Pergamon

Tetrahedron Letters 42 (2001) 8303–8306

TETRAHEDRON LETTERS

Liquid-phase synthesis of a cyclic hexameric peptide nucleic acid

Geoffrey Depecker, Caroline Schwergold, Christophe Di Giorgio, Nadia Patino and Roger Condom*

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

Received 23 July 2001; accepted 19 September 2001

Abstract—A cyclic fully *N*-protected hexameric (aminoethylglycinamide) can be readily obtained by using a divergent approach in liquid phase and consists of coupling orthogonal fragments of suitable oligomers. This cyclic *N*-protected backbone is then converted into a peptide nucleic acid by a series of selective deprotection–coupling steps affording the desired structure in good overall yields. Such a procedure could provide a general approach for targeting any short cyclic peptide nucleic acids (containing every kind of nucleobase).

1. Introduction

Some functional RNAs biologically act as folding stem-loop (or hairpin) structures which can be highly conserved and play an active role by interacting with proteins and/or nucleotidic sequences. For example, HIV trans-activation is dependent, among others, on the formation of a complex between TAR loop RNA and cellular factors.[1] Two complementary stem-loops may form 'kissing-loop' complexes which were found to regulate the replication[2] of ColE1 *Escherichia coli* plasmid and to promote dimerization of viral RNAs (SL1 loop of HIV-1).[3] Several studies[4] have further shown that kissing-loop interactions involving only a few nucleotide residues in each loop can be specific and selective. Recently, formation of such complexes was shown to occur between the natural TAR loop of HIV RNA (which only contains six bases) and its complementary synthetic one.[5] Moreover, kissing-loop complexes exhibit a particularly enhanced stability as compared to complexes resulting from the interaction between a loop and its linear complementary sequence.

RNA loops which are involved in pathogenic cellular, viral or bacterial expression processes therefore represent potential targets in biomedicinal chemistry. A large number of linear RNA inhibitors have been synthesized in the context of antisense strategies. However, their effectiveness is hampered by their poor cellular uptake and low in vivo stability.[6] Furthermore, owing to the required length of the antisense (15–20 mer) to ensure selectivity, their high cost limits considerably their therapeutic potential.

With the view of addressing these drawbacks, we are exploring the feasibility of targeting loops that are involved in crucial steps of viral expression by means of short and cyclic complementary oligonucleotide analogues. Specificity and selectivity could be brought by their cyclic nature, thus mimicking kissing-loop interactions. Furthermore, the use of RNA data bank 3D structures and molecular design could allow a perfect fitting between both the RNA target and the synthetic loop.

In this paper, we describe the synthesis of a cyclic hexameric PNA (peptide or polyamide nucleic acids) derivative **1** bearing two different pyrimidine nucleobases (see Fig. 1). This hexameric PNA derivative **1** represents the first generation of this type of synthetic loop. It can serve as a model to validate a more general

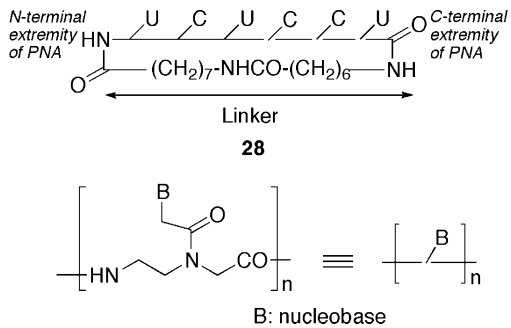


Figure 1. Cyclic hexameric PNA structure.

Keywords: kissing-loop complex; peptide nucleic acids; cyclization; orthogonally synthons; coupling steps.

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

0040-4039/01/$ - see front matter
PII: S0040-4039(01)01776-2

approach including natural and synthetic nucleobases to target natural biologically functional loops. Compound **1** is constituted by the U-C-U-C-C-U PNA sequence containing a linker tethered to *N*- and *C*-extremities of the PNA backbone through amide bonds. This hexameric PNA is complementary of a HIV RNA loop which is involved in the packaging of the HIV virus.[7] PNAs, which consist of *N*-(2-aminoethyl)glycine repeating units attached to a nucleic base through a methylenecarbonyl linker, have been chosen for their remarkable ability to recognize their complementary oligonucleotides (especially RNAs) with high specificity and affinity.[8] Moreover, their remarkable resistance to nucleases, their achiral, uncharged pseudopeptidic backbone and their increased lipophilicity make them more attractive than conventional charged antisense inhibitors for therapeutic use.[9] The length of the linker (17 atoms) was optimized by molecular modelization with the aim of allowing the head-to-tail cyclization to occur favourably.

Syntheses of linear PNA oligomers using solid-phase chemistry have been largely described and consist of an iterative process (coupling monomers or submonomers) to achieve elongation.[10,11] To synthesize the cyclic hexameric PNA derivative **1**, we used the fully protected backbone (FPB) approach that was been recently developed in our laboratory.[12] This approach consists of elaborating, using liquid-phase synthetic procedures, a fully protected linear poly(aminoethylglycinamide) containing as many different and orthogonal protecting groups as they are different potential elongation sites (P_1, P_2) and different base units to introduce (P_{A-D}) (Fig. 2). The hexameric backbone is then connected, at one or at both of its *N*- and/or *C*-extremities, to a linker. The following head-to-tail cyclization affords the cyclic *N*-protected hexameric backbone. Selective and sequential deprotection then allows the simultaneous condensation of the required number of identical nucleobase acetic acid units on to the backbone (Fig. 2).

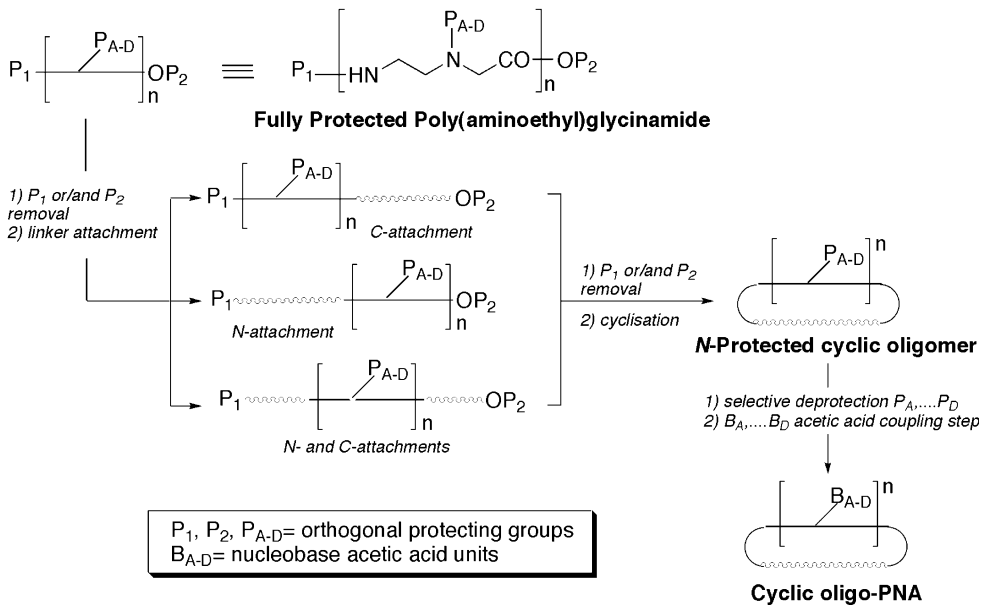
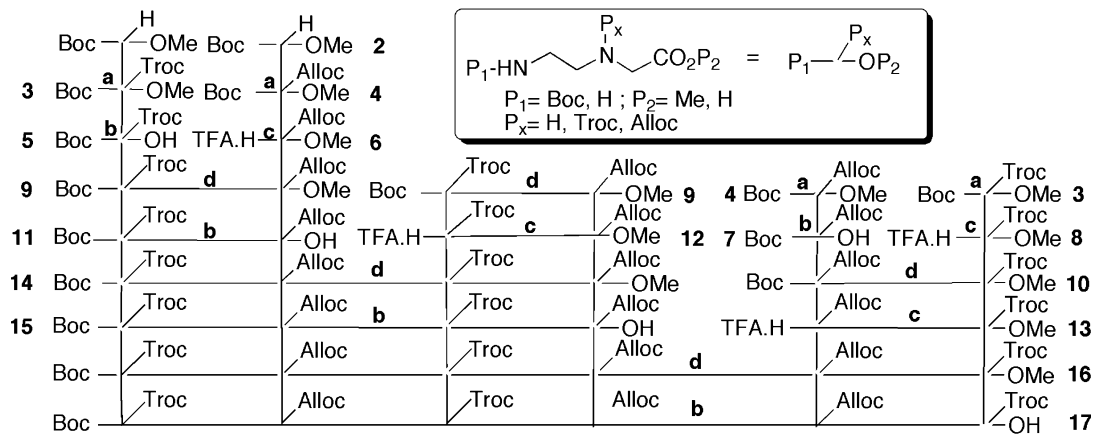


Figure 2. Synthetic strategy to cyclic PNA analogues.

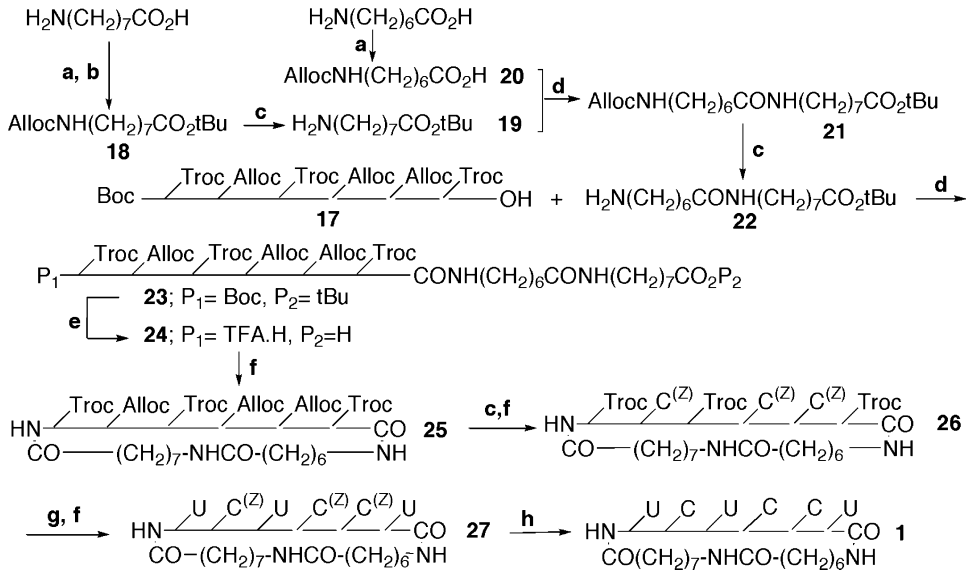


Scheme 1. *Reagents and conditions*: (a) Troc-Cl or Alloc-Cl, TEA, CH_2Cl_2, O°C, 10 min (90%); (b) $LiOH/CaCl_2$, *i*$PrOH/H_2O$, 2 h (90%); (c) TFA, CH_2Cl_2, 1 h (95%); (d) Cl-CO_2*i*Bu, NMM, CH_2Cl_2, −15°C, 15 min (80%).

The synthesis of the cyclic *N*-protected hexameric backbone, shown in Scheme 1, was performed via a divergent approach using classical deprotection–activation coupling reactions in peptide chemistry. The orthogonally *N*-protected aminoethylglycine **5**, **6**, and **7**, **8** monomer synthons were first prepared and condensed two by two to afford the fully protected **9** and **10** dimers, respectively, and, after suitable deprotections, their corresponding dimers **11**, **12** and **13**. The two dimer units **11** and **12** that were obtained from **9** were condensed together to produce tetramer **14** which, after C-terminal deprotection, was conjugated to dimer **13** to yield linear hexamer **16**. This hexamer, as its acid **17**, was then grafted onto the α-amino-ω-ester linker **22** giving derivative **23**. All Boc deprotections were performed with TFA/CH_2Cl_2. All saponifications of methyl esters were Ca^{2+} catalyzed using LiOH (1N), $CaCl_2$, *i*PrOH/H_2O medium. The coupling reactions were achieved via isobutylchloroformate activation. Linker **22** was synthesized, as shown in Scheme 2, from ω-aminooctanoic and ω-aminoheptanoic acids using routine deprotection–activation processes (six-step procedure from ω-aminooctanoic acid).

The one-step and simultaneous *N*- and *C*-deprotection of **23** with TFA/CH_2Cl_2 afforded the bifunctional TFA ω-amino-α-acid **24** salt in quasi quantitative yield. The key cyclization step of **24** leading to the cyclic fully protected *N*-poly(aminoethylglycinamide) backbone **25** was achieved using HOAt/HATU activation. The head-to-tail cyclization of **24** was performed under semi-high dilution conditions (10 mM) in order to minimize intermolecular condensation and afford **25**. Its structure and purity were clearly attested by NMR, mass spectrometry and HPLC analyses. Conversion of the orthogonally hexa-protected cyclic precursor **25** into the target PNA loop **1** required the introduction of the protected *Z*-cytosine and uracyl acetic acid base units **28** and **29**, respectively, followed by cytosine deprotection. First, the three Alloc protecting groups were cleanly removed from **25** with $Pd[PPh_3]_4$ as allyl acceptor and DEA, and three *N*-*Z*-cytosine acetic acid units **28** were then condensed onto the triamine thus produced by means of HATU/HOAt activation giving the mixed polyamide compound **26**. Next, the selective cleavage of the three Troc groups by means of cadmium in acetic acid followed by the HATU/HOAt-mediated condensation of three uracyl acetic acid units **29** on to the triamine thus generated, afforded the cyclic *Z*-cytosine and uracyl hexameric PNA **27**. Benzyloxycarbonyl (*Z*) removal was performed by treatment with HBr in glacial acetic acid.

In conclusion, a cyclic protected hexameric aminoethylglycinamide framework carrying as many different and orthogonal protecting groups as there are different types of nucleobases in the target PNA sequence has been prepared. This backbone allowed a cyclic hexameric PNA complementary to the PACK HIV RNA loop to be successfully synthesized following a liquid-phase procedure (17 steps starting from monomers, acetic acid unit bases and linker afforded 5 mg of pure **1**[†] in 6% overall yield). Although this particular hexameric PNA loop **1** contains only two different nucleobases, our FPB approach could be applied to any short cyclic PNA (less than decamer) containing the four natural or synthetic bases as a large palette of orthogonal protecting groups is available. This liquid-phase FPB approach allows selective, sequential and simultaneous attachment of several nucleobase units on the protected hexameric (aminoethylglycinamide) backbone. It also offers several advantages such as a greater choice of solvents and reagents, high solubility of the different protected backbones, high coupling efficiencies, intro-



Scheme 2. *Reagents and conditions*: (a) Alloc-Cl, 1N NaOH, dioxane, 1 h (90%); (b) *N*,*N*-dimethylformamide di-*tert*-butyl acetal, toluene, 80°C, 2 h (90%); (c) $Pd[P(Ph)_3]_4$, DEA, CH_2Cl_2, 10 min (90%); (d) Bop, NMM, CH_2Cl_2, 15 min (80%); (e) TFA, CH_2Cl_2, 2 h (95%); (f) HOAt, HATU, DIPEA, DMF, 12 h (90%) and **28** ($C^{(Z)}$-CH_2CO_2H) for compound **26**, 15 min (84%) or **29** (U-CH_2CO_2H) for compound **27**, 15 min (70%); (g) Cd/AcOH, DMF, 5 h (80%); (h) HBr/AcOH, 2 days (70%).

[†] HPLC (A/B 97:3 to 40:60 over 30 min): $t_R = 16.60$ min, $\lambda_{max} = 201.5$ and 270.9 nm. MS (ESI+) m/z 1778.7 $(M+H)^+$.

duction of the same natural nucleobase units in one step and at advanced stages of the synthesis.

2. Experimental

Analytical HPLC chromatograms were obtained using a WATERS 600 equipped with a 996 photodiode array detector (PDA, UV detector from 195 to 290 nm) and a column (250×4 mm) packed with Lichrospher 100-RP-18 (5 μm). A gradient with water (0.1% TFA) as solvent A and acetonitrile (0.1% TFA) as solvent B was used with a 1 ml/min flow. Mass spectrometry analyses were carried out on a TSQ 7000 Finnigan MAT (ESI) performed by J. M. Guigonis of GUMPAC, Nice (France).

Acknowledgements

We thank the 'Agence Nationale de Recherche sur le SIDA' (ANRS) and SIDACTION for their support and the LARTIC (Professor Cabrol-Bass, D.) laboratory for molecular design.

References

1. Gatignol, A. *Gene Expr*. **1996**, *5*, 217.
2. Tomizawa, J. I. *J. Mol. Biol*. **1990**, *212*, 695.
3. Mujeed, A.; Clever, J. L.; Billeci, T. M.; James, T. L.; Parslow, T. G. *Nat. Struct. Biol*. **1998**, *5*, 432.
4. Eisinger, J. *Biochem. Biophys. Res. Commun*. **1971**, *43*, 854.
5. Comolli, L. R.; Pelton, J. G.; Tinoco, I. *Nucleic Acids Res*. **1998**, *26*, 4688.
6. Crooke, S. *Adv. Pharmacol*. **1997**, *40*, 1.
7. Hayashi, T.; Shioda, T.; Iwakura, Y.; Shibuta, H. *Virology* **1992**, *188*, 590.
8. Nielsen, P. E. *Pure Appl. Chem*. **1998**, *70*, 105.
9. Uhlmann, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem. Int. Ed*. **1998**, *37*, 2796.
10. Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K. H.; Hansen, H. F.; Koch, T.; Egholm, M.; Buchardt, O.; Nielsen, P. E.; Coull, J.; Berg, R. H. *J. Pept. Sci*. **1995**, *3*, 175.
11. Richter, L. S.; Zuckermann, R. N. *Bioorg. Med. Chem. Lett*. **1995**, *5*, 1159.
12. Di Giorgio, C.; Pairot, S.; Schwergold, C.; Patino, N.; Condom, R.; Farese-Di Giorgio, A.; Guedj, R. *Tetrahedron* **1999**, *55*, 1937.