RNA-Binding Agents Hot Paper

DOI: 10.1002/anie.201306665

Structure-Based Design of an RNA-Binding *p*-Terphenylene Scaffold that Inhibits HIV-1 Rev Protein Function**

Luis González-Bulnes, Ignacio Ibáñez, Luis M. Bedoya, Manuela Beltrán, Silvia Catalán, José Alcamí, Santos Fustero,* and José Gallego**

Numerous antibiotics bind to ribosomal RNA, and many functional RNA motifs have considerable therapeutic potential. However, the development of small RNA-binding agents has been hampered by the difficulties posed by these structures, which have limited physicochemical diversity and are often flexible.[1] In order for this approach to be successful, it is essential to identify new chemical scaffolds recognizing RNA.

The Rev response element (RRE) is a strongly conserved 350-nucleotide structure located in the *env* gene of human immunodeficiency virus type-1 (HIV-1) RNA. Within subdomain IIB of the RRE, the unusually widened major groove of a 4:6 internal loop forms a high-affinity complex[2] with the arginine-rich α-helix of Rev, a virally encoded 116 amino acid protein that adopts a helix-turn-helix conformation[3] (Figure 1). This interaction between internal loop IIB and the RNA-binding α-helix of Rev ($Rev_{34\text{-}50}$) is essential for virus viability because it triggers a cascade of events that allow the transport of unspliced or incompletely spliced viral RNA molecules into the cytoplasm of the infected cell in the late phase of the viral infectious cycle. These events include the cooperative incorporation of additional Rev molecules into the complex through interactions with further sites on the RRE and protein–protein contacts,[4] and the tethering of the RRE–Rev ribonucleoprotein to the Crm1 host export factor. In addition to RNA nuclear export, Rev has been shown to enhance translation and packaging[5] and to control the nucleocytoplasmic shuttling of the HIV-1 integrase.[6] Clearly, this protein represents a pivotal target for HIV-1 therapy, but to date, the development of Rev-based inhibitors has remained an elusive goal. Herein we report the structure-based design of new RNA-binding *p*-terphenyl Rev mimics that inhibit RRE-Rev function and HIV-1 replication.

We aimed to generate organic ligands that mimic the three-dimensional distribution of the side chains of $Rev_{34\text{-}50}$ complexed with internal loop IIB of the RRE[2] (Figure 1a). Some reports had shown that tris-3,2′,2″-substituted *p*-ter-

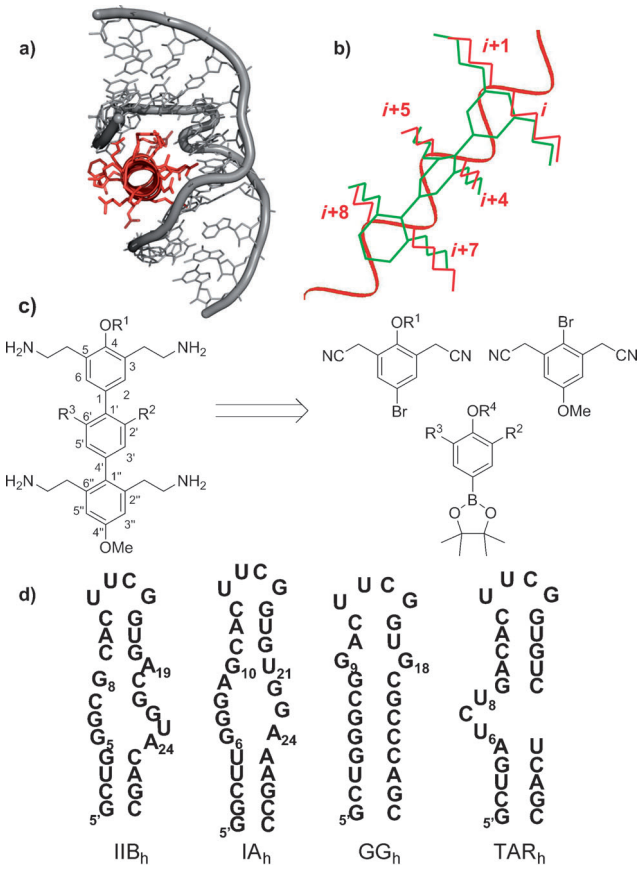

Figure 1. Structure-based design of RNA-binding *p*-terphenyls. a) View of the complex formed by internal loop IIB of the RRE and $Rev_{34\text{-}50}$. The α-helix (red) is deeply embedded in its RNA receptor (gray). b) Representation of a bilaterally substituted *p*-terphenyl molecule (green) superposed on an α-helix (red). c) Chemical structure and synthesis strategy of a 3,5,2′,6′,2″,6″-substituted *p*-terphenyl molecule. d) Secondary structures of the RNA oligonucleotides utilized in this study: hairpins IIB_h and IA_h containing internal loops IIB and IA of the RRE, and control hairpins GG_h and TAR_h.

[*] L. González-Bulnes,[+] Dr. J. Gallego
Universidad Católica de Valencia
C/Quevedo 2, 46001 Valencia (Spain)
E-mail: jose.gallego@ucv.es
I. Ibáñez,[+] Dr. S. Catalán, Prof. S. Fustero
Universidad de Valencia
Avda. V. A. Estellés s/n, 46100 Burjassot (Spain)
and
Centro de Investigación Príncipe Felipe
Avda. Autopista del Saler 16, 46012 Valencia (Spain)
E-mail: santos.fustero@uv.es
Dr. L. M. Bedoya, M. Beltrán, Dr. J. Alcamí
Instituto de Salud Carlos III
Carretera Majadahonda-Pozuelo km 2, 28220 Majadahonda (Spain)
E-mail: ppalcami@isciii.es
[+] These authors contributed equally to this work.

[**] We received support from MEC (BFU2012-30770, CTQ2010-19774) and ISCIII (RD12/0017, PI 12/0506) of Spain, Generalitat Valenciana (PROMETEO 2010/061, AP-225/2011), and Universidad Católica de Valencia, and thank A. Velázquez for help with ITC, and C. Cabrera for providing plasmids.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201306665.

phenyl molecules could mimic one face of an α-helical peptide by adopting a staggered conformation that reproduces the angular orientation of three α-helix side chains.[7] We envisioned that the introduction of substituents on both sides of a *p*-terphenyl scaffold would ensure a 360° side-chain projection in space similar to that observed in the IIB–Rev_{34-50} complex, where two thirds of the α-helix are surrounded by RNA. After docking different biphenyl and terphenyl ligands into the IIB structure, the best results were obtained for 3,5,2′,6′,2″,6″-substituted *p*-terphenyls (Figure 1), the binding poses of which approximately reproduced the orientation of Rev_{34-50} in its complex with IIB. Subsequent calculations indicated that synthetically accessible 2-aminoethyl lateral substituents would be well suited to interact with the RNA sugar–phosphate backbone.

The syntheses were based on sequential palladium-based Suzuki couplings[8] of aryl halides and aryl triflates with aryl boronic esters (Figure 1c). We first generated a group of tetrakis(2-aminoethyl) biphenyls where the side chains occupied bilateral 3,5 and 2′,6′ (**1**); 3,5 and 3′,5′ (**2**); or 2,6 and 2′,6′ (**3**) positions. Our first terphenyl molecule (**4b**) contained four bilateral 2-aminoethyl side chains in positions 3,5,2″,6″ but lacked substituents in the central ring. Subsequent 3,5,2″,6″ tetrakis(2-aminoethyl) terphenyls contained a single methoxy (**5a**, **5b**, **5c**, and **5d**) or ethyl (**5e** and **5f**) group in the 2′ position of the central benzene. The last terphenyl series contained two methyl (**6a** and **6b**) or ethyl (**6c** and **6d**) groups in positions 2′ and 6′ (Figure 2). These molecules had bilateral groups in all three rings and were predicted to be the best Rev_{34-50} mimics.

To examine the strength of the interactions and the location of the binding sites, we used NMR spectroscopy to monitor the titration of an RNA hairpin containing internal loop IIB (IIB_h; Figure 1d) with the biphenyl and terphenyl molecules or the RRE-binding antibiotic neomycin B, which was used as a reference.[9] The biphenyl **1** required a high ligand/IIB_h molar ratio (6:1) to produce detectable changes in the IIB_h TOCSY spectrum, whereas **2**, **3**, and neomycin B induced chemical shift perturbations in stem nucleotides outside the loop (Figure S1 in the Supporting Information). By contrast, terphenyl molecules containing substituents in the central benzene induced chemical shift changes in the internal loop and adjacent nucleotides only, and these shifts were apparent at lower ratios (1:1 and 2:1; Figure 3a and Figure S1 in the Supporting Information). The terphenyls that gave rise to the greatest RNA chemical shift variations and sharper complex resonances were **6b** and **6d**, which contain two bilateral methyl or ethyl groups, respectively, in positions 2′ and 6′. Spectral line shape analyses of the complexes of IIB_h with **6b** and **5b** indicated binding of two ligand molecules to internal loop IIB, with equilibrium dissociation constants K_d of 11 and 25 μM for **6b**, and 25 and 33 μM for **5b** (Figure S2).

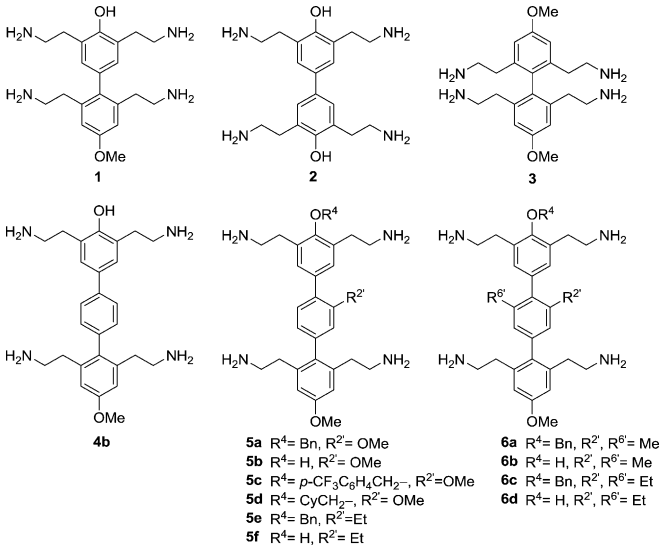

Figure 2. Chemical structures of the bilaterally substituted *p*-biphenyl and *p*-terphenyl compounds.

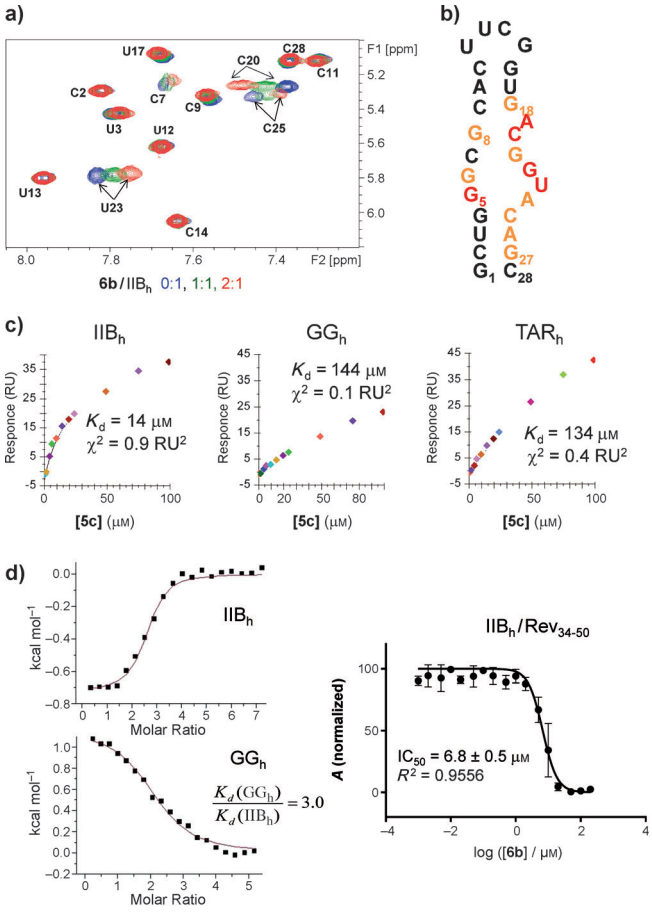

Figure 3. Loop IIB recognition and inhibition of the IIB–Rev_{34-50} interaction by terphenyl compounds. a) Titration of IIB_h with **6b** monitored by NMR spectroscopy. The H5–H6 region of the TOCSY spectrum of unbound IIB_h (blue) is superposed on the spectra of complexes with increasing **6b**/IIB_h molar ratios. b) Map of the **6b** binding site in the IIB_h hairpin. The nucleotides with aromatic or H1′ protons undergoing chemical shift variations upon the addition of two equivalents of **6b** are highlighted in orange and red ($\Delta\delta \geq 0.05$ and 0.1 ppm, respectively). c) SPR steady-state equilibrium binding curves for the interaction between **5c** and IIB_h, GG_h, and TAR_h. d) Titration curves of IIB_h and GG_h with **6b**, as monitored by ITC. e) Inhibition of the IIB_h–Rev_{34-50} interaction by **6b**, as measured by FP.

The biphenyl and terphenyl interactions were further characterized with surface plasmon resonance (SPR) and isothermal titration calorimetry (ITC). To evaluate specificity, the SPR experiments employed two control GG_h and TAR_h hairpins, in which the IIB loops are substituted with a G:G opposition and the HIV-1 Tat-binding UCU bulge,[10] respectively (Figure 1d). The SPR methodology allowed us to study the 2′-methoxy terphenyl subset as well as **1** and **3**. The IIB_h–terphenyl binding curves were best fit with a two-site model, comprising a higher-affinity event involving the binding of one or two ligand molecules to loop IIB (see below), and lower-affinity association of additional molecules to the hairpin. We obtained a K_d of 13.0 μM for the higher-affinity interaction between IIB_h and **5b**, together with $K_d(GG_h)/K_d(IIB_h)$ and $K_d(TAR_h)/K_d(IIB_h)$ specificities of 1.9 and 6.0, respectively. **5c** had specificity ratios of 10.3 and 9.6, respectively, comparable to those observed for $Rev_{34\text{-}50}$ with the same methodology (14.5 and 4.4; Figure 3c, and Figure S3 and Table S1). **5d**, **1**, and **3** bound to IIB_h with higher K_d values.

ITC analyses of IIB_h–terphenyl interactions also revealed a higher-affinity transition followed by lower-affinity binding at high ligand/IIB_h ratios. The titrations focused on the higher-affinity transition, which involved the association of approximately two terphenyl molecules with loop IIB. The K_d values obtained with this approximation were useful for qualitatively ranking the ligands and evaluating specificities (Figure 3d, and Figure S4 and Table S2). The compounds with the greatest IIB_h affinities were **6b** and **6d**, which had $K_d(GG_h)/K_d(IIB_h)$ specificities of 3.0 and 4.4, respectively. The affinities were lower for **5b** and **6a**. Compound **2** interacted with IIB_h with lower affinity, no specificity and higher stoichiometry. All of these results are consistent with the NMR observations described above.

The complexes of IIB_h with **6b** and **6d** were further studied by NMR spectroscopy, and were very similar. Examination of the IIB_h chemical shift perturbations allowed us to map the residues forming the binding site of the ligands. These spanned all of the GGCG:ACGGUA loop and several adjacent residues in the lower stem (Figure 3b). The spectra for both the isolated and ligand-bound IIB_h showed that G22 and A24 were stacked on top of each other and U23 had an extrahelical location. However, the intensity of the intra-residue H8–H1′ NOE of G22 was significantly reduced upon **6b** and **6d** association (Figure S5a), thus indicating a switch from the *syn* conformation adopted by this nucleotide in the unbound loop[11] to an *anti* conformation. The binding of **6b** and **6d** weakened the sequential connectivities of A19 (Figure S5a), thus favoring an extrahelical conformation for this nucleotide. These conformational changes involving G22 and A19 were also observed upon $Rev_{34\text{-}50}$ binding to subdomain IIB.[11]

Intermolecular RNA–terphenyl NOEs (described in Figure S5b) demonstrated that one of the two terphenyl molecules forming the complex lies diagonally across the major groove of the IIB loop, with its first benzene ring located close to G6 in the GGCG strand of the loop and its third benzene ring located in proximity to A19 in the opposite $A_{19}CGGUA_{24}$ loop strand (Figure 4a).

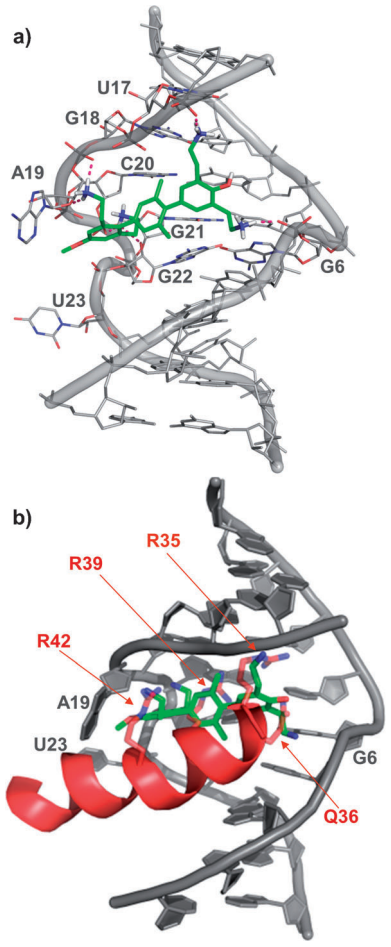

Figure 4. Structure of the loop IIB–terphenyl complexes. a) View of the best-scored complex between loop IIB and **6b**, built by NOE-restrained docking from the structure of the RRE–Rev complex (PDB ID:1ETG).[2] b) Overlay of this model with the $Rev_{34\text{-}50}$ helix (red ribbon representation) present in the 1ETG structure. The amino acids matching the position of the terphenyl 2-aminoethyl substituents are explicitly represented.

We next tested whether the biphenyl and terphenyl compounds inhibited the IIB_h–$Rev_{34\text{-}50}$ interaction by using an assay based on fluorescence polarization (FP). **6b** turned out to be the most potent inhibitor, with an IC_{50} value of 6.8 μM, followed by **4b** and other terphenyl molecules containing methyl or ethyl groups in the central benzene (Figure 3e, and Figure S6 and Table S3). The 2′-methoxy terphenyl subset exhibited higher IC_{50} values consistent with with the detection of lower IIB_h affinity by ITC and NMR.

When we checked whether the compounds were able to block HIV-1 replication in cell cultures, all terphenyls containing methyl or ethyl groups in the central benzene ring were active and **6b**, with an EC_{50} value of 3.4 μM, was the most potent inhibitor. The remaining ligands and neomycin B were not active or exhibited much weaker activities at the assay concentrations, and none of the compounds showed cellular toxicity. We also carried out an assay based on transfecting a full-length replication-competent HIV-1 clone, whereby early steps of infection are bypassed and only post-integration events of the viral infectious cycle occur. **6b** was

again the most potent inhibitor, followed by **5e** and **6a** (Figure 5, and Figure S7 and Table S4). The EC_{50} values of these terphenyls were similar to those obtained from the infection experiment, thus indicating that their main target was contained in the transcriptional or post-transcriptional steps of the viral infectious cycle, as is the case for the RRE–Rev system.

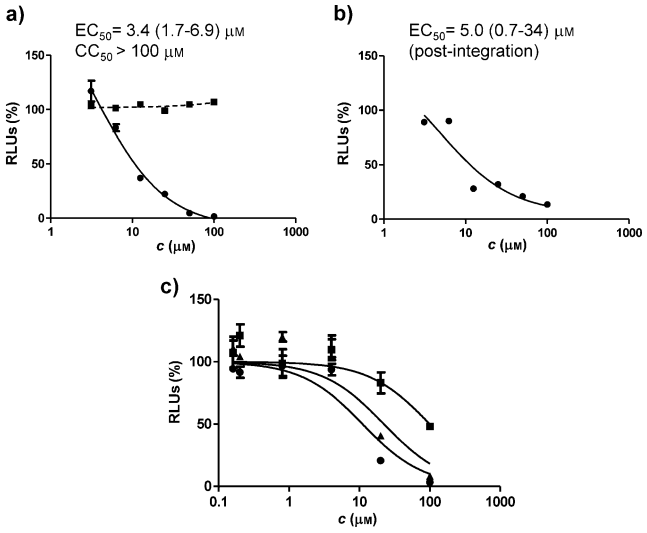

Figure 5. Results of cellular assays with terphenyl **6b**. In all cases, results are expressed as percentage of relative luminescence units (RLUs), where 100% is the level of viral replication obtained in the presence of the vehicle used to dissolve the compounds. a) Antiviral activity in HIV-1 infection assays (EC_{50}) and cellular toxicity (CC_{50}). ----■---- viability; ——●—— infection. b) Antiviral activity in HIV-1 transfection assays. c) Inhibition of Rev-mediated transport of RRE-containing RNA into the cytoplasm. Inhibition by **6b** was tested at three different concentrations of the plasmid encoding Rev (pCMV-Rev): ——■—— 500 ng/well ($50 < IC_{50} < 100$ μM); ——▲—— 200 ng/well ($IC_{50} = 21.4(\pm 6.3)$ μM); ——●—— 20 ng/well ($IC_{50} = 10.4(\pm 2.4)$ μM).

To test whether **6b** acted on the complete RRE–Rev system in cells, we used an assay based on transfecting plasmids encoding Rev and an RRE–luciferase reporter system. The results indicated that **6b** inhibited Rev-mediated transport of RRE-containing RNA from the nucleus to the cytoplasm. The IC_{50} values ranged from 10.4 to 21.4 μM and depended on the concentration of Rev-encoding plasmid used in the assay (Figure 5c). This clearly established cellular inhibition of RRE–Rev ribonucleoprotein function. In this context, additional Rev monomers bind to different sites in the 350-nucleotide RRE structure in addition to loop IIB. One of the additional Rev binding sites in the RRE, a purine-rich internal loop located in subdomain IA (Figure 1d), has recently been identified.[4] By using NMR spectroscopy, we found that **6b** specifically recognized internal loop nucleotides of the IA_h hairpin at low ligand/RNA ratios (Figure S8).

Since an effect on HIV-1 transcription is also possible, we studied the interaction between the terphenyl molecules and DNA. SPR and NMR experiments detected DNA association through the minor groove (Figure S9). DNA binding may contribute to the observed antiviral activity, but the available data are consistent with an effect based on RRE–Rev inhibition, as explained in Figure S9.

In conclusion, the data obtained after synthesis and evaluation of a set of biphenyl and terphenyl molecules indicate that we have rationally designed a novel RNA-binding bilaterally substituted *p*-terphenylene scaffold that mimics the $Rev_{34\text{-}50}$ helix of Rev and is able to inhibit RRE–Rev function and HIV-1 replication. This is important, as most of the small RNA-binding agents described to date are related to peptides or antibiotics, or were discovered by screening.[1]

In agreement with the computational predictions, NMR analyses indicated that **6b** and **6d** bound to the major groove of loop IIB with their side chains projected in a wide spatial angle, thus occupying the binding site of the N-terminal segment of $Rev_{34\text{-}50}$ (Figure 4). The 3,5-ethylene amino groups of the first benzene ring of the ligands contact the phosphate groups of G6 in one strand and U17 or G18 in the opposite strand. This same two-strand interaction is established by Rev.[2] The third benzene is inserted deep into the pocket formed by the S-turn G21 and G22 residues and extrahelical A19, and its 2″,6″-ethylene amino groups contact the unusually bent sugar–phosphate backbone of these nucleotides, where several phosphates are close to each other and are therefore stabilized by the positively charged amino groups of the ligands. In this respect, both terphenyls induced conformational changes in the A19 and G22 nucleotides strikingly similar to those brought about by Rev binding.

In agreement with the models, loop IIB affinity and specificity increased for 2'- and 2',6'-substituted terphenyls relative to **1**, **2**, **3**, and **4b**, thus indicating that both an appropriately staggered terphenyl conformation and an adequate spacing of the two pairs of 2-aminoethyls were important for the interaction with loop IIB. The terphenyl **6b**, together with its 2′,6′-diethyl analogue **6d**, exhibited the greatest affinity for IIB_h. FP experiments indicated that terphenyls containing 2'- or 2',6'-methyl or ethyl groups displaced $Rev_{34\text{-}50}$ from the IIB loop more efficiently than those with a 2′-methoxy group, and **6b** was the most potent inhibitor. In agreement with these data, methyl- and ethyl-substituted terphenyls were able to inhibit HIV-1 replication at post-integration steps of the viral infectious cycle, and **6b** was the most active molecule. In a cellular assay, this terphenyl was able to block the complete RRE–Rev system with low micromolar IC_{50} values similar to those measured for HIV-1 inhibition and $Rev_{34\text{-}50}$ displacement from loop IIB.

Altogether, the compiled data provide strong evidence that bilaterally substituted terphenyls efficiently mimic the RNA-binding α-helix of Rev and provide clues on how these small molecules may accomplish the blocking of RRE–Rev function and HIV-1 replication in a cellular context. The RNA specificity of the terphenyls is in the same range as that measured for $Rev_{34\text{-}50}$ itself, and the compounds bind to at least two specific Rev sites within the RRE; loops IIB and IA. By mimicking Rev, the compounds likely compete with the protein by binding to several sites in the RRE, thereby interfering with the assembly of Rev monomers on the RRE. In addition, we have detected an interaction with the DNA

minor groove that may contribute to the observed antiviral effect. The *p*-terphenylene scaffold offers many opportunities for improvement, and all of these issues are under investigation.

Received: July 30, 2013
Revised: September 25, 2013
Published online: November 8, 2013

Keywords: drug design · HIV · RNA · small-molecule inhibitors · terphenyls

[1] M. Mayer, T. L. James, *Methods Enzymol.* **2005**, *394*, 571 – 587; J. R. Thomas, P. J. Hergenrother, *Chem. Rev.* **2008**, *108*, 1171 – 1224; L. Guan, M. D. Disney, *ACS Chem. Biol.* **2012**, *7*, 73 – 86.
[2] J. L. Battiste, H. Mao, N. S. Rao, R. Tan, D. R. Muhandiram, L. E. Kay, A. D. Frankel, J. R. Williamson, *Science* **1996**, *273*, 1547 – 1551.
[3] M. D. Daugherty, B. Liu, A. D. Frankel, *Nat. Struct. Mol. Biol.* **2010**, *17*, 1337 – 1342; M. A. DiMattia, N. R. Watts, S. J. Stahl, C. Rader, P. T. Wingfield, D. I. Stuart, A. C. Steven, J. M. Grimes, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 5810 – 5814.
[4] M. D. Daugherty, I. D'Orso, A. D. Frankel, *Mol. Cell* **2008**, *31*, 824 – 834.
[5] H. C. Groom, E. C. Anderson, A. M. Lever, *J. Gen. Virol.* **2009**, *90*, 1303 – 1318; M. Blissenbach, B. Grewe, B. Hoffmann, S. Brandt, K. Uberla, *J. Virol.* **2010**, *84*, 6598 – 6604.
[6] A. Levin, Z. Hayouka, A. Friedler, A. Loyter, *Nucleus* **2010**, *1*, 190 – 201.
[7] B. P. Orner, J. T. Ernst, A. D. Hamilton, *J. Am. Chem. Soc.* **2001**, *123*, 5382 – 5383; H. Yin, G. I. Lee, H. S. Park, G. A. Payne, J. M. Rodriguez, S. M. Sebti, A. D. Hamilton, *Angew. Chem.* **2005**, *117*, 2764 – 2767; *Angew. Chem. Int. Ed.* **2005**, *44*, 2704 – 2707.
[8] A. Suzuki, *Chem. Commun.* **2005**, 4759 – 4763.
[9] M. L. Zapp, S. Stern, M. R. Green, *Cell* **1993**, *74*, 969 – 978; M. Hendrix, E. S. Priestley, G. F. Joyce, C. H. Wong, *J. Am. Chem. Soc.* **1997**, *119*, 3641 – 3648.
[10] C. Dingwall, I. Ernberg, M. J. Gait, S. M. Green, S. Heaphy, J. Karn, A. D. Lowe, M. Singh, M. A. Skinner, *EMBO J.* **1990**, *9*, 4145 – 4153.
[11] R. D. Peterson, J. Feigon, *J. Mol. Biol.* **1996**, *264*, 863 – 877.

