQRRR), constituted by non-canonical a-amino acid residues with all their side-chains either one methylene longer (Tat(+)1) or shorter (Tat(−)1) than the natural ones (Fig. 1A and B). Such peptides constitute a model of ligands based on a common peptidic scaffold, connected to RNA recognition motives *via* flexible linkers of different lengths. Some studies report that altering the side-chain length of one or several R

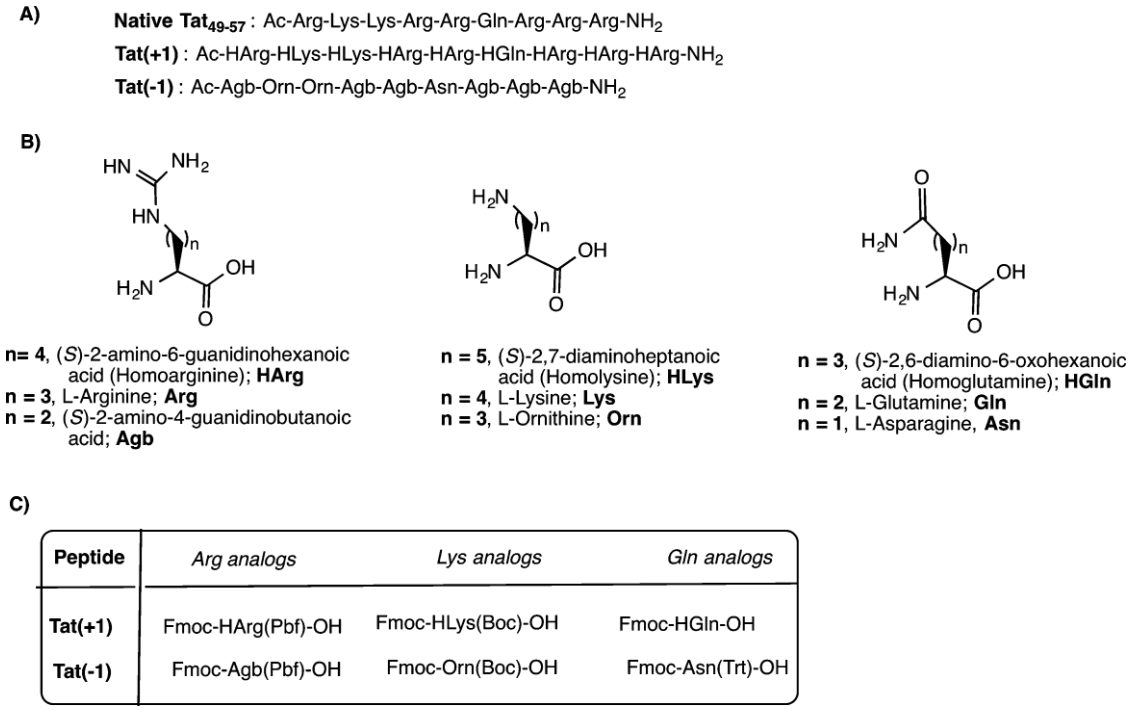


Fig. 1 (A) Design of $Tat_{49–57}$-derived peptides containing in their amino acid side-chains either one methylene group more (Tat(+1)) or less (Tat(−1)), than Tat. (B) Chemical structures of Arg, Lys, Gln and of the corresponding analogs with varying side-chain lengths. (C) Protected amino acids used for the solid-phase syntheses of Tat(+1) and Tat(−1).

and K residues of the basic Tat domain[23,24] and of other cationic peptides[25] has an impact both on the RNA binding affinity and specificity, as well as on the HIV Tat/TAR-mediated protein expression. However, modifying the length of only some residues while keeping the others unchanged has an unpredictable effect that is highly position-dependent. By contrast, adding (Tat(+1)) or removing (Tat(−1)) one methylene group to all amino acid side-chains of a peptide ligand ($Tat_{49–57}$) can not modify the RNA interaction network and comparing binding thermodynamics should allow a quantification of the impact of the side-chain length on the RNA recognition mode. We have performed the synthesis of Tat(−1) and Tat(+1) and established the thermodynamics of association with TAR RNA, comparatively to the natural $Tat_{49–57}$ peptide. The ability of Tat(−1) and Tat(+1) to compete with a fluorescent $Tat_{48–57}$ fragment has also been investigated. Molecular modeling studies have been carried out in order to support our assumptions based on experimental data. Lastly, the inhibitory effects of the three peptides on the HIV-1 replication in cells have also been measured.

Materials and methods

Synthetic procedures of new compounds and their characterization (NMR, HPLC, ESI MS and HRMS) are given in the ESI.†

TAR binding studies

Materials and equipment. Unless otherwise stated, all reagents and solvents were of analytical grade and are purchased from Sigma (St Louis, U.S.A.). HEPES [4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid] and all inorganic salts for buffers were purchased from Calbiochem-Merck Millipore (Fontenay sous Bois, France) (molecular biology grade). 5′-labelled and unlabeled RNA oligonucleotides (TAR: CCAGAUCUGAGCCUGGGAGCUCU CUGG; TARab: CCAGAGAGCCUGGGAGCUCUCUGG and IRES: GCCGAGUAGUGUUGGGUCGCGAAAGGC) were purchased from IBA GmbH (Gottingen, Germany) and used without further purification. The native *N*-acetyl Tat fragment derivative ($Tat_{49–57}$: RKKRRQRRR) was purchased from Eurogentec. The Fluorescein labeled $Tat_{48–57}$ peptide (GRKKRRQRRR) was purchased from EZBiolab (Carmel, U.S.A.) and used without further purification. All buffers were filtered through 0.22 mm Millipore filters (GP ExpressPLUS membrane).

All standard fluorescence measurements were performed in buffer A (20 mM HEPES (pH 7.4 at 25 °C), 20 mM NaCl, 140 mM KCl and 3 mM $MgCl_2$).

FRET experiments were performed in buffer B (50 mM tris buffer (pH 7.4 at 25 °C), 20 mM KCl and 0.005% tween 20).

UV melting experiments were performed in buffer C (10 mM sodium cacodylate, 10 mM NaCl (pH 7.5) and 0.1 mM EDTA). Prior to all experiments, refolding of the RNA was performed using a thermocycler (ThermoStatPlus Eppendorf) as follows: the RNA, diluted in 1 mL of the appropriate buffer, was first denatured by heating to 90 °C for 2 min then cooled to 4 °C for 10 min followed by incubation at 20 °C for 15 min. After refolding, the RNA was diluted to the appropriate working concentration.

Fluorescence binding assays. Ligand solutions were prepared as serial dilutions using an epMotion automated pipetting

system (eppendorf) in buffer A at a concentration twice higher than the desired final concentration to allow for the subsequent dilution during the addition of the 5′ Alexa488 labelled RNA (TAR, TAR_{ab} or $IRES_{IIID}$) solution. The appropriate ligand solution (30 mL) was added to a well of a non-treated black 384-well plate (Nunc 237105), in triplicate. 30 mL of the 10 nM RNA solution was then added to each well containing ligand. This subsequent dilution lowered the final RNA concentration to 5 nM. The fluorescence was measured on a GeniosPro (Tecan) with an excitation filter of 485 10 nm and an emission filter of 535 15 nm. Each point was measured 5 times with a 500 ms integration time and averaged. Binding was allowed to proceed at least 30 min at room temperature to achieve equilibrium.

To study the temperature dependence, the plates were incubated for 30 min at different temperatures ranging from 5 °C to 35 °C.

FRET competitive assays. Ligand solutions and RNA (40 nM working solutions) were prepared as described above in buffer B. The labeled fluorescein Tat peptide (40 nM in buffer B) was mixed to an equal volume of 5′-labelled DABCYL TAR RNA for 20 min at room temperature to form a Tat/TAR complex before adding the ligand. The appropriate ligand solution (30 mL) was added to a well of a 384-well plate, in triplicate, followed by 30 mL of the Tat/TAR solution. Fluorescence was measured as described above after 30 min of incubation at room temperature.

Data analysis

Binding data (K_D and FRET experiments) were analyzed using Prism 5 (GraphPad Software) by nonlinear regression following eqn (1):

$$Y = \text{Bottom} + (\text{Top}-\text{Bottom})/(1 + 10^{((\log A - X)*\text{HillSlope})}) \quad (1)$$

where A is either K_D or IC_{50}, K_D values were converted to $\Delta G°$ values as $\Delta G° = RT \ln K_D$.

For thermodynamic analysis, $\Delta G°$ values were plotted *versus* T. Nonlinear regression using the three-parameter fit in Prism 5 was used to fit the following equation to the data:

$$\Delta G°_T = \Delta H°_{Tr} + \Delta C_p(T - Tr) - T\Delta S°_{Tr} - T\Delta C_p \ln(T/Tr) \quad (2)$$

where Tr is a constant reference temperature (in our study Tr = 293.15 K), and the three fit parameters are $\Delta H°_{Tr}$ the change in enthalpy upon binding at Tr; $\Delta S°_{Tr}$, the change in entropy upon binding at Tr; and ΔC_p, the change in heat capacity. ΔC_p was assumed to be independent of temperature; inclusion of a $\Delta C_p/\Delta T$ term in the analysis did not improve the quality of the fits and gave larger standard errors for the returned parameters.

$\Delta H°_T$ and $\Delta S°_T$ were calculated from the results derived from the fitting of the curve $\Delta G°$ values *versus* T by eqn (3a) and (3b) using:

$$\Delta H°_T = \Delta H°_{Tr} + \Delta C_p(T - Tr) \quad (3a)$$

$$\Delta S°_T = \Delta S°_{Tr} + \Delta C_p \ln(T/Tr) \quad (3b)$$

where $\Delta H°_T$ is the change in enthalpy upon binding at T (25 °C) and $\Delta S°_T$, the change in entropy upon binding at T.

Temperature-dependent UV spectroscopy (UV melting). Thermal denaturation scans were obtained using a Cary 300 (Varian) spectrophotometer equipped with an electrothermal multicell holder. Absorbance *versus* temperature profiles were recorded at 260 nm. After structuration of unlabeled TAR RNA and incubation (1 h) with the corresponding ligand, the temperature was raised from 20 to 90 °C, at a heating rate of 0.5 °C min^{-1}. Thermal denaturation studies were carried out at 2 mM TAR RNA with 2 mM of the ligand or without ligand (TAR alone). The experiments were performed in buffer C. The melting temperature (T_m) value was taken as the midpoint of the melting transition as determined by the maximum of the first-derivative plot using Prism software.

Molecular dynamics simulations

Flexible alignment from MOE software (2014.06, Chemical computing group Inc.) was used to obtain the starting complex conformations for each peptide using PDBID 2 kdq as a template. Starting from the best scored superposition complex, a molecular dynamics system was prepared using the explicit TIP3P water solvent for its equilibration, simulations and free energy evaluation.

Complexes were prepared with *xleap* using the modified AMBER force field ff12SB. Force field parameters for non-standard amino acids were calculated by B3LYP/6-31G* DFT using Gaussian03.

Each structure was minimized in two steps: first, all residues except the solvent were held fixed with a restraint force of 100 kcal mol^{-1} $Å^{-2}$. Steepest descent minimization followed by the conjugate gradient was performed using 2500 steps in both cases. The same minimization protocol was applied in 10 000 steps without positional restraints. After minimization, the temperature was raised from 0 to 300 K in 100 ps using constant volume dynamics. A restraint force of 10 kcal mol^{-1} $Å^{-2}$ was applied to the complex and SHAKE was turned on for bonds involving hydrogen atoms. Then, 200 ps of constant pressure dynamics were applied for density equilibration. Finally, each production run lasted 20 ns, performed using Langevin dynamics without heavy atom positional restraints and defining a collision frequency of 1 ps^{-1}. An atom-based long range cutoff of 9 Å was applied during all the simulations.

Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) thermodynamic analysis was performed to predict the relative binding energy between peptides and RNA. An ensemble of conformations was extracted from the last 2 ns of MD simulations and normal modes were calculated to approximate vibrational entropies. MM/GBSA energies were calculated for each complex of the last 2 ns of the MD simulations and the salt concentration was set to 0.15 M.

Antiviral assays

The infectivity was assayed on HeLa P4 cells expressing CD4 receptors and the *lacZ* gene under the control of the HIV-1 LTR.[26] HeLa P4 cells were plated using 200 mL of DMEM medium

(Invitrogen) supplemented with 10% inactivated Fetal Calf Serum (FCS), 1 mg mL^{-1} geneticin (G418, Gibco-BRL), gentamicin (45 mg mL^{-1}), in 96 multiwell plates, at a density of 10 000 cells per well. After overnight incubation at 37 °C, the supernatant was discarded and 200 mL of the fresh medium containing HIV-1 in the presence or absence of peptides was added. Twenty-four hours later, the supernatant was discarded and the wells were washed three times with 9% NaCl. Each well was refilled with 200 mL of a reaction buffer containing 50 mM Tris–HCl (pH 8.0), 100 mM b-mercaptoethanol, 0.05% Triton X-100, and 5 mM 4-methyl-umbelliferyl-b-D-galactoside (4-MUG). After 24 h at 37 °C, the reaction products were assessed in a fluorescence microplate reader (Cytofluor II) at 360/460 nm excitation/emission.

Results

Chemistry

The structures of the $Tat_{49–57}$-derived peptides containing all the side-chains either one methylene longer (Tat(+1)) or shorter (Tat(−1)), as compared to $Tat_{49–57}$, are described in Fig. 1(A and B).

The syntheses of Tat(+1) and Tat(−1) were performed following classical solid-phase protocols[27] on a Rink-Amide MBHA resin, using HBTU as a coupling reagent and starting from commercially available Na-Fmoc amino acid residues (Fig. 1B and C), except for Fmoc-HLys(Boc)-OH and Fmoc-HGln-OH whose preparations are described below (Schemes 1 and 2, respectively). Tat(+1) and Tat(−1) were purified by reverse-phase high performance liquid chromatography (HPLC) and their structure was confirmed using High Resolution Mass Spectroscopy (HRMS).

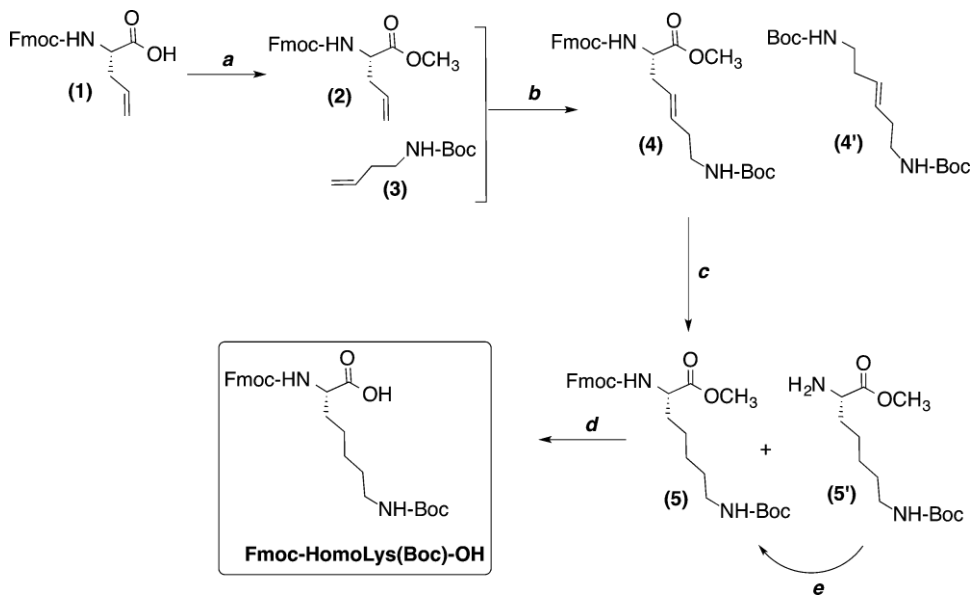
Fmoc-HLys(Boc)-OH and Fmoc-HGln-OH were prepared according to slightly modified published procedures, *via* olefin cross metathesis[28] and selective oxidation of Z-Ne-lysine, respectively.[29] Fmoc-HLys(Boc)-OH was obtained in 4 steps (Scheme 1), starting from Fmoc-Gly(Allyl)-OMe (2) and *N*-Boc allylamine (3). The olefin cross-metathesis reaction was performed by means of the Grubbs catalyst (1st generation), yielding compounds (4) and (4^0) in 40% and 46%, respectively. The low yield obtained in (4) is likely due to the presence of the hundred Fmoc group, since the same reaction starting from Z-Gly(Allyl)-OMe afforded the expected product in 69%.[28,29] Hydrogenolysis of (4) in THF led to Fmoc-HLys(Boc)-OMe (5) in 41% yield together with H-HLys(Boc)-OMe (5^0) (40% yield). Although the Fmoc group is generally recognized as a stable protective group under hydrogenation conditions,[30] its cleavage has been nevertheless reported in some cases.[31,32] Finally, saponification of (5) with LiOH 1 N afforded the Fmoc-HLys(Boc)-OH residue in 95% yield.

On the other hand, Fmoc-HGln-OH was prepared in 31% overall yield (Scheme 2), starting from the commercially available H-Lys(Z)-OH residue (6) following a three-step procedure: (i) $KMnO_4$ oxidation of the Ne-Z protected side-chain of the lysine residue, (ii) cleavage of the Z group under strong acidic conditions (TFMSA/TFA) (iii) protection of the a-amino group using Fmoc-OSu.

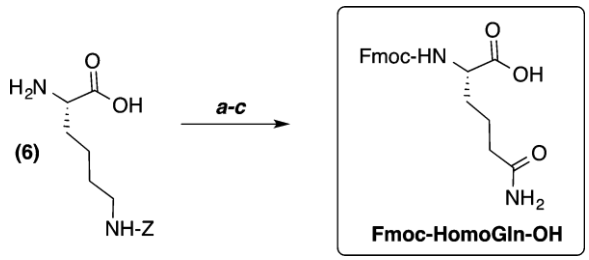
Interaction studies

Affinity and specificity. Dissociation constants (K_D(TAR); Table 1) associated with TAR-peptide complexes were determined by monitoring the fluorescence change of a TAR (18–44) fragment labeled with a fluorescent group (Alexa 488), as previously described.[13] All three peptides bind to TAR with sub-micromolar affinities. The apparent K_D for the peptide–RNA complexes decreased as the side-chain length increases (Table 1). Thus, Tat(+1) is the most affine compound while Tat(−1) displays the weakest affinity.

We then measured the binding affinities of the peptides for two other hairpin RNA fragments, the bulgeless TAR sequence (TAR_{ab}) and the hepatitis C virus IRES IIId stem-loop ($IRES_{IIID}$)[33]



Scheme 1 *Reagents and conditions*: (a) Cs_2CO_3, MeI, DMF, 85%; (b) $RuCl_2(PCy_3)_2(QCHPh)$ (0.2 eq.), DCM, 40%; (c) H_2, Pd/C (0.2 eq.), THF, 41%; (d) LiOH 1 N, dioxane, 0 °C, 75%. (e) Fmoc-OSu, DCM, DIPEA, 95%.



Scheme 2 *Reagents and conditions*: (a) $KMnO_4$, 1.8 M H_2SO_4 in aqueous 50% AcOH; (b) TFA/TIS/TFMSA (20/2.5/2.5); (c) Fmoc-OSu, DIEA, CH_3CN/H_2O (3/1).

Table 1 Dissociation constants K_D(nM), (specificity ratios),[a] IC_{50} (nM), and T_m[b] (1C) values of peptide/TAR complexes[c]

	K_D(TAR)		K_D(TAR_{ab})		K_D($IRES_{IIId}$)		IC_{50}		T_m
Tat(−1)	224	50	246	53	77	12 (2.9)[a]	40	5	70.6
Tat	112	33	114	35	37	7 (3.0)[a]	28	3	67.5
Tat(+1)	27	4	24	2	12	2 (2.2)[a]	18	4	63.0

[a] Defined as average K_D(TAR)/K_D($IRES_{IIID}$). [b] T_m of TAR alone is 58.5 1C. [c] All experiments are done in triplicate and the results are expressed as average SD.

(Table 1, K_D(TAR_{ab}) and K_D($IRES_{IIID}$), respectively). Regardless of the RNA sequence, Tat(+1) is always the most affine ligand and Tat(−1) the less one. None of the three peptides discriminates between TAR and TAR_{ab}, the K_D(TAR) and K_D(TAR_{ab}) values being very close in the three cases. On the other hand, the three peptides bind to $IRES_{IIID}$ two to three times stronger than to TAR. Calculating the K_D(TAR)/K_D($IRES_{IIId}$) ratio gives the specificity of the three peptides for $IRES_{IIID}$ over wild-type TAR RNA. Tat(+1), which has the longest side-chains, is the less specific ligand. These results are commented in the discussion part (see above).

Thermodynamic profiles. To gain further insight into the TAR binding modes of the three peptides, thermodynamic profiles associated with each peptide–TAR equilibrium were determined. Free energies of Gibbs (DG1) were first calculated at several temperatures (278–308 K) from the dissociation constants (DG1 = RT ln K_D). Enthalpy (DH1) and entropy changes (DS1) were then determined as described in materials and methods (see above). The results are summarized in Table 2.

Clearly, the change in the side-chain length of amino acid residues influences the mode of interaction. An increase in the chain length is associated with an increase both in enthalpy and entropy. Thus, while Tat(−1)/TAR and Tat/TAR complexes are strongly enthalpy driven and entropically disfavored, Tat(+1)/TAR one is both enthalpy and entropy driven, signifying that Tat(−1) is the tightest ligand and at the opposite, Tat(+1) the loosest one. An Enthalpy Entropy Compensation phenomenon occurs,[34] resulting in close but nevertheless distinct DG1 and K_D values for the three studied peptides (Fig. 2). DC_p values are negative but increase as the side-chain length increases.

FRET displacement assays

We assessed the ability of peptides Tat, Tat(+1) and Tat(−1) to displace a fluoresceinated $Tat_{48–57}$ peptide fragment from a preformed complex established with a Dabcyl-labeled $TAR_{18–44}$ fragment, *via* a FRET assay.[13,35] In the absence of ligands, the association of these two partners results in an efficient quenching of the dye. Upon addition of increasing amounts of the studied peptides, the fluorescence rises in all cases, demonstrating that all of them are able to displace the fluoresceinated Tat fragment from the complex. As shown in Table 1, IC_{50} values decrease from 40 to 18 nM as the chain length increases, Tat(+1) being the best inhibitor of the Tat/TAR interaction and Tat(−1) the weakest.

Table 2 Thermodynamic parameters[a] of peptide/TAR complexes[b]

DG1		DH1		TDS1		DC_p	
−38.0	2.4	−55.7	1.9	−17.7	0.5	−2.27	0.07
−39.7	1.7	−46.4	0.4	−6.75	0.03	−1.76	0.15
−43.1	1.3	−20.3	1.2	+22.8	0.2	−1.06	0.28

[a] DG1, DH1 and TDS1 are expressed in kJ mol^{-1}. DC_p is expressed in kJ mol^{-1} K^{-1}. [b] All experiments are done in triplicate and the results are expressed as average SD.

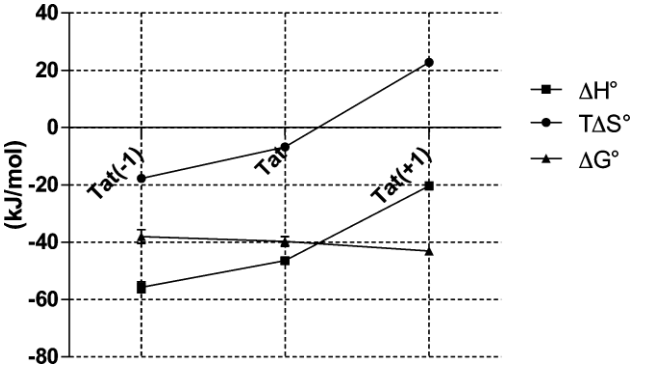


Fig. 2 DH1, TDS1 and DG1 values associated with the three peptide/TAR RNA equilibria. All experiments are done in triplicate and the results are expressed as average SD.

UV melting studies

To evaluate the ability of the peptides to stabilize the TAR RNA structure, thermal denaturation of TAR RNA and TAR RNA/ligand complexes was monitored using UV spectroscopy at 260 nm. In all cases, the TAR melting temperature (T_m = 58.5 1C) increases upon ligand binding at a 1 : 1 TAR/ligand ratio (Table 1 and Fig. 3), demonstrating a stabilizing effect on the RNA secondary structure. While Tat binding leads to a shift in T_m of 9 1C, Tat(−1) and Tat(+1) induced a shift of respectively 12.1 1C and 4.5 1C, suggesting that the tighter is the ligand, the stronger is the stabilization of the TAR structure. Thus, even if the TAR/Tat(+1) complex is the most stable from a thermodynamic point of view, Tat(+1) binding induces the weakest stabilization of the RNA structure.

Molecular modeling studies

Molecular Dynamics (MD) were performed to examine the formation of Tat derivatives/TAR complexes at an atomistic level and to analyze the effect of the side-chain length on the interaction mechanism. The preliminary results proved that docking procedures were not able to predict the preferred

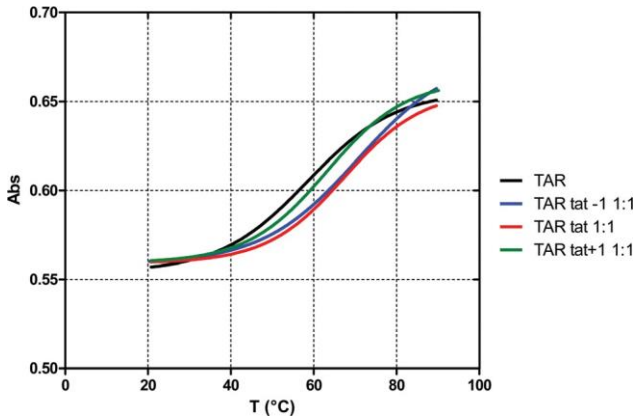


Fig. 3 UV melting curves at 260 nm of TAR RNA and its complex with Tat, Tat(+1) and Tat(−1).

interaction mechanism in the experimentally available Tat/TAR complex. Thus, MD starting conformations for Tat, Tat(−1) and Tat(+1) peptides were obtained by flexible alignment, considering as a reference the crystal structure of the HIV-1 TAR RNA bulge region, complexed with a b-hairpin cyclic peptide mimic of HIV-1 Tat protein (L-22), reported in the literature.[36]

MD simulations suggest significant differences in the peptide interaction mechanism, in good agreement with thermodynamic studies. Although all compounds tend to remain in the bulge region, they induce distinct conformational changes in the RNA structure (Fig. 4). The analysis of the full MD trajectory suggested that Tat(+1) – and, in a minor extent, Tat, – widen the RNA bulge in order to enhance their interaction patterns.

Conversely, Tat(−1) promotes the establishing of a hydrogen bond between C8 and C14 base pairs that induces a contraction of the bulge region contributing to the stabilization of the TAR structure. This conformation is long lasting due to the hydrophobic interaction between Tat(−1) and A19 that encloses the peptide into the bulge region (Fig. 5). These results suggest that Tat(−1) could strengthen the stabilization of the TAR secondary structure, as pointed out previously using UV melting studies.

A thorough analysis of MD trajectories reveals that the interaction mechanism of all peptides is mainly achieved by side-chain charge-based and hydrophobic interactions. Electrostatic interactions are primarily established due to the positive charges of arginine and lysine analogs.

The molecular modeling results agree with experimental data, considering a non-negligible entropy–enthalpy compensation for these systems. On the one hand, Tat(−1) shows abiding interactions with TAR RNA as a result of this confinement, might significantly contribute to D*H*1, but may also contribute to the stabilization of the RNA secondary structure. On the other hand, Tat(+1) would provide less interactions with TAR, but might produce more pronounced changes in the RNA conformation than Tat.

Antiviral activity

The antiviral activity of the three peptides was assayed on HeLa P4 cells expressing CD4 receptors and the *lacZ* gene under the control of the HIV-1 LTR (Fig. 6). Tat(−1) and Tat(+1) are stronger inhibitors than Tat, with IC_{50} values equal to 10 3 mM, 18 2 mM and 49 4 mM, respectively.

Discussion

To investigate the impact of the side-chain length on RNA/peptide interactions, we have synthesized two analogs of the basic region of the HIV-1 Tat protein, containing a-amino acid residues with side-chains either one methylene longer or shorter than the natural ones, and compared their interaction mode with TAR RNA. Varying the molecular flexibility of all side-chains at the same time and in the same degree modulates the strength of the interaction at the binding interface of the complex. The shorter the side-chain length, the more intimate are the contacts between the peptide and RNA. Nevertheless, the looser ligand (Tat(+1)) gives the most stable complex, in total agreement with theoretical predictions.[14] Thus, contrary to what one might think, more flexibility at the RNA–peptide

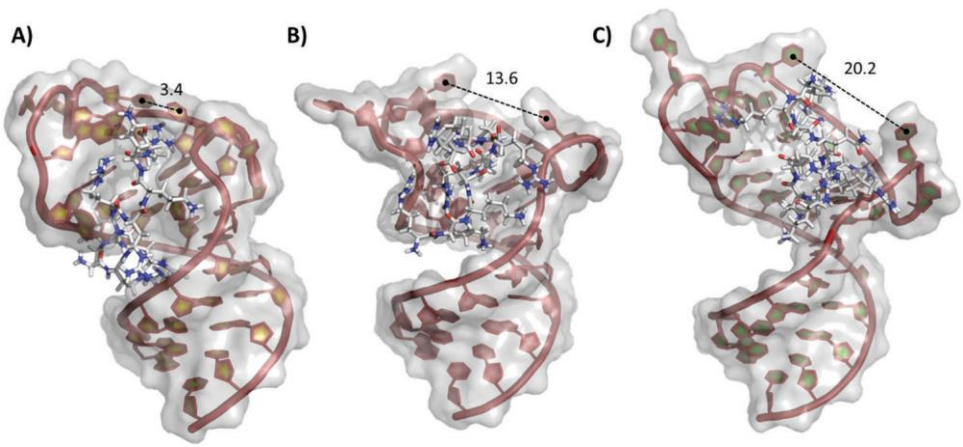


Fig. 4 Final MD frame representations of Tat(−1) (A), Tat (B) and Tat(+1) (C). The bulge closure distance was calculated as N3(C8)–N3(C14) for each system.

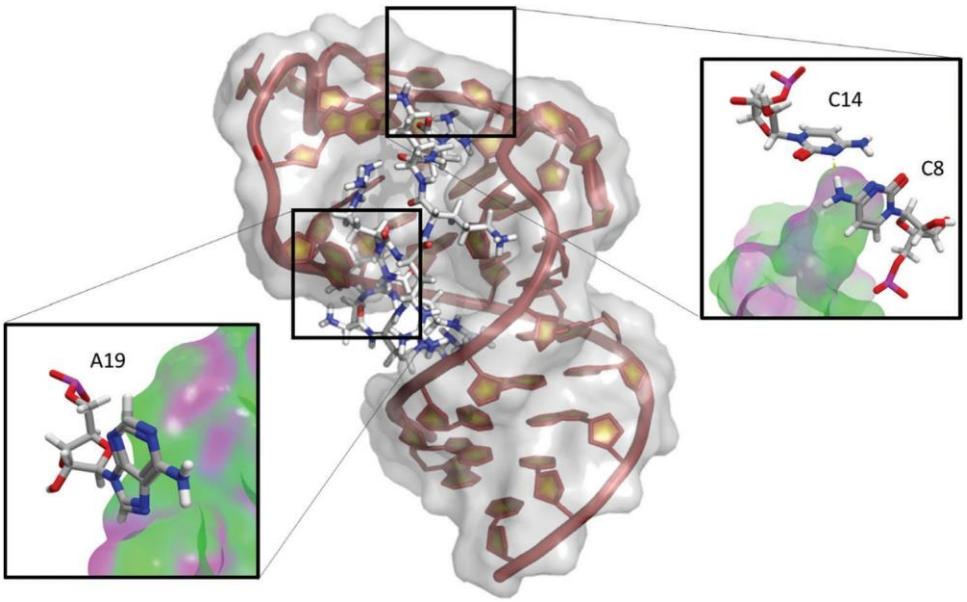


Fig. 5 Graphical representation of the predicted conformation of the Tat(−1)/TAR complex. The Tat(−1) structure was confined within the RNA due to the contraction of the bulge region produced by the C8–C14 hydrogen bond and the hydrophobic interaction with A19. Molecular surface for Tat(−1) is colored according to its lipophilicity (green: hydrophobic, pink: hydrophilic).

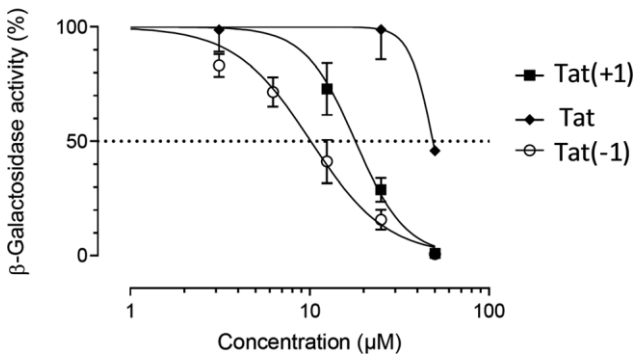


Fig. 6 Inhibition of b-galactosidase activity by Tat(−1), Tat and Tat(+1).

binding interface does not necessarily result in lower affinity interactions. The enhanced thermodynamic stability observed as the number of methylene groups/side-chain increases has an entropic origin. Indeed, adding one methylene group/side-chain in Tat(−1) then another one in Tat result in enthalpy losses (D(D*H*)) of 9.25 and 26.15 kJ mole^{-1} overbalanced by entropy gains (*T*D(D*S*)) of 10.95 and 29.55 kJ mole^{-1}, respectively. Considering the structural similarity of the ligands, it seems unlikely that these favorable entropic effects are linked to differences in ion release (the polyelectrolyte effect) or in the protonation state. They rather exhibit an increased hydrophobic effect[37,38] associated with the increase in ligand hydrophobicity (addition of methylene). Concerning the heat capacity changes, DC_p are negative in all cases, in line with other studies related to the binding of small molecules to nucleic acids.[39,40] This mainly reflects both a reduction in the non-polar surface accessible to the solvent upon peptide binding (hydrophobic effect), and a RNA conformational change.[37,41] However, following these criteria, the most hydrophobic peptide, Tat(+1), which induces the highest distortion of the RNA structure (as shown by molecular modeling studies), should display the lowest DC_p value whereas the less hydrophobic peptide, Tat(−1), should be associated with the highest one. Since the opposite is observed, other criteria than the hydrophobic effect and the conformational change have to be taken into account, such as differences in the ordering of water molecules[42] and/or in changes in RNA and/or ligand vibrational modes upon peptide binding.[37] Indeed, intuitively, water molecules trapped in a loose complex should be less ordered than in a tight one and the decrease in vibrational modes should be lower.

Regarding the binding specificity, the lack of specificity of the Tat basic peptide fragment for wild-type TAR over TAR$_{ab}$ has been previously reported[43,44] and the Tat(−1) and Tat(+1) derivatives show the same behavior. On the other hand, all the three peptides bind to the IRES IIId hairpin with higher affinities than to TAR RNA but is noteworthy that Tat(+1), which has the longest side-chains residues, is the less discriminating ligand, as was the case of Tat derivatives containing HArg comparatively to those containing Agb ones.[23,25] This could be a consequence of the higher conformational flexibility of the side chains, which could adopt a larger number of orientations suitable for binding different RNA structures[12,23,25]

Thermal melting studies reveal a stabilization effect upon binding of the three peptides but Tat(+1) induces the lowest stabilization and Tat(−1) the highest. These results point out that an increase in T_m should not be systematically linked to an increase in the affinity. In our case, the T_m increase is associated with a decrease in DH_1 values (C, Fig. 7), indicating that the melting temperature reflects the strength of the interaction at the binding interface but also the degree of compactness of

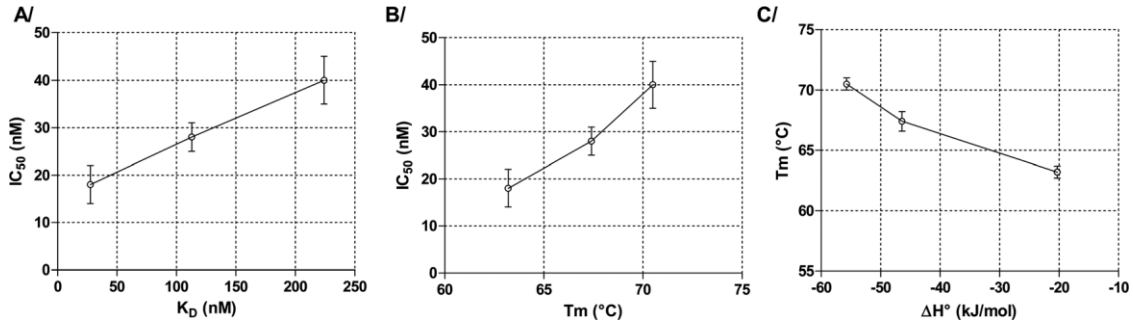


Fig. 7 Plots of: (A) IC_{50} against K_D; (B) IC_{50} against T_m; (C) T_m against DH_1. All experiments are done in triplicate and the results are expressed as average SD.

the TAR secondary structure at the bulge level. For Tat(−1), the high contraction at the bulge level is induced by two supplementary interactions as compared with Tat and Tat(+1), one intra-molecular C8–C14 hydrogen bond and one hydrophobic interaction of the ligand with A19 (Fig. 4), which may account for the strong negative DH_1 value obtained in this case.

The ability to displace the labeled-Tat fragment (IC_{50}) increases as both K_D and T_m decrease (A and B, Fig. 7). Interestingly, Tat(+1) is the strongest Tat competitor, although the RNA stabilization induced upon its binding is the weakest. Molecular modeling studies suggest that increasing the side-chain length of Tat leads to a higher distortion of TAR RNA at the bulge level (Fig. 4) and it is possible that this particular conformational change impedes the Tat/TAR recognition process. These results demonstrate how it is hazardous to predict the competitive behavior of a RNA ligand only on the basis of K_D and/or T_m values.

At last, Tat(−1) and Tat(+1) display respectively a 5- and 2.6-fold increase in antiviral activity in cell assays compared to the parent $Tat_{49–57}$ peptide. However, there is no trend between the ability to displace the Tat fragment (see IC_{50} values in Table 1) and the antiviral activity. Accordingly, these results likely reflect a combination of factors, including resistance to proteolytic degradation. Indeed, peptides containing non-canonical arginine residues (HArg and Agb) display a higher protease resistance, as compared to the natural sequences.[45] Another important factor to be taken into account is cellular uptake. Indeed, it has been shown that altering the length of all Arg side-chains in poly-arginine peptoid oligomers $[N\text{-arg}]_7$,[46] in the poly-arginine $[Arg]_7$ peptide[47] and in Tat-derivatives[23] has an unpredictable impact on cellular penetration, depending on the nature of the studied system: while lengthening all the side-chains in $[N\text{-arg}]_7$ and in $[Arg]_7$ increases the amount of cellular uptake, the Tat-derived peptide with the shortest side-chain (Agb) residue shows the most efficient uptake. In our case, it would be interesting to explore both the biological stability and the cellular uptake of Tat(+1) and Tat(−1) as compared to Tat, to analyse to what point these two parameters affect the antiviral activity.

To conclude, our investigations provide useful information regarding the impact of the side-chain length of a peptide ligand on the adaptive RNA binding mechanism, allowing us to highlight non-intuitive relationships between thermodynamic parameters and to validate theoretical models. Thus, varying identically the length of all side-chains of a natural peptide RNA ligand modulates the binding affinity, the strength of the interaction at the binding interface of the resulting peptide/RNA complex and the degree of compactness of the RNA at the bulge level. This results in modulating the conformational change in the RNA structure at the interaction site that in turn has an impact on the ability of the ligand to compete with the cognate RNA partner. Clearly, the length of chains connecting the central scaffold of a ligand to RNA recognition motives is an important determinant that should be considered when designing new RNA ligands as competitors of RNA complexes of biological interest, and modifying in the same way the length of all chains while maintaining the same interaction potential is one way to achieve this goal.

Acknowledgements

This work was supported by the CNRS (Centre National de la Recherche Scientifique) and UCA (Universit´e Cˆote d'Azur).

References

1 J. R. Thomas and P. J. Hergenrother, *Chem. Rev.*, 2008, 108, 1171–1224.
2 T. Hermann, *Wiley Interdiscip. Rev.: RNA*, 2016, 7, 726–743.
3 F. Aboul-ela, *Future Med. Chem.*, 2010, 2, 93–119.
4 J. A. Cruz and E. Westhof, *Cell*, 2009, 136, 604–609.
5 W. Hong, J. Zeng and J. Xie, *Acta Pharm. Sin. B*, 2014, 4, 258–265.
6S. Fulle and H. Gohlke, *J. Mol. Recognit.*, 2010, 23, 220–231.
7 T. Hermann, *Biochimie*, 2002, 84, 869–875.
8 N. Leulliot and G. Varani, *Biochemistry*, 2001, 40, 7947–7956.
9 Y. Savir and T. Tlusty, *PLoS One*, 2007, 2, e468.
10 J. R. Williamson, *Nat. Struct. Biol.*, 2000, 7, 834–837.
11 Y. Maeda, R. Iwata and T. Wada, *Bioorg. Med. Chem.*, 2013, 21, 1717–1723.
12 M. Murtola, S. Zaramella, E. Yeheskiely and R. Stromberg, *ChemBioChem*, 2010, 11, 2606–2612.

13 L. Pascale, S. Azoulay, G. A. Di, L. Zenacker, M. Gaysinski, P. Clayette and N. Patino, *Nucleic Acids Res.*, 2013, 41, 5851–5863.
14 C. Forrey, J. F. Douglas and M. K. Gilson, *Soft Matter*, 2012, 8, 6385–6392.
15 A. Re, T. Joshi, E. Kulberkyte, Q. Morris and C. T. Workman, *Methods Mol. Biol.*, 2014, 1097, 491–521.
16 F. A. Abulwerdi, M. D. Shortridge, J. Sztuba-Solinska, R. Wilson, S. F. Le Grice, G. Varani and J. S. Schneekloth, Jr., *J. Med. Chem.*, 2016, 59, 11148–11160.
17 A. Blond, E. Ennifar, C. Tisne and L. Micouin, *ChemMedChem*, 2014, 9, 1982–1996.
18 J. Pai, T. Yoon, N. D. Kim, I. S. Lee, J. Yu and I. Shin, *J. Am. Chem. Soc.*, 2012, 134, 19287–19296.
19 S. N. Richter and G. Palu, *Curr. Med. Chem.*, 2006, 13, 1305–1315.
20 S. Bannwarth and A. Gatignol, *Curr. HIV Res.*, 2005, 3, 61–71.
21 K. A. Roebuck and M. Saifuddin, *Gene Expression*, 1999, 8, 67–84.
22 G. Mousseau, S. Mediouni and S. T. Valente, *Curr. Top. Microbiol. Immunol.*, 2015, 389, 121–145.
23 C. H. Wu, Y. P. Chen, C. Y. Mou and R. P. Cheng, *Amino Acids*, 2013, 44, 473–480.
24 C. H. Wu, Y. P. Chen, S. L. Liu, F. C. Chien, C. Y. Mou and R. P. Cheng, *Org. Biomol. Chem.*, 2015, 13, 11096–11104.
25 S. J. Lee, S. Hyun, J. S. Kieft and J. Yu, *J. Am. Chem. Soc.*, 2009, 131, 2224–2230.
26 D. Rocancourt, C. Bonnerot, H. Jouin, M. Emerman and J. F. Nicolas, *J. Virol.*, 1990, 64, 2660–2668.
27 R. Behrendt, P. White and J. Offer, *J. Pept. Sci.*, 2016, 22, 4–27.
28 T. P. Boyle, J. B. Bremner, J. A. Coates, P. A. Keller and S. G. Pyne, *Tetrahedron*, 2005, 61, 7271–7276.
29 K. Wisniewski and A. S. Kolodziejczyk, *Org. Prep. Proced. Int.*, 1997, 29, 338–341.
30 L. A. Carpino and G. Y. Han, *J. Am. Chem. Soc.*, 1970, 92, 5748–5749.
31 T. Maegawa, Y. Fujiwara, T. Ikawa, H. Hisashi, Y. Monguchi and H. Sajiki, *Amino Acids*, 2009, 36, 493–499.
32 J. Martinez, J. C. Tolle and M. Bodanszky, *J. Org. Chem.*, 1979, 44, 3596–3598.
33 R. Klinck, E. Westhof, S. Walker, M. Afshar, A. Collier and F. Boul-Ela, *RNA*, 2000, 6, 1423–1431.
34 J. D. Chodera and D. L. Mobley, *Annu. Rev. Biophys.*, 2013, 42, 121–142.
35 L. Pascale, A. L. Gonzalez, G. A. Di, M. Gaysinski, C. J. Teixido, R. E. Tejedor, S. Azoulay and N. Patino, *J. Biomol. Struct. Dyn.*, 2016, 34, 2327–2338.
36 A. Davidson, T. C. Leeper, Z. Athanassiou, K. Patora-Komisarska, J. Karn, J. A. Robinson and G. Varani, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, 106, 11931–11936.
37 J. M. Sturtevant, *Proc. Natl. Acad. Sci. U. S. A.*, 1977, 74, 2236–2240.
38 D. H. Williams and B. Bardsley, *Perspect. Drug Discovery Des.*, 1999, 17, 43–59.
39 M. M. Islam, P. Pandya, S. Kumar and G. S. Kumar, *Mol. BioSyst.*, 2009, 5, 244–254.
40 R. Stolarski, *Acta Biochim. Pol.*, 2003, 50, 297–318.
41 D. S. Pilch, M. Kaul, C. M. Barbieri and J. E. Kerrigan, *Biopolymers*, 2003, 70, 58–79.
42 P. W. Snyder, J. Mecinovic, D. T. Moustakas, S. W. Thomas, III, M. Harder, E. T. Mack, M. R. Lockett, A. Heroux, W. Sherman and G. M. Whitesides, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, 108, 17889–17894.
43 M. A. Gelman, S. Richter, H. Cao, N. Umezawa, S. H. Gellman and T. M. Rana, *Org. Lett.*, 2003, 5, 3563–3565.
44 J. Kamine, P. Loewenstein and M. Green, *Virology*, 1991, 182, 570–577.
45 J. Izdebski, E. Witkowska, D. Kunce, A. Orlowska, B. Baranowska and E. Wolinska-Witort, *J. Pept. Sci.*, 2004, 10, 524–529.
46 P. A. Wender, D. J. Mitchell, K. Pattabiraman, E. T. Pelkey, L. Steinman and J. B. Rothbard, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, 97, 13003–13008.
47 D. J. Mitchell, D. T. Kim, L. Steinman, C. G. Fathman and J. B. Rothbard, *J. Pept. Res.*, 2000, 56, 318–325.