Pergamon

Tetrahedron Letters 43 (2002) 1641–1644

TETRAHEDRON LETTERS

Liquid-phase combinatorial synthesis of diPNA-arginine conjugates

Mary-Lorène Leroux, Christophe Di Giorgio, Nadia Patino and Roger Condom*

Laboratoire de Chimie Bio-Organique, UMR 6001, Université de Nice Sophia-Antipolis, F-06108 Nice Cedex 2, France

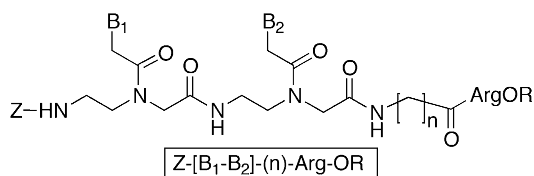
Received 4 December 2001; revised 7 January 2002; accepted 8 January 2002

Abstract—Considering the promising anti HIV activity displayed by some diPNA-arginine conjugates, a library has been generated to determine the target(s) and mode(s) of action of these presumed multi targets drugs and to optimize the antiviral properties of lead compounds. This library has been prepared using a combinatorial liquid-phase strategy, involving easily available *N*-protected PNA dimeric backbones as building blocks. © 2002 Elsevier Science Ltd. All rights reserved.

The generation of combinatorial chemical libraries has emerged as a highly efficient method for the identification of novel chemical leads or for the optimization of a promising lead compound.[1]

In this context, we recently reported[2] a new anti-HIV family constituted by a Polyamide Nucleic Acid[3] (PNA) dimer moiety bearing the adenine–guanine or the guanine–adenine base-pair, tethered to an arginine residue through a linker of variable length (Fig. 1).

Several of these PNA-arginine conjugates were indeed found to display a micromolar anti-HIV activity (IC_{50} = 1 μM) without any cellular toxicity (CC_{50} >200 μM). However, these compounds seemed to act by at least two different modes of action. To elucidate the target(s) and mode(s) of action of these presumed multi-targeted drugs, to establish structure–activity relationships and to optimize the antiviral activity of lead compounds, we planned to elaborate an extended library of these compounds.



(B_1, B_2)= (adenine, guanine) or (guanine, adenine)
n= 2, 5, 7
R= CH_3 or H; Z= benzyloxycarbonyl

Figure 1. diPNA-Arginine conjugate of HIV-1.

Combinatorial syntheses of PNA libraries have been described using solid-phase procedures.[4–7] However, in our case, a liquid-phase combinatorial approach seemed better adapted, considering the small size of the conjugates and the large amounts (multi-milligrams) required for extensive and broad screening purposes. We report here the combinatorial synthesis of a diPNA-arginine conjugates library, following an easy and rapid liquid-phase FPB (Fully Protected Backbone) strategy.

Molecular diversity for the family of the diPNA-arginine conjugates shown in Fig. 1 can be generated by varying their structural elements such as the nature of the nucleobases, the length of the spacer linking the arginine residue to the PNA dimer backbone, and the nature of *C*-extremity of arginine (R). Therefore, we envisaged to prepare such a library using (i) the five natural nucleobases, adenine (A), cytosine (C), guanine (G), thymine (T), uracile (U), and the universal base analogue 5-nitroindole[8] (In), (ii) a spacer that would be composed of 2, 3, 4 and 5 methylenes, and (iii) the arginine residue under its acid and methyl ester forms. The complete size of this library will be of 288 components (Fig. 2). Indeed, there are six^2 nucleobase combinations for PNA dimer conjugates, which further display two different arginine *C*-extremities (R = H and Me.). If we consider the four linkers, we get 288 structures.

To anticipate deconvolution problems, the main pool is partitioned into four sublibraries with respect to each of the four-spacer lengths. Each of these sublibraries is

* Corresponding author. Tel.: +3-349-207-6152; fax: +3-349-207-6151; e-mail: condom@unice.fr

0040-4039/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0040-4039(02)00067-9

split in two groups, in respect to the *C*-extremity of the arginine residue (acid or ester). Each of them is constituted by six batches of six compounds. Moreover, each batch contains a well-identified PNA at the second position (Fig. 2).

To generate this library, we applied a strategy, named Fully Protected Backbone (FPB), that was recently developed for the synthesis of linear[9] and cyclic[10] polyPNA.

Starting from fully orthogonally protected derivatives of type **1**, this strategy consists of the selective deprotection of the P_1-protected amino function, which is then condensed stoichiometrically with an equimolar mixture of the six B_{1-6} nucleobase acetic acid units. After subsequent release of the second P_2 amino protecting group, the B_i unit chosen among the six B_{1-6} nucleobase acetic acid units is then introduced. A first sublibrary of diPNA-arginine conjugates is obtained after deprotection of the guanidinium group of arginine and of nucleobases (ester groups) followed by a subsequent hydrolysis (acid groups) (Fig. 3).

As an illustration of this original method, we describe herein the preparation of the sublibrary deriving from the aminobutanoic acid ($n=3$) spacer. The key orthogonally fully protected di(aminoethylglycinamide) backbone **1** (Scheme 1) was prepared by condensing the dimeric moiety **2** and aminobutanoyl arginine methyl ester **3** by means of DCC/HOSu, in 69% yield. The arginine-based linker **3** was prepared from commercial N_G-(Pmc)-arginine and *N-Z* aminobutanoic acid, following a procedure previously described.[2]

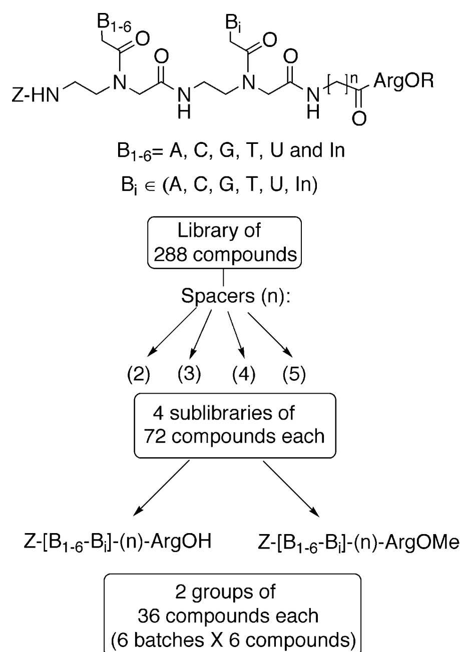


Figure 2. Composition and subdivision of the libraries of diPNA-arginine conjugates.

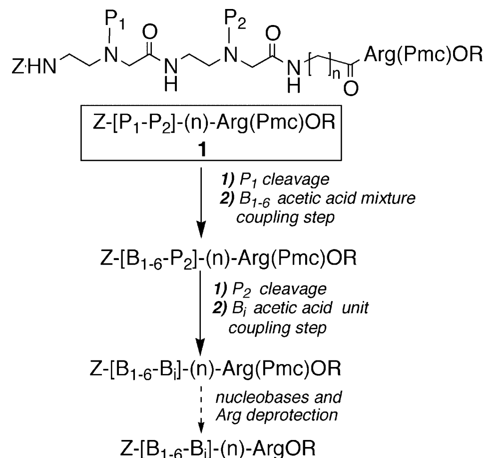
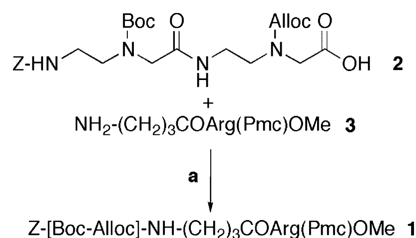
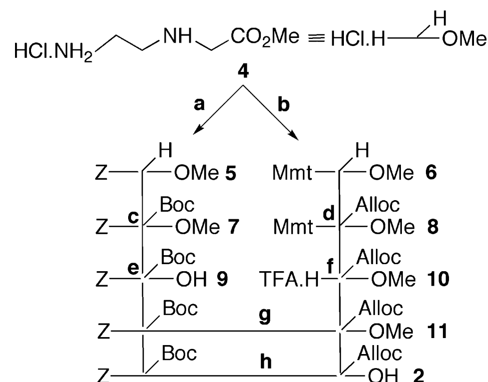


Figure 3. General strategy for the generation of diPNA-arginine conjugates library. P_1 and P_2 are two orthogonal protecting groups.

The synthesis of the *N*-protected dimer **2** was achieved in eight steps in 37% overall yield (Scheme 2). *Z*- and Mmt-protection of the primary amino function of backbone **4** gave **5** and **6**, respectively. Boc- and Alloc-protection of the secondary amine of **5** and **6**, respectively, led to the corresponding fully protected monomers **7** and **8**. Saponification of **7** and detritylation of **8** afforded acid **9** and the TFA ammonium salt



Scheme 1. Synthesis of key compound **1**. (a) DCC/HOSu, DMF (69%).



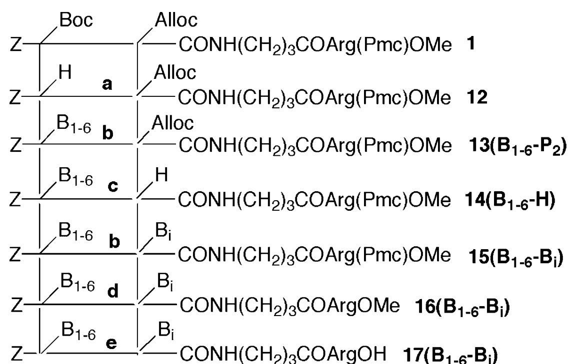
Scheme 2. Preparation of Boc- and Alloc-protected backbone **2**. (a) ZOSu, NMM, CH_3CN (66%); (b) Mmt-Cl, TEA, CH_2Cl_2 (90%); (c) $(Boc)_2O$, TEA, CH_2Cl_2 (100%); (d) Alloc-Cl, TEA, CH_2Cl_2 (88%); (e) 1 M LiOH, THF (82%); (f) TFA/CH_2Cl_2 (92%); (g) $ClCO_2i$Bu, DIEA, CH_2Cl_2 (96%); (h) 1 M LiOH, 0.8 M $CaCl_2$, *i*PrOH/H_2O 7/3 (85%).

10, respectively. Subsequent condensation of **9** and **10** using isobutyl chloroformate led to the fully protected dimer **11** which, after saponification, gave **2**.

The last steps of the preparation of the targeted sublibrary are described in Scheme 3. Starting from **1**, the Boc group was selectively and smoothly cleaved using a 5.5 M methanolic solution of HCl, generated in situ by addition of acetyl chloride to methanol.[11] The resulting hydrochloride salt **12** was then reacted stoichiometrically with an equimolar mixture of the six base-acetic acids [protected ($G^{(OBn)}$, $A^{(diBoc)}$, $C^{(diBoc)}$) or free (T, U, In)] by means of HATU in DMF. The depletion of **12** and base units $\mathbf{B_{1-6}}$ was monitored by HPLC.[12] As shown in Fig. 4B, HPLC analysis of the mixture isolated after a rough purification (DMF removal and triturating with water) clearly indicated the presence of the six neo-formed **13** ($\mathbf{B_{1-6}}$-$\mathbf{P_2}$) conjugates.

The selective Alloc-deprotection of **13** ($\mathbf{B_{1-6}}$-$\mathbf{P_2}$) was cleanly performed using $Pd[PPh_3]_4$ and DEA. The corresponding mixture of the six compounds **14** ($\mathbf{B_{1-6}}$-**H**) (as attested by HPLC, data not shown) was purified by triturating with diethylether. Then, condensation with each of the six B_{1-6} base units onto **14** ($\mathbf{B_{1-6}}$-**H**), via a HATU activation, gave six **15** ($\mathbf{B_{1-6}}$-$\mathbf{B_i}$) batches which were purified as described for mixture **13** ($\mathbf{B_{1-6}}$-$\mathbf{P_2}$). Treatment of each **15** ($\mathbf{B_{1-6}}$-$\mathbf{B_i}$) batch with 1/1 TFA/CH_2Cl_2 induced the simultaneous deprotection of the Boc and OBn-nucleobases ($A^{(diBoc)}$, $G^{(OBn)}$, $C^{(diBoc)}$) and of the N_G-Pmc arginine residue, affording the corresponding **16** ($\mathbf{B_{1-6}}$-$\mathbf{B_i}$) batch. Hydrolysis with 2 M HCl led to the corresponding acids **17** ($\mathbf{B_{1-6}}$-$\mathbf{B_i}$). The six batches **16** ($\mathbf{B_{1-6}}$-$\mathbf{B_i}$) and the six **17** ($\mathbf{B_{1-6}}$-$\mathbf{B_i}$) ones were purified by semi-preparative HPLC. The HPLC analysis of each of these clearly indicated the presence of the six conjugates. This is illustrated in Fig. 4C for the batch **16** ($\mathbf{B_{1-6}}$-**In**).

Sublibraries containing β-alanine ($n=2$), aminopentanoic acid ($n=4$) and aminocaproic acid ($n=5$) spacers have been prepared following the same strategy as described for $n=3$.



Scheme 3. Generation of the sublibrary Z-[B_{1-6},B_i]-(3)-ArgOR: (a) AcCl/MeOH (100%); (b) $B_{1-6}CH_2CO_2H$, HATU, DMF (87%); (c) $Pd[P(Ph)_3]_4$, DEA, DMF (98%); (d) $B_iCH_2CO_2H$, HATU, DMF (68–96%); (e) TFA/CH_2Cl_2 1/1 (100%); (f) 2 M HCl, CH_3CN (98%).

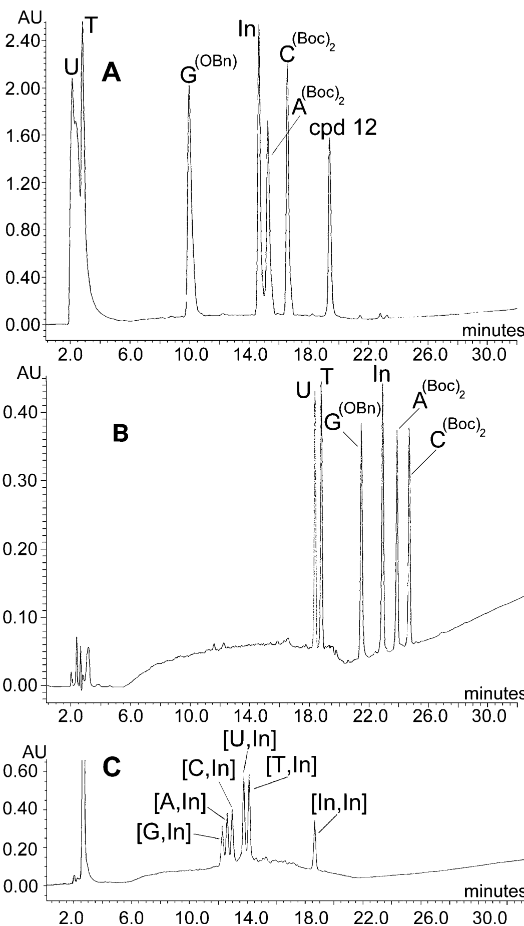


Figure 4. HPLC traces of: (**A**) the six B_{1-6} base acetic acids and compound **12**. (**B**) Z-[B_{1-6}-Alloc]-(3)-Arg(Pmc)OMe **13** ($\mathbf{B_{1-6}}$-$\mathbf{P_2}$). (**C**) Batch **16** ($\mathbf{B_{1-6}}$-**In**) Z-[B_{1-6}-In]-(3)-ArgOH.[12]

All batches of type **16** ($\mathbf{B_{1-6}}$-$\mathbf{B_i}$) (esters) and **17** ($\mathbf{B_{1-6}}$-$\mathbf{B_i}$) (acids) constituting the four sublibraries are currently under biological evaluation, for their antiviral activity on cells as well as for their ability to inhibit viral targets (RT, protease, integrase….). Moreover, to allow the easy identification of active compounds in batches showing some biological activity, without the need to resynthezise every component separately, the corresponding orthogonal library containing a well-identified PNA at the first position (i.e. Z-[$\mathbf{B_i}$-$\mathbf{B_{1-6}}$]-(n)-ArgOR) is under preparation. Indeed, any batch of a given sublibrary from library Z-[$\mathbf{B_{1-6}}$-$\mathbf{B_i}$]-(n)-ArgOR and any one of a given sublibrary from library Z-[$\mathbf{B_i}$-$\mathbf{B_{1-6}}$]-(n)-ArgOR share only one synthon. Therefore, for a given screen, an active compound will confer activity to one single batch of each libraries.

In conclusion, to establish structure–activity relationships for some promising antiHIV diPNA-arginine conjugates, we generated a library using a FPB strategy, which requires easily obtainable building blocks. This method allowed the preparation of 288 components, split into four sublibraries. Each group constituting these sublibraries has been synthesized in five steps (esters) or six steps (acids) from a fully protected back-

bone, in an overall yield of 70% and with a high final mixture purity. Biological evaluation of this library should permit us to determine the target(s) and the mode(s) of action of this family of compounds, as well as to identify new active molecules and/or to optimize the structure of active conjugates.

It should be noted that one could extend this strategy to prepare libraries of short linear and/or cyclic polyPNAs (less than decamer). Moreover, the monomeric and polymeric protected backbones could also find an application as building blocks in combinatorial solid-phase syntheses of PNA libraries.

Acknowledgements

We thank the 'Agence Nationale de Recherche sur le Sida' (ANRS), SIDACTION, CNRS, the 'Région Provence-Alpes-Côte d'Azur' and the Association pour la Recherche sur le Cancer (ARC) for their support.

References

1. (a) Gallop, M. A.; Barrett, R. W.; Dower, W. J.; Fodor, S. P. A.; Gordon, E. M. *J. Med. Chem.* **1994**, *37*, 1233–1280; (b) Terrett, N. K.; Gardner, M.; Gordon, D. W.; Kobylecki, R. J.; Steele, J. *Tetrahedron* **1995**, *51*, 8135–8173; (c) Ellman, J. A.; Gallop, M. A. *Curr. Opin. Chem. Biol.* **1998**, *2*, 317–319; (d) Guillier, F.; Orain, D.; Bradley, M. *Chem. Rev.* **2000**, *100*, 2091–2158.
2. Farese-Di Giorgio, A.; Pairot, S.; Patino, N.; Condom, R.; Di Giorgio, C.; Aumelas, A.; Aubertin, A. M.; Guedj, R. *Nucleosides Nucleotides* **1999**, *18*, 263–275.
3. (a) Nielsen, P. E.; Egholm, M. *Curr. Issues Mol. Biol.* **1999**, *1*, 89–104; (b) Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 2796–2823; (c) Dueholm, K. L.; Nielsen, P. E. *New J. Chem.* **1997**, *21*, 19–31.
4. (a) Cook, P. D.; Kiely, J.; Sprankle, K. WO Patent 9,523,163, **1995**; (b) Cook, P. D.; Kiely, J.; Sprankle, K. US Patent 5,831,014, **1998**; (c) Cook, P. D.; Kiely, J.; Sprankle, K. US Patent 5,539,083, **2001**.
5. Soonkap, H.; Patron, A.; Fagnani, R. WO Patent 2000 002 899, **2000**.
6. (a) Matysiak, S.; Reuthner, F.; Hoheisel, J. D. *BioTechniques* **2001**, *31*, 896–904; (b) Matysiak, S.; Hauser, N. C.; Wurtz, S.; Hoheisel, J. D. *Nucleosides Nucleotides* **1999**, *18*, 1289–1291 and references cited therein.
7. Nielsen, P. E. *Method Enzymol.* **1996**, *267*, 426–433.
8. Challa, H.; Styers, M. L.; Woshi, S. A. *Org. Lett.* **1999**, *10*, 1639–1641.
9. Di Giorgio, C.; Pairot, S.; Schwergold, C.; Patino, N.; Condom, R.; Farese-Di Giorgio, A.; Guedj, R. *Tetrahedron* **1999**, *55*, 1937–1958.
10. (a) Verheijen, J. C.; Grotenbreg, G. M.; Hart de Ruyter, L.; van der Klein, P. A. M.; van der Marel, G. A.; van Boom, J. H. *Tetrahedron Lett.* **2000**, *41*, 3991–3995; (b) Depecker, G.; Schwergold, C.; Di Giorgio, C.; Patino, N.; Condom, R. *Tetrahedron Lett.* **2001**, *42*, 8306–8308.
11. Nudelman, A.; Bechor, Y.; Falb, E.; Fisher, B.; Wexler, B. A.; Nudelman, A. *Synth. Commun.* **1998**, *28*, 471–474.
12. HPLC analyses were performed with a 996 photodiode array detector (195–290 nm), using a RP-18 (5 μm) Licrospher 100 (250×4 mm) column as support. Elution gradient A/B: 80/20 to 0/100 in 30 mn of solvents A (0.1% TFA H_2O) and B (0.1% TFA CH_3CN) was used with a flow of 1 mL/min. (**B**) Peaks attribution was determined by means of UV spectra (λ_{max} are caracteristic of each nucleobases). (**C**) Peaks attribution was performed by LC/MS.