Design and synthesis of intrinsically cell-penetrating nucleopeptides†

Piero Geotti-Bianchini,[a,b] Julien Beyrath,[a] Olivier Chaloin,[a] Fernando Formaggio*[b] and Alberto Bianco*[a]

Received 8th July 2008, Accepted 5th September 2008
First published as an Advance Article on the web 17th September 2008
DOI: 10.1039/b811639c

Nucleopeptides, which are constituted of α-amino acids bearing nucleobases at their side chains, are able to penetrate into cells and to reach the nucleus without cytotoxic effects.

Nucleopeptides were first introduced at the beginning of the 70's,[1,2] but only recently have they emerged as a promising alternative to peptide nucleic acids (PNAs).[3–5] In 1991 the group of Nielsen demonstrated the ability of PNAs to strongly interact with DNA and RNA thus paving the way to their potential biomedical applications.[3–5] The PNAs may be constituted of many structural pseudopeptide motifs.[5] The most important and exploited scaffolds are based on the repetition of *N*-(2-aminoethyl)glycine residues containing the nucleobase linked to the central nitrogen *via* an amide bond.[3,4] Nucleopeptides, less explored than PNAs, are instead constituted of sequences of amino acids bearing nucleobases covalently bound to their side chains.[6] Typically, the nucleobase is linked in the β-position of the alanyl side chain.[7,8] The α- and β-nucleopeptides are able to form rigid and well defined double strands stabilized by H-bonding with complementary sequences.[9–12] The capability of these synthetic hybrids to form stable base-pairing might be useful for DNA displacement, specific recognition and modulation of the interactions with the natural nucleic acids. Alternatively, the nucleoamino acids were introduced at defined positions within a synthetic peptide backbone, constituted of protein α-amino acids, to take advantage of the conformational changes of the peptide during its interaction with complementary single-strand DNA/RNA sequences.[13,14] Nucleoamino acids are also versatile building blocks for the preparation of nucleobase substituted analogues of natural products as DNA intercalating agents,[15] or for the synthesis of novel and efficient double strand DNA binding ligands,[16] or new opioid selective mimetics.[17]

In the context of the biological applications, including gene regulation, one of the disadvantages of PNAs, the closest to drug development, is their poor cellular uptake.[5,18] Indeed, very few examples have shown that PNAs penetrate efficiently into the cells.[19,20] Cellular uptake can be facilitated by conjugating the PNAs to a cell penetrating peptide or to other carrier systems.[21–26] On the contrary the α- and β-nucleopeptides have never been studied for their capacity to cross the cell membranes without provoking undesired effects in terms of cell viability. Indeed, an easy, robust and efficient method to deliver nucleobase-rich oligomers into cells is very much needed.

To address this issue, we have designed a series of α-nucleopeptides based on the sequence Ala-AlaT-Ala [AlaT, β-(thymin-1-yl)alanine] repeated four times (Fig. 1). The nucleopeptides contain four nucleobases which were previously found sufficient for a considerable duplex stability.[11] Moreover, the polyalanine scaffold should allow enough structural flexibility for a possible interaction with complementary DNA or RNA strands. The total sequence of the Ala-rich 14-mers comprises also two Lys residues at both the N- and C-terminal ends of the nucleopeptides. The insertion of Lys has a dual purpose: i) to improve the solubility in physiological conditions; and ii) to allow the introduction of the fluorescent probe FITC (fluorescein isothiocyanate) or biotin (Fig. 1, **T4FIT** and **T4Bt**, respectively) for cell penetration and surface plasmon resonance studies.

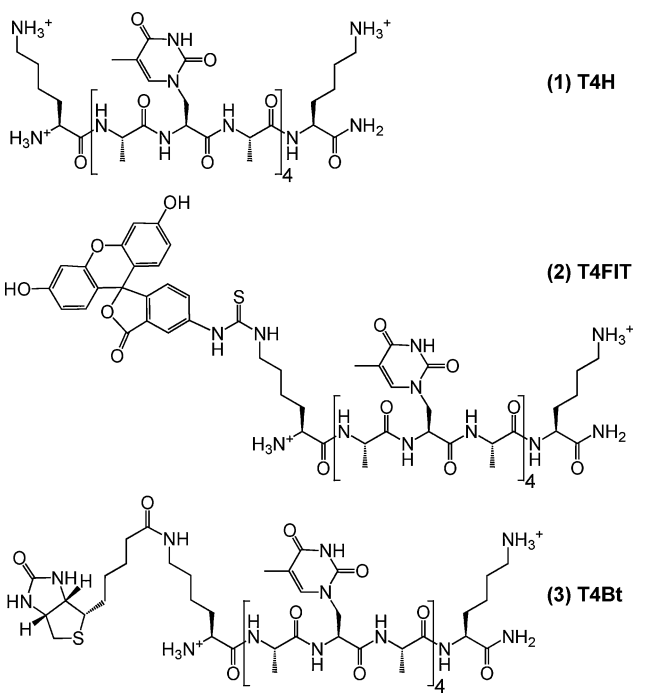

Fig. 1 Molecular structures of the nucleopeptides.

Initially, the nucleoamino acid AlaT was prepared following the approach of nucleophilic ring opening of the lactone of the Boc protected serine using the thymine base.[27,28] The nucleopeptides were subsequently prepared by solid-phase synthesis using the Merrifield strategy (see Supporting Information†).[29] Following the cleavage from the resin, the crude compounds were purified by HPLC to afford the desired products in a fairly good yield and purity higher than 94%.

[a]*CNRS, Institut de Biologie Moléculaire et Cellulaire, Laboratoire d'Immunologie et Chimie Thérapeutiques, 15 Rue René Descartes, 67000, Strasbourg, France. E-mail: A.Bianco@ibmc.u-strasbg.fr*
[b]*Dipartimento di Chimica, Università di Padova, via Marzolo 1, 35131, Padova, Italy. E-mail: fernando.formaggio@unipd.it*

† Electronic supplementary information (ESI) available: Details on the solid phase synthesis, characterization, cell uptake and viability tests, and surface plasmon resonance of the nucleopeptides. See DOI: 10.1039/b811639c

To assess the biological behaviour of our nucleopeptides, we initially studied their ability to penetrate into cells. RENCA cells were incubated for times between 5 and 60 min and up to 12 hours at 37 °C with concentrations of **T4FIT** and **T4Bt** ranging from 0.5 to 50 μM. The cells treated with **T4FIT** were directly analyzed using fluorescence microscopy and flow cytometry. As shown in Fig. 2A the nucleopeptide rapidly penetrates into the cells. We have found that this happens in a time (Fig. 2B) and dose dependent manner (Fig. S2–S4). At the highest concentration of **T4FIT**, more than half of the cells are stained. The fluorescent oligomers seem to be confined into vesicular compartments, particularly abundant in the perinuclear region. Cellular uptake of **T4FIT** is energy-dependent since the internalization is significantly reduced treating the cells with sodium azide and 2-deoxy-D-glucose (Fig. 2B).[30]

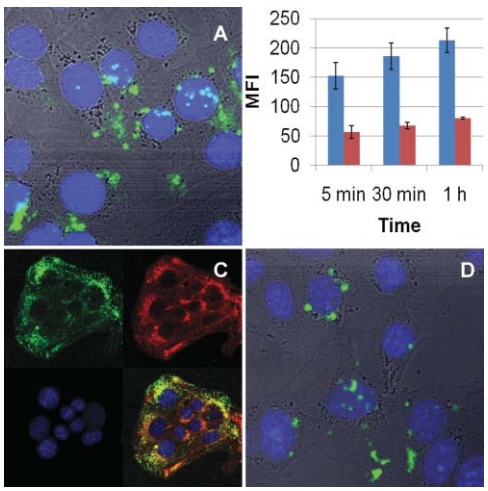

Fig. 2 (A) Confocal microscopy image of RENCA cells incubated with 50 μM of **T4FIT** for 12 h. (B) Time dependent uptake of 50 μM of **T4FIT** by Raji cells at 37 °C in the absence (blue bars) or in the presence (red bars) of endocytosis inhibitors NaN_3 and 2-deoxy-D-glucose, quantified by flow cytometry. Plots display the mean fluorescence signal of three experiments ± the standard error. (C) Confocal microscopy images of RENCA cells co-incubated with 50 μM of **T4FIT** (green) and 20 μg mL^{-1} of transferrin-Alexa-546 (red) at 37 °C for 30 min. Overlapping of the images evidences endosomal co-localization (yellow). (D) Confocal microscopy image of RENCA cells incubated with 50 μM of **T4Bt** at 37 °C for 12 h. Nucleopeptide was revealed with streptavidin-FITC (green). The cell nucleus was labelled with DAPI (blue).

Confocal images of the cells incubated with transferrin, a marker of endocytosis, confirmed a substantial co-localization with **T4FIT** (Fig. 2C).[31] This type of entrapment into endosomal vesicles was found also for the fluorescent PNAs.[19] However, our images show a diffusion of the nucleopeptide into the cytoplasm, suggesting either an endosomal escape or a possible contribution of an alternative route of cellular entry. In addition, nucleopeptides were found into the nucleus (Fig. 2A and D). The confocal analysis confirmed that they are inside and not simply associated to the nuclear membrane (see also Fig. S6). This behaviour is reminiscent of cationic cell penetrating peptides, as our nucleopeptides contain two positively charged residues which impart high water solubility.[30,31]

Similarly, the biotinylated nucleopeptide **T4Bt**, which was detected using a fluorescent streptavidin, presents the same behaviour (Fig. 2D). The capability of our nucleopeptides to enter the cells, and particularly to enter the nucleus, is an extremely relevant finding, as we may envisage using such conjugates to bind RNA or DNA and modulate their activity. Preliminary binding studies, using surface plasmon resonance, of the thyminyl-based oligomer **T4H** to its complementary sequence containing four adenylyl-nucleoamino acids (details on the synthesis are reported in Supporting Information) allowed us to calculate an apparent equilibrium dissociation constant in the micromolar range (Fig. 3). For this purpose, nucleopeptide **A4Bt** was immobilized on the sensor chip *via* streptavin previously covalently fixed to the dextran carboxylic functions on the chip gold surface, activated in turn with carbodiimide and *N*-hydroxysuccinimide. After this coupling, increasing concentrations of **T4H** from 12.5 to 50 μM were injected to evaluate the interaction capability of the two complementary nucleopeptides. Nucleopeptide **T4H** gradually complexed **A4Bt** as evidenced by an increase of the signal during the association phase. During the dissociation phase, **T4H** was very slowly released. The (apparent) association rate constant (k_a) and the (apparent) dissociation rate constant (k_d) corresponded to 3.18×10^2 M^{-1} s^{-1} and 1.09×10^{-2} s^{-1}, respectively. The resulting dissociation constant at the equilibrium (K_D) is 34 μM, which represents the affinity between the two complementary nucleopeptide chains. Comparable results were obtained using two synthetic complementary oligonucleotide sequences (dT_8 and dT_{10}) (Fig. S7). To prove that the recognition between the complementary bases on the nucleopeptides and DNA follows the Watson–Crick rules, and it is not only due to non specific base-stacking, we immobilized on the gold chip nucleopeptides containing four thymines or four cytosines (details on the synthesis are reported in Supporting Information). Such nucleopeptide sequences did not show any interaction either with **T4H** or with thymine-based DNA sequences dT_8 and dT_{10} (Fig. S8 and S9). This supports the capacity of nucleopeptides to specifically bind nucleotide sequences.

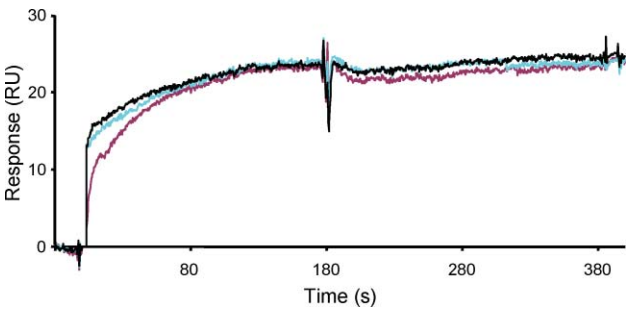

Fig. 3 Sensorgrams obtained by allowing various concentrations of nucleopeptide **T4H** to interact with the nucleopeptide **A4Bt** immobilized onto sensor chip. RU corresponds to the resonance unit (1000 RU = 1 ng mm^{-2} of analyte) [12.5 μM (magenta), 25 μM (cyan), 50 μM (black)].

To further investigate the impact of the nucleopeptides on cell behaviour, cell viability was evaluated using three different cell lines. Increasing doses of nucleopeptides **T4H** and **T4Bt** were added to the cell cultures. The cells were analyzed after 24 hour incubation and no significant decrease in the number of living cells was observed in comparison to the doxorubicin-treated control (Fig. 4). The viability was the same for Raji, Jurkat and BL41 cell lines (see also Fig. S10 and S11). This is another interesting

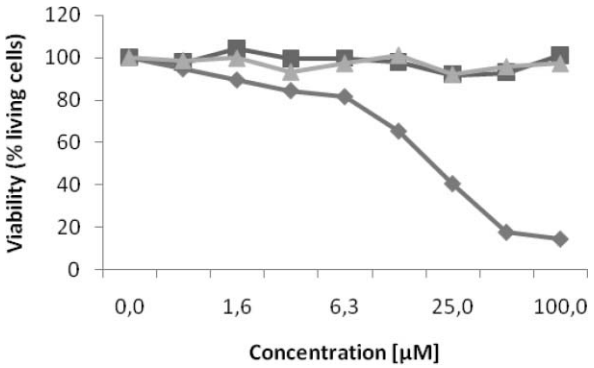

Fig. 4 Effect of **T4H** and **T4Bt** on the viability of Raji cell line. Concentration–response curves: Raji cells were treated with increasing concentrations of nucleopeptides for 24 h, followed by MTS test. For each experiment, the percentages are expressed relative to the mean value of the untreated cells. Square: **T4H**; Triangle: **T4Bt**; Diamond: Positive control (doxorubicin).

characteristic of our nucleopeptides which exhibit a remarkably low cytotoxicity when compared to PNAs.[18]

In summary, we have designed and prepared a series of highly water-soluble nucleopeptides which are easily taken up by the cells without toxic effects. Very importantly, they are able to pass the nuclear membrane thus paving the way for their use as DNA modulators. The improvement of cell penetration is certainly extremely beneficial for increasing the diagnostic and therapeutic efficacy of nucleopeptide-based drugs.

Acknowledgements

This work has been supported by the CNRS. P. G.-B. is grateful to Università italo-francese/Université franco-italienne for supporting his international PhD (VINCI program 2005). The authors wish to thank Marion Decossas and Stéphanie Lacotte for her precious help and fruitful discussions. Confocal images were recorded at RIO Cellular Imaging Platform of Esplanade (Strasbourg, France).

Notes and references

1 H. de Koning and U. K. Pandit, *Recl. Trav. Chim. Pays-Bas*, 1971, **91**, 1069.
2 J. D. Buttrey, A. S. Jones and R. T. Walker, *Tetrahedron*, 1975, **31**, 73.
3 P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497.
4 M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden and P. E. Nielsen, *Nature*, 1993, **365**, 566.
5 P. E. Nielsen, *Mol. Biotechol.*, 2004, **26**, 233.
6 B. Falkiewicz, *Acta Biochim. Pol.*, 1999, **46**, 509.
7 P. Lohse, B. Oberhauser, B. Oberhauser-Hofbauer, G. Baschang and A. Eschenmoser, *Croat. Chim. Acta*, 1996, **69**, 535.
8 U. Diederichsen, D. Weicherding and N. Diezemann, *Org. Biomol. Chem.*, 2005, **3**, 1058.
9 U. Diederichsen, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 445.
10 U. Diederichsen, *Angew. Chem., Int. Ed. Engl.*, 1997, **36**, 1886.
11 A. M. Brückner, P. Chakraborty, S. H. Gellman and U. Diederichsen, *Angew. Chem., Int. Ed.*, 2003, **42**, 4395.
12 P. Chakraborty and U. Diederichsen, *Chem.–Eur. J.*, 2005, **11**, 3207.
13 P. Garner, S. Dey and Y. Huang, *J. Am. Chem. Soc.*, 2000, **122**, 2405.
14 Y. Huang, S. Dey, X. Zhang, F. Sonnichsen and P. Garner, *J. Am. Chem. Soc.*, 2004, **126**, 4626.
15 K. B. Lorentz and U. Diederichsen, *J. Org. Chem.*, 2004, **69**, 3917.
16 Z. Zhan, P. Chaltin, A. Van Aerschot, J. Lacey, J. Rozenski, R. Busson and P. Herdewijn, *Bioorg. Med. Chem.*, 2002, **10**, 3401.
17 D. A. Kharkevich, N. V. Sumbatyan, A. N. Topin, O. N. Chichenkov, S. V. Zaitsev and G. A. Korshunova, *FEBS Lett.*, 1994, **351**, 308.
18 E. Uhlmann, A. Peyman, G. Breipohl and D. W. Will, *Angew. Chem., Int. Ed.*, 1998, **37**, 2796.
19 S. A. Noble, M. A. Bonham, J. E. Bisi, D. A. Bruckenstein, P. H. Brown, S. C. Brown, R. Cadilla, M. D. Gaul, J. C. Hanvey, C. F. Hassman, J. A. Josey, M. J. Luzzio, P. M. Myers, A. J. Pipe, D. J. Ricca, C. W. Su, C. L. Stevenson, S. A. Thomson, R. W. Wiethe and L. E. Babiss, *Drug Dev. Res.*, 1995, **34**, 184.
20 R. W. Taylor, P. F. Chinnery, D. M. Turnbull and R. N. Lightowels, *Nat. Genet.*, 1997, **15**, 212.
21 K. Wang, Q-F. Zhang, X-S. Wang, Y-W. Xue, D. Pang and S-B. Fu, *Chin. Med. J.*, 2004, **117**, 566.
22 M. Eriksson, P. E. Nielsen and L. Good, *J. Biol. Chem.*, 2002, **277**, 7144.
23 J. Oehlke, T. G. Wallukat, Y. Wolf, A. Ehrlich, B. Wiesner, K. Berger and M. Bienert, *Eur. J. Biochem.*, 2004, **271**, 3043.
24 K. Braun, P. Peschke, R. Pipkorn, S. Lampel, M. Wachsmuth, W. Waldeck, E. Friedrich and J. Debus, *J. Mol. Biol.*, 2002, **318**, 237.
25 L. C. Boffa, S. Scarfi, M. R. Maraini, G. Damonte, V. G. Allfrey, U. Benatti and P. L. Morris, *Cancer Res.*, 2000, **60**, 2258.
26 G. Cutrona, E. M. Carpaneto, M. Ulivi, S. Roncella, O. Landt, M. Ferrarini and L. C. Boffa, *Nat. Biotechnol.*, 2000, **18**, 300.
27 P. Lohse, PhD Thesis, ETH Zürich, 1992.
28 L. Lescrinier, C. Hendrix, L. Kerrenmans, J. Rozenski and P. Herdewijn, *Chem.–Eur. J.*, 1998, **4**, 425.
29 M. Goodman, A. Felix, L. Moroder and C. Toniolo, *Methods of Organic Chemistry (Houben-Weyl)*, *Vol. E22a*, Thieme, Stuttgart, 2002.
30 B. A. Smith, D. S. Daniels, A. E. Coplin, G. E. Jordan, L. M. McGregor and A. Schepartz, *J. Am. Chem. Soc.*, 2008, **130**, 2948.
31 F. Duchardt, M. Fotin-Mleczek, H. Schwarz, R. Fischer and R. Brock, *Traffic*, 2007, **8**, 848.