Replacement of Ala by Aib improves structuration and biological stability in thymine-based α-nucleopeptides

Piero Geotti-Bianchini,[a,b] Alessandro Moretto,[a] Cristina Peggion,[a] Julien Beyrath,[b] Alberto Bianco*[b] and Fernando Formaggio*[a]

Received 28th September 2009, Accepted 24th November 2009
First published as an Advance Article on the web 21st January 2010
DOI: 10.1039/b920211k

Three thymine-based nucleo-heptapeptides, each containing two nucleo-amino acids and zero, one or four Aib residues, respectively, have been synthesized. A single Aib residue is enough to promote the adoption of a helical structure in our nucleopeptides and to increase significantly their resistance towards enzymatic degradation. The insertion of four Aib residues, out of seven residues in the sequence, affords a rigid, 3_{10}-helical nucleopeptide that is substantially unaffected by serum enzymes and is not cytotoxic.

Introduction

Naturally occurring amino acids carrying nucleobases at their side chains were discovered several decades ago,[1] giving rise to the first interest in the biological properties of peptides containing such amino acids, which were named *nucleopeptides*.[2] However, only in the last two decades the research on nucleopeptides has significantly developed, mainly in view of the possibility to specifically design synthetic peptide/nucleotide hybrids for gene therapy applications.[3–7] The most known and investigated oligonucleotide analogues with a polyamide backbone are the *peptide nucleic acids* (PNA),[8] in which the nucleobases are anchored to the α-nitrogen through a linker. Nucleopeptides consisting of α-amino acids are often called αPNA, as they share with the PNA the polyamide backbone.[4d,7]

Peptide-based constructs have remarkable advantages with respect to the phosphate–sugar backbone of the natural polynucleotides in view of *in vivo* applications. Indeed, peptide backbones are not degraded by nucleases (active on phosphodiester bonds and ubiquitous). In addition, they can often more easily cross the cell membranes, as they lack the large negative charge of polynucleotides. Finally, hydrophobicity, chemical stability, 3D-structure and additional features can be tailored to specific needs by an appropriate choice of the amino acid building blocks.

We have recently reported that sequential nucleopeptides, characterized by the insertion of α-alanyl nucleoamino acids at (*i*, *i*+3) positions, are able to penetrate the cells and reach the nucleus without cytotoxic effects.[9] As conformation plays a major role in determining the functionality of oligonucleotides[10] and peptide/nucleobase hybrids,[4d] we have also explored[11] the possibility to build nucleopeptides with rigid and predictable 3D-structures by exploiting the helix induction properties of α-aminoisobutyric acid (Aib, Fig. 1).[12] This latter achievement is quite remarkable as it offers a tool to spatially arrange the nucleobases at well defined positions. However, the syntheses of these conformationally controlled nucleopeptides[11] had to be performed by time-consuming solution methodologies, since the faster solid-phase synthesis approach gives unsatisfactory results when a large number of sterically hindered Aib residues is present. Therefore, in the present work we combine the faster solid-phase synthesis technique with the advantages offered by the insertion of Aib residues into our (*i*, *i*+3)-functionalized nucleopeptides. We aim at verifying if as few as one Aib residue is sufficient to impart conformational and biological stabilities to a nucleo-heptapeptide containing two β-thyminyl-alanine (AlaT, Fig. 1) residues. In particular, we have synthesized and studied H-(Ala-AlaT-Ala)$_2$-Lys-NH$_2$ (**T2**), H-Ala-AlaT-Ala-**Aib**-AlaT-Ala-Lys-NH$_2$ (**T2U**), with one Aib in the middle of the sequence, and H-(**Aib**-AlaT-**Aib**)$_2$-Lys-NH$_2$ (**T2U4**), with all Ala replaced by Aib (Fig. 1).

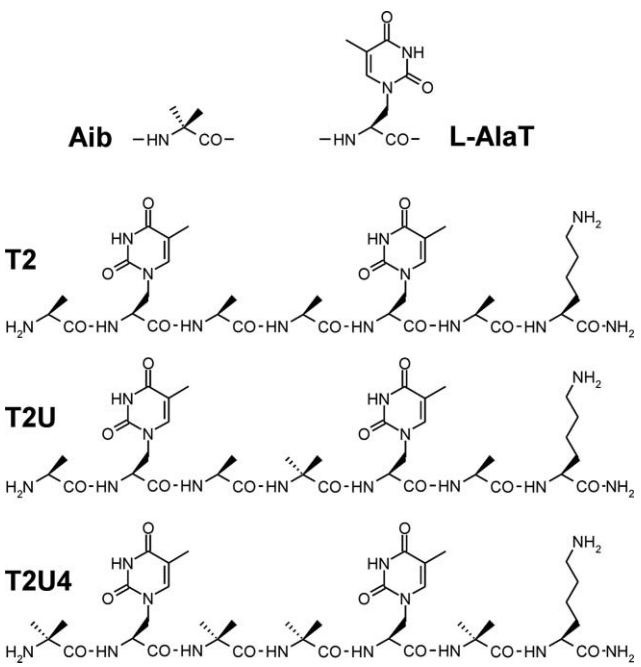


Fig. 1 Chemical structures of Aib, AlaT and the three nucleo-heptapeptides synthesized.

[a]*Department of Chemistry, University of Padova, 35131 Padova, Italy. E-mail: fernando.formaggio@unipd.it*
[b]*CNRS, Institut de Biologique Moléculaire et Cellulaire, Laboratoire d'Immunologie et Chimie Thérapeutiques, 67000 Strasbourg, France. E-mail: A.Bianco@imbc-cnrs.unistra.fr*

Results and discussion

Design

According to the encouraging results of our previous work on nucleopeptides,[9,11] we decided to maintain the two AlaT residues at the (*i*,*i*+3) relative positions. This arrangement would allow for the alignment of the two nucleobases on the same side of the helix, provided the peptide adopts a 3_{10}-helical conformation. We added one Lys residue to impart water solubility, as it has been shown to be sufficient for peptides of this length.[4d] Finally, we chose to have a primary amide at the C-terminus since the negative charge of the carboxylate could hamper the membrane penetration and decrease the affinity towards the polyanionic nucleotides.

We then decided to insert one Aib in the middle of the sequence of the reference nucleopeptide **T2**, in the hope that this residue could exert its conformational bias on the entire nucleopeptide. Given the lower reactivity of Aib at its amino function, as compared to its carboxylic moiety, the Ala[3]–Aib[4] coupling was expected to be easier than the AlaT[2]–Aib[3], so that position 4 was chosen for the insertion.

Synthesis

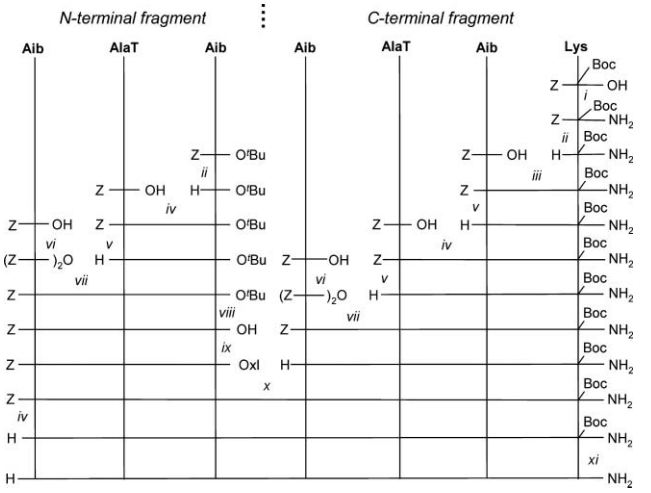
Z- and Boc-L-AlaT-OH were prepared according to literature procedures,[11,13] starting from the appropriate N^{α}-protected L-Ser β-lactone.

The synthesis of **T2** and **T2U** was accomplished on solid phase by applying the Boc/Bzl strategy on a MBHA resin. Yields in AlaT introduction were much lower as compared to an Ala residue, probably because of the relatively large side-chain and to a possible H-bonding interference by the thymine moiety. No significant improvement was observed when the BOP/HOBt activation protocol was replaced by the often more efficient HATU/HOAt mixture. As Aib reacts sluggishly in amide bond formation, triple coupling, longer reaction times and the HATU/HOAt activation were used. Longer reaction times were applied also for the introduction of Ala[3]. These procedures allowed us to obtain **T2** and **T2U** in comparable yields (29% **T2**, 26% **T2U**).

Because of the four Aib residues, two of them consecutive, **T2U4** was prepared in solution by a segment condensation approach (Scheme 1), previously successfully applied to a similar nucleopeptide.[11] Firstly, the C-terminal tetrapeptide amide *Z*-Aib-AlaT-Aib-Lys(Boc)-NH_2 and the N-terminal tripeptide *Z*-Aib-AlaT-Aib-O*t*Bu were synthesized step-by-step. Then the N-terminal segment was C-deprotected by acidolysis, transformed into its reactive 4*H*-oxazolidin-5-one by treatment with a carbodiimide, and coupled with the N-deprotected tetrapeptide amide to yield the protected **T2U4**. The pure final nucleopeptide was obtained after removal of the Z and Boc protecting groups (by catalytic hydrogenation and TFA acidolysis, respectively), followed by a HPLC purification step.

Conformational analysis

The three nucleo-heptapeptides were studied by 2D-NMR techniques in d_6-DMSO solution, at 5 mM concentration. All proton resonances could be assigned with the only exception of the heterocyclic N(3)H proton of every couple of thymines, due to overlapping, and of the two NH_3^+ in **T2U4**.



Scheme 1 Synthetic strategy adopted for the solution synthesis of **T2U4**. (i) EDC/HOBt, $NH_{3(g)}$ in DCM; (ii) H_2, Pd/C, DCM; (iii) EDC/HOBt, NMM, pH 8, DCM; (iv) EDC/HOAt, NMM, pH 8, DMF; (v) H_2, Pd/C, MeOH; (vi) EDC, DCM; (vii) NMM, pH 8, MeCN; (viii) 50% TFA in DCM; (ix) EDC in DMF; (x) NMM, pH 8 in MeCN, reflux; (xi) 20% TFA in DCM.

Conformationally informative data were obtained from the ROESY spectra. As summarized in Table 1, for all nucleopeptides we observed most of the possible sequential cross-peaks of the types NH(*i*)→NH(*i*+1) and C^{α}H(*i*)→NH(*i*+1) [C^{β}H(*i*)→NH(*i*+1) in the case of the Aib residues], indicative of folded conformations.[14] In particular, none of such cross-peaks is missing in the ROESY spectrum of **T2U4**. Moreover, in this latter case two medium range, C^{β}H(*i*)→NH(*i*+2) spatial connectivities were observed (Fig. 2), which strongly support[14] the conclusion that **T2U4** adopts a 3_{10}-helical[15] conformation in solution. Medium range connectivities are not observed for **T2** and only two quite weak (and between nuclei three residues apart) are found for **T2U**. Overall, our NMR data indicate that **T2U4**

Table 1 Sequential cross-peaks detected in the 2D-NMR ROESY spectra of **T2**, **T2U** and **T2U4**

Cross-peak	*i*	T2	T2U	T2U4
NH(*i*)-NH(*i*+1)	1	[b]	[b]	[b]
	2	+	+	+
	3	+	+	+
	4	*n. d.*[c]	+	+
	5	+	+	+
	6	+	+	+
	7	+	+	+
C^{α}H(*i*)-NH(*i*+1)[a]	1	*n. d.*	+	+
	2	+	+	+
	3	+	*n. d.*	+
	4	*n. d.*	+	+
	5	+	+	+
	6	+	+	+
	7	++[d]	++[d]	++[d]

[a] In the case of Aib, lacking the C^{α} proton, the C^{β}H(*i*)→NH(*i*+1) cross peaks are considered. [b] The rapid exchange of the ammonium protons with the solvent moisture suppresses most of their cross-peaks. [c] *n. d.*: cross-peak not detected. [d] Cross-peaks with both non-equivalent $CONH_2$ protons are observed

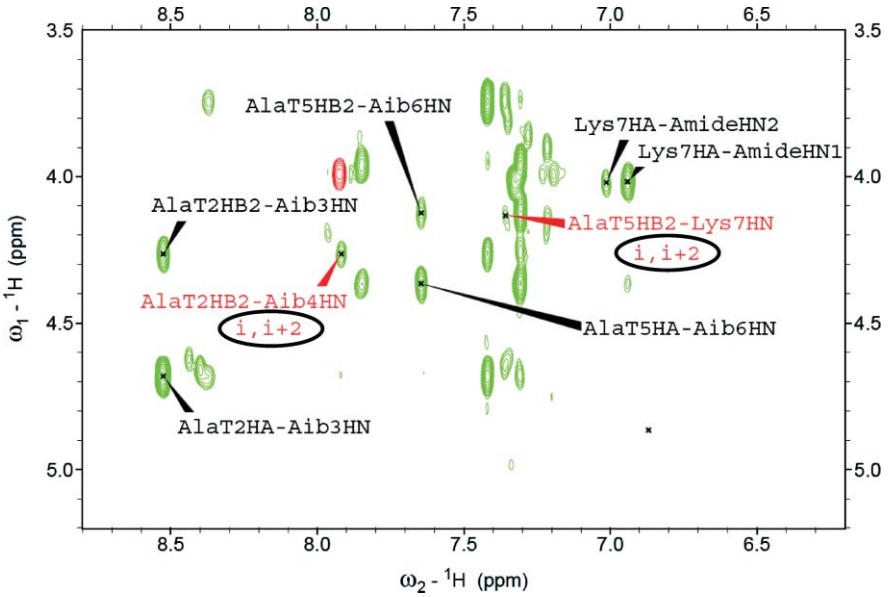


Fig. 2 Section of the ROESY spectrum of **T2U4** (5 mM in d_6-DMSO). The two medium range $C^{\beta}H(i)\rightarrow NH(i+2)$ connectivities, diagnostic of a 3_{10}-helix, are highlighted.

is folded into a helical conformation, **T2U** folds to a significant extent, whereas **T2** has the lowest degree of helical content.

A 3_{10}-helical disposition would locate the two thymine moieties on the same side of the helix, after one complete helix turn. This 3D arrangement is indeed confirmed for **T2U4**, in the ROESY spectrum of which 4 spatial connectivities between the two thymines are observed (Fig. 3). Only one of these is seen in the spectrum of **T2U** and none in that of **T2**.

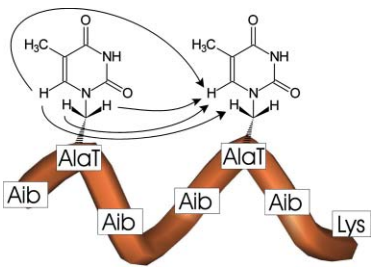


Fig. 3 The four connectivities (indicated by the arrows) observed between the two thymines in the ROESY spectrum of **T2U4** (5 mM in d_6-DMSO).

Circular dichroism spectra of the three nucleo-heptapeptides were recorded in phosphate buffer (10 mM NaH_2PO_4/Na_2HPO_4, pH 7.4), at 40 μM peptide concentration (Fig. 4). All spectra display a negative band, mostly due to the nucleobase chromophore, in the range 250–300 nm, and a more intense negative band in the low-wavelength region ($\lambda < 230$ nm). The nucleobase band is slightly more intense for **T2U** than for **T2**, but definitely much larger and 10 nm red-shifted in the case of **T2U4**. The increased intensity of the dichroic signal is likely due to a lower degree of conformational freedom experienced by the two nucleobases in the side chain of **T2U4**,[5c] as a consequence of the rigidified backbone structure.

Informative hints come also from the CD spectra recorded in the same environment, but at different temperatures (0–60 °C range). Shape and intensity of the spectrum of **T2U4** are almost

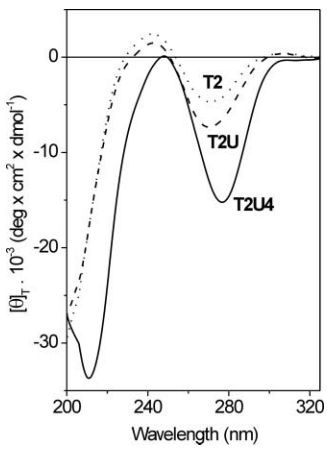


Fig. 4 CD spectra of **T2**, **T2U** and **T2U4** in phosphate buffer, pH 7.4 (room temperature, 40 μM nucleopeptide concentration).

unaffected by the temperature variation, whereas the spectra of **T2** and **T2U** undergo a significant change, particularly in the low-wavelength spectral region ($\lambda < 230$ nm). Fig. 5 illustrates the effect of the temperature on the total molar ellipticity value at 208 nm, the typical absorption wavelength for the amide $\pi\rightarrow\pi^*$ transition (parallel component) of a helical peptide.[16,17] It is clear that **T2U4** is endowed with a superior structural stability as compared to **T2** and **T2U**.

Concerning the nature of the secondary structures adopted by the three nucleopeptides, we must rely mainly on the NMR data. Indeed, it is difficult to estimate the contribution of the nucleobase chromophore to the amide absorption region (190–250 nm), the most conformationally informative for a peptide. However, keeping this advice in mind we can tentatively draw some conclusions, as previously done by other authors on different nucleopeptides.[5c,6] We observe (Fig. 4) a negative maximum for **T2U4**, at about 215 nm, and for **T2U**, at about 203 nm, but not for **T2** in the same spectral region. The presence of such a

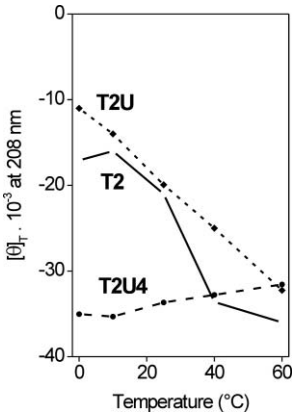


Fig. 5 $[\theta]_T$ values at 208 nm for **T2**, **T2U** and **T2U4** at different temperatures.

negative maximum is often indicative of a helical (α or 3_{10}) peptide structure, whereas extended ("unordered", or poly-Pro-II type) conformations are characterized by a negative maximum below 200 nm (usually around 195 nm).[17] Therefore, **T2U** and **T2U4** are likewise folded into a helical structure, as indicated also by our NMR investigation. Altogether, our CD analysis suggests that nucleobase rigidity and helical content markedly increase when Ala residues are replaced by Aib.

Biological tests

The stability of the three nucleopeptides towards enzymatic hydrolysis was tested in murine serum. HPLC analysis at increasing incubation times showed that **T2** is rapidly degraded, while **T2U** is significantly more resistant and **T2U4** is completely stable even after 48 h incubation (Fig. 6). In particular, after 1 h incubation only 5% of the initial amount of **T2** is left, whereas the quantity of non-degraded **T2U** is 10 times higher and **T2U4** is still intact. The observed behavior is probably due to the higher efficacy of hydrolytic enzymes in cleaving amide bonds among coded residues, rather than peptide bonds involving non-coded amino acids.[18] Indeed, similar results have been recently reported for Aib-containing peptides.[19]

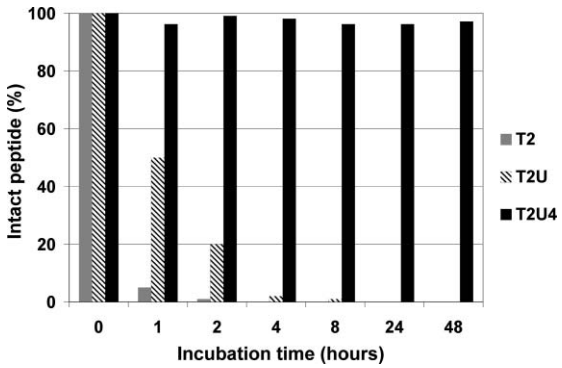


Fig. 6 Fraction of intact nucleopeptide after different incubation times in 50% murine serum, evaluated by analytical HPLC.

We have previously shown that Ala-rich nucleopeptides bearing AlaT residues at (*i*,*i*+3) positions are not cytotoxic.[9] Therefore, a cytotoxicity test was performed only on **T2U4**, the nucleopeptide with the highest content of Aib residues.

Human lymphoid tumor Raji cells were incubated for 24 h with different concentrations of **T2U4** and their viability was evaluated by the spectrophotometric MTS assay.[20] As shown in Fig. 7, **T2U4** has no cytotoxic effects up to 100 μM concentration.

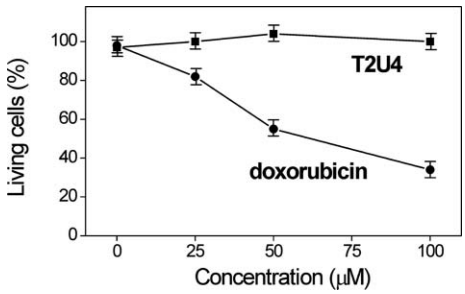


Fig. 7 Raji cells viability, evaluated by the MTS assay, after 24 h incubation with different amounts of **T2U4** or doxorubicin (positive control). The average of three experiments and their standard deviation are plotted.

Conclusions

Three thymine-based nucleo-heptapeptides (**T2**, **T2U**, **T2U4**) containing two AlaT residues at (*i*,*i*+3) relative positions and zero, one or four Aib residues, respectively, have been synthesized by solid-phase (**T2**, **T2U**) or solution (**T2U4**) procedures. The conformational analysis indicates that **T2U4** adopts a very stable helical structure, most likely of the 3_{10} type. Also **T2U**, containing only one Aib residue, is folded to a significant extent, thus confirming the known ability of Aib to promote helical structures. The nucleopeptide **T2** appears to be the less folded among the three. Tests in mouse serum have shown that a single Aib residue is sufficient to impart a tenfold resistance increase towards enzymatic degradation. The insertion of four Aib, out of seven residues, affords a nucleopeptide that is stable even after 48 h incubation without detected cytotoxicity.

At this stage of our research it would be premature to conclude that these biological results are directly due to the presence of Aib residues. Indeed, the observed behavior could arise from the (helical) secondary structure formed as a consequence of the Aib introduction. In any case, the encouraging results of this work stimulate us to develop longer Aib-containing nucleopeptides, amenable to be synthesized by means of the faster solid-phase technique. We believe that a small percentage of Aib residues could be sufficient to impart resistance to enzyme degradation and a well defined 3D-structure. This latter feature, in turn, might allow us to design nucleopeptides with a nucleobase spatial arrangement specifically tailored for interactions with a given target (*e.g.*, a polynucleotide sequence).

Experimental

Materials and methods

Solvents. Reagent grade *N*,*N*-dimethylformamide (DMF) from Baker (Phillipsburg, USA) was stored over 3 Å molecular sieves under N_2. Reagent grade dichloromethane (DCM) from Carlo Erba (Milan, Italy), freshly distilled over phosphoric anhydride, was used for the couplings.

Acetic acid (AcOH), 1-butanol, chloroform, dichloromethane (DCM), diethyl ether (Et_2O), ethanol, ethyl acetate (AcOEt), methanol (MeOH), petroleum ether, boiling range 40–60 °C (PE) and toluene were at least reagent grade and used without further purification.

Acetonitrile (MeCN) was HPLC grade and used without further purification.

Reagents. *Para*-cresol and trifluoroacetic acid (TFA) were purchased from Acros-Janssen (Geel, Belgium). Potassium hydrogensulfate, sodium hydrogencarbonate, sodium hypochlorite and sodium sulfate anhydrous were purchased from Carlo Erba (Milan, Italy). *N*,*N*-diisopropylethylamine (DIEA), *N*-methylmorpholine (NMM), ninhydrin and palladium catalyst (10% on activated charcoal) were purchased from Fluka (Buchs, Switzerland). 7-Aza-1-hydroxy-1*H*-benzotriazole (HOAt) and *O*-(7-aza-benzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HATU) were purchased from GL Biochem Lt (Shangai, China). *N*-Ethyl-*N*′-(3-dimethylamino)propyl-carbodiimide chlorohydrate (EDC·HCl), *O*-(benzotriazol-1-yl)- *N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (BOP) were purchased from Iris Biotech (Marktredwitz, Germany). Disodium hydrogenphosphate dihydrate and sodium dihydrogenphosphate monohydrate were purchased from Merck (Darmstadt, Germany). Formic acid, iodine, *N*,*N*,*N*′,*N*′-tetramethyl-4,4′-diamino-diphenylmethane (TDM) and trimethylsilyl trifluoromethansulfonate (TMSOTf) were purchased from Sigma-Aldrich (Milwaukee, USA).

Silica gel 60 (40-63 μm diameter, mesh 230-400) from Merck was used for flash-chromatographic purifications.

Characterization. Melting points were measured on a Leitz apparatus model Laborlux 12 and are uncorrected.

Silica gel 60 F_{254} (Merck) on glass was used for TLC characterization. Retention factors (R_f) were measured using the following solvent systems as eluents: $CHCl_3$–ethanol 9 : 1 (R_f1), 1-butanol–AcOH–H_2O 3 : 1 : 1 (R_f2), toluene–ethanol 7 : 1 (R_f3). Products were detected either by UV lamp irradiation or by exposition to I_2 vapors or by warming with a heat gun and spraying firstly with a 1.5% NaClO solution, then with a ninhydrin–TDM solution.[21] All compounds were chromatographically pure.

Optical rotations were measured on a Perkin-Elmer model 241 polarimeter at the sodium D line wavelength, using a cell with an optical pathlength of 10 cm. Concentrations are expressed in g/100 mL. $[\alpha]_D^{25}$ are calculated using the formula $[\alpha] = \alpha/(c \cdot l)$, where c is the concentration (in g mL^{-1}) and l is the optical path (in dm). Spectrophotometric grade MeOH (Fluka) was used as a solvent.

IR absorption measurements on KBr pellets were performed on a Perkin-Elmer model 580 B spectrophotometer, equipped with a Perkin-Elmer 3600 IR data station.

Analytical chromatograms were recorded either on an Agilent Technologies 1200 Series instrument, provided with a Kromasil Phenomenex C18 column (4.6 × 250 mm, 100 Å, flow 1.0 mL min^{-1}) or on a Varian ProStar 410 instrument, provided with a Nucleosil 100-5 C18 column (4.6 × 150 mm, 100 Å, flow 1.2 mL min^{-1}) and a Varian ProStar 330.71 PDA detector.

Preparative scale chromatographic purifications were performed either on a on a Beckmann System Gold HPLC 166 instrument, provided with a SDS C18 Macherey-Nagel column (10 × 250 mm, 100 Å, flow 6.0 mL min^{-1}) and a Beckmann System Gold UV detector operating at 230 nm, or on a Shimadzu SCL-6B instrument, provided with Shimadzu LC-8A pumps, a Delta Pack C18 column (19 × 300 mm, 100 Å, flow 17 mL min^{-1}) and a Shimadzu SPD-64 detector operating at 265 nm. As eluents, the following solvent mixtures were used: (1) A: H_2O + 0.1% TFA, B: MeCN + 0.08% TFA (HPLC Varian and Beckmann); (2) A′: H_2O–MeCN 9 : 1 + 0.05% TFA; B′: MeCN–H_2O 9 : 1 + 0.05% TFA (HPLC Agilent and Shimadzu). MilliQ H_2O was used for eluent preparation.

1D ^{1}H-NMR spectra were recorded on Bruker AC 200, Bruker AC 250 or DRX 400 spectrometers. 2D NMR spectra were recorded on AMX 500 Bruker or DMX 600 spectrometers. Chemical shifts (δ) are expressed in parts per million (ppm) with regard to the tetramethylsilane signal. Solvent residual peaks ($CHCl_3$, δ 7.26 ppm, or d_5-DMSO, δ 2.50 ppm) were used for calibration. Peak multiplicity is described as follows: s (singlet), br s (broad singlet), d (doublet), dd (doublet of doublets), m (multiplet), coupling constant (J) values are given in Hz. Deuterated solvents ($CDCl_3$, d_6-DMSO) were purchased from Cambridge Isotope Laboratories (Cambridge, England).

ESI-TOF mass spectra were collected on a Mariner ESI-TOF mass spectrometer (Perseptive Biosystems). Prior to injection, samples were dissolved in a 1 : 1 mixture of H_2O–MeOH containing 0.1% formic acid. The positive ions were accelerated at 10, 15, 20 or 30 keV.

Circular dichroic spectra were recorded on a Jasco (Tokyo, Japan) model J-715 dichrograph provided with a Haake model F3 thermostat. Fused quartz cells (Hellma, Müllheim, Germany) of 0.5 cm path length were employed. Data are expressed in terms of total molar ellipticity [θ_T] (deg cm^2 $dmol^{-1}$).

Solid-phase peptide synthesis

Solid-phase synthesis of the nucleopeptide sequences was accomplished using a PSP 4000 semiautomatic peptide synthesizer for Boc/Bzl strategy.[22] 4-Methylbenzhydrylamine hydrochloride salt resin (loading 0.62 mmol g^{-1}), purchased from Applied Biosystems (Carlsbad, USA) was used as solid support. Boc-Ala-OH, Boc-Lys(2Cl-Z)-OH and Boc-Aib-OH were purchased from NeoMPS (Strasbourg, France). Boc-AlaT-OH was prepared according to literature procedures.[13] Couplings were allowed to proceed for 20 min for activated protein amino acids on N-terminal protein amino acids, 40 min for activated protein amino acids on N-terminal AlaT, 60 min for activated protein amino acids on N-terminal Aib, for activated Boc-AlaT-OH and for activated Boc-Aib-OH. Each coupling was repeated twice (three times for activated Boc-Aib-OH) and subsequently the qualitative Kaiser test[23] was used to verify the completion of the reaction. Boc deprotection was accomplished by treating with neat TFA twice (firstly for 1 min then for 3 min). After completing the peptide sequences, the resin beads were transferred to syringes provided with sintered glass filters, a TFA–TMSOTf–*para*-cresol (80 : 25 : 10, v/v/w) cleavage mixture[24] was added and cleavage was allowed to proceed overnight at room temperature. Crude nucleopeptides were recovered by precipitation with cold Et_2O, dissolved in milliQ H_2O and lyophilized prior to HPLC purification (Beckmann HPLC, 1-21% B in 30 min).

2*TFA·H-(Ala-AlaT-Ala)$_2$-Lys(H)-NH$_2$ [T2]. Synthesis scale: 25 μmol. Activation protocol: 5 eq each of Boc-protected protein amino acid, BOP, HOBt and 15 eq DIEA or 3 eq each of Boc-AlaT-OH, HATU, HOAt, 10 eq DIEA. Yield: 5.9 mg (29%). HPLC (Varian, 0–50% B in 20 min): t_r 8.41 min. NMR δ_H (500 MHz; d_6-DMSO) 11.22-11.20 (2H, 2 br s, N(3)H thymine AlaT^2 and AlaT^5), 8.80-8.79 (1H, d, *J* 7.0, αNH AlaT^2), 8.30-8.28 (1H, d, *J* 6.5, NH Ala^3), 8.25-8.24 (1H, d, *J* 6.5, NH Ala^4), 8.19 (3H, br s, NH_3^+ Ala^1), 8.14-8.13 (1H, d, *J* 5.5, NH Ala^6), 8.06-8.04 (1H, d, *J* 8.0, αNH AlaT^5), 7.95-7.93 (1H, d, *J* 4.5, αNH Lys^7), 7.87 (3H, br s, NH_3^+ Lys^7), 7.44 (1H, s, C(6)H thymine AlaT^2), 7.40 (1H, s, C(6)H thymine AlaT^5), 7.31 (1H, br s, 1 CONH), 7.01 (1H, br s, 1 CONH), 4.66-4.63 (1H, m, αCH AlaT^2), 4.61-4.57 (1H, m, αCH AlaT^5), 4.27-4.24 (1H, m, αCH Ala^3), 4.24-4.21 (1H, m, αCH Ala^6), 4.17-4.14 (1H, m, αCH Ala^4), 4.13-4.11 (1H, m, αCH Lys^7), 4.09-4.06 (1H, m, 1 βCH AlaT^5), 4.04-4.03 (1H, m, 1 βCH AlaT^2), 3.84-3.79 (2H, m, αCH Ala^1 and 1 βCH AlaT^2), 3.74-3.69 (1H, m, 1 βCH AlaT^5), 2.77-2.74 (2H, m, εCH$_2$ Lys^7), 1.72 (3H, s, C^{5}H$_3$ thymine AlaT^2), 1.71 (3H, s, C^{5}H$_3$ thymine AlaT^5), 1.68-1.64 (2H, m, βCH$_2$ Lys^7), 1.55-1.52 (2H, m, δCH$_2$ Lys^7), 1.34-1.33 (3H, d, *J* 6.5, CH$_3$ Ala^1), 1.31-1.25 (2H, m, γCH$_2$ Lys^7), 1.23-1.21 (6H, 2 d, CH$_3$ Ala^3 and Ala^6), 1.19-1.18 (3H, d, *J* 7.0, CH$_3$ Ala^4). *m/z* (ESI) 820.50 (M + H$^+$; $C_{34}H_{54}N_{13}O_{11}$ requires 820.41).

2*TFA·H-Ala-AlaT-Ala-Aib-AlaT-Ala-Lys(H)-NH$_2$ [T2U]. Synthesis scale: 15 μmol. Activation protocol: 5 eq each of Boc-protected protein amino acid, BOP, HOBt and 15 eq DIEA, 5 eq each of Boc-Aib-OH, HATU, HOAt and 15 eq DIEA or 3 eq each of Boc-AlaT-OH, BOP, HOBt and 10 eq DIEA. Yield: 3.4 mg (26%). HPLC (Varian, 0-50% B in 20 min): t_r 9.10 min. NMR δ_H (500 MHz, d_6-DMSO) 11.24-11.23 (2H, 2 br s, N(3)H thymine AlaT^2 and AlaT^5), 8.86-8.84 (1H, d, *J* 8.0, αNH AlaT^2), 8.43-8.42 (1H, d, *J* 6.0, NH Ala^3), 8.36 (1H, br s, NH Aib^4), 8.21 (3H, br s, NH_3^+ Ala^1), 7.92-7.91 (1H, d, *J* 8.0, αNH AlaT^5), 7.86-7.84 (1H, d, *J* 8.5, αNH Lys^7), 7.84-7.82 (3H, br s, NH_3^+ Lys^7), 7.74-7.72 (1H, d, *J* 7.0, NH Ala^6), 7.50 (1H, d, *J* 1.5, C(6)H thymine AlaT^2), 7.36 (1H, s, C(6)H thymine AlaT^5), 7.22 (1H, br s, 1 CONH), 7.02 (1H, br s, 1 CONH), 4.66-4.62 (1H, m, αCH AlaT^2), 4.49-4.44 (1H, m, αCH AlaT^5), 4.28-4.24 (1H, m, 1 βCH AlaT^5), 4.22-4.21 (1H, m, αCH Ala^3), 4.21-4.19 (1H, m, αCH Ala^6), 4.15-4.11 (1H, m, αCH Lys^7), 4.02-3.98 (1H, m, 1 βCH AlaT^2), 3.92-3.87 (1H, m, 1 βCH AlaT^2), 3.86-3.80 (2H, m, αCH Ala^1 and 1 βCH AlaT^5), 2.78-2.72 (2H, m, εCH$_2$ Lys^7), 1.73 (3H, d, *J* 1.5, C^{5}H$_3$ thymine AlaT^2), 1.71 (3H, s, C^{5}H$_3$ thymine AlaT^5), 1.68-1.65 (2H, m, βCH$_2$ Lys^7), 1.58-1.52 (2H, m, δCH$_2$ Lys^7), 1.36-1.34 (2H, m, γCH$_2$ Lys^7), 1.33-1.32 (3H, d, *J* 7.0, CH$_3$ Ala^1), 1.29 (6H, s, 2 CH$_3$ Aib^4), 1.27-1.25 (3H, d, *J* 7.0, CH$_3$ Ala^6), 1.25-1.24 (3H, d, *J* 7.5, CH$_3$ Ala^3). *m/z* (ESI) 834.49 (M + H$^+$; $C_{35}H_{56}N_{13}O_{11}$ requires 834.42).

Solution-phase syntheses

The syntheses of Z-Aib-AlaT-Aib-O*t*Bu and Z-Aib-AlaT-Aib-OH have been described elsewhere.[11]

Z-AlaT-Aib-Lys(Boc)-NH$_2$. Z-AlaT-OH (0.41 g, 1.2 mmol) was dissolved in anhydrous DMF, cooled in a water/ice bath, and HOAt (0.19 g, 1.4 mmol) and EDC·HCl (0.28 g, 1.5 mmol) were added. H-Aib-Lys(Boc)-NH$_2$ (obtained by catalytic hydrogenolysis of 0.89 g, 1.7 mmol, of the corresponding Z-protected derivative) was added and the pH was adjusted to 8 with NMM. The mixture was stirred at room temperature for 6 days. The solvent was removed at reduced pressure, the resulting oil was taken up with AcOEt, washed with $KHSO_4$ 5%, H_2O, $NaHCO_3$ 2% and H_2O, then desiccated on Na_2SO_4 and evaporated to an oil. The crude product was purified by flash-chromatography (eluent: DCM–MeOH, gradient from 18 : 1 to 9 : 1) to yield a colorless solid (0.46 g, 58%). Mp: 142-145 °C (from PE). $[\alpha]_D^{25} = -35.9$ (*c* = 0.34, MeOH). TLC: R_f1 0.30, R_f2 0.85, R_f3 0.20. IR (cm^{-1}): 3410, 1675, 1521. NMR δ_H (200 MHz, d_6-DMSO) 8.32 (1H, br s, N(3)H thymine), 7.64-7.60, 7.47-7.43 (2H, 2 d, *J* 7.6 and 7.8, αNH AlaT and αNH Lys), 7.35-7.30 (7H, m, C$_6$H$_5$-Z, C(6)H thymine and NH Aib), 7.11 (1H, br s, 1 CONH), 7.00 (1H, br s, 1 CONH), 6.76-6.72 (1H, t, *J* 4.2, εNH Lys), 5.06-4.89 (2H, m, CH$_2$-Z), 4.41-4.33 (1H, m, 1 αCH), 4.23-4.19 (2H, m, 1 αCH and 1 βCH AlaT), 3.64-3.55 (1H, m, 1 βCH AlaT), 2.88-2.80 (2H, m, εCH$_2$ Lys), 1.69-1.65 (4H, m, 4H, C^{5}H$_3$ thymine and 1 βCH Lys), 1.59-1.49 (1H, m, 1 βCH Lys), 1.38-1.33 (16H, m, 1 CH$_3$ Aib, 3 CH$_3$ Boc, γCH$_2$ and δCH$_2$ Lys), 1.30 (3H, s, 1 CH$_3$ Aib).

Z-Aib-AlaT-Aib-Lys(Boc)-NH$_2$. (Z-Aib)$_2$O (0.58 g, 1.26 mmol) dissolved in anhydrous MeCN and cooled in a water/ice bath was added to H-AlaT-Aib-Lys(Boc)-NH$_2$ (obtained by catalytic hydrogenolysis of 0.46 g, 0.65 mmol, of the corresponding Z-protected peptide). The pH was adjusted to 8 with NMM and the mixture was stirred at room temperature for 7 days. The solvent was evaporated at reduced pressure, the residue was taken up with AcOEt, washed with $KHSO_4$ 5%, H_2O, $NaHCO_3$ 2%, and H_2O, desiccated on Na_2SO_4 and evaporated to a solid foam, which was purified by flash-chromatography (eluent: DCM–MeOH, gradient from 15 : 1 to 13 : 1) to afford a colorless solid (0.36 g, 75%). Mp: 124–126 °C (from Et$_2$O–PE). $[\alpha]_D^{25} = -55.6$ (*c* = 0.25, MeOH). TLC: R_f 1 0.25, R_f 2 0.80, R_f 3 0.15. IR (cm^{-1}): 3329, 1675, 1525. NMR δ_H (250 MHz, $CDCl_3$) 9.22 (1H, br s, N(3)H thymine), 8.63-8.62 (1H, d, *J* 3.5, αNH AlaT), 7.92 (1H, s, C(6)H thymine), 7.34-7.33 (5H, m, C$_6$H$_5$-Z), 7.17-7.13 (1H, d, *J* 8.0, αNH Lys), 7.03 (2H, br s, CONH$_2$), 6.59 (1H, br s, NH Aib^3), 5.78 (1H, br s, NH Aib^1), 5.05 (2H, s, CH$_2$-Z), 4.62 (1H, br s, εNH Lys), 4.28-4.15 (3H, m, αCH Lys, αCH and 1 βCH AlaT), 4.06-3.98 (1H, m, βCH AlaT), 3.14-3.09 (1H, m, 1 εCH Lys), 3.03-2.95 (1H, m, 1 εCH Lys), 2.05-1.95 (1H, m, 1 βCH Lys), 1.88 (3H, s, C^{5}H$_3$ thymine), 1.87-1.80 (1H, m, 1 βCH Lys), 1.52-1.43 (13H, m, 3 CH$_3$ Aib, γCH$_2$ and δCH$_2$ Lys), 1.42-1.41 (12H, 2 s, 1 CH$_3$ 1 Aib and 3 CH$_3$ Boc). *m/z* (ESI) 745.38 (M+H$^+$; $C_{35}H_{53}N_8O_{10}$ requires 745.39).

2*TFA·H-(Aib-AlaT-Aib)$_2$-Lys(H)-NH$_2$ [T2U4]. Z-Aib-AlaT-Aib·Oxl (0.26 g, 0.51 mmol, obtained by activation of 0.30 g, 0.58 mmol, of Z-Aib-AlaT-Aib-OH with 0.13 g, 0.64 mmol EDC·HCl) was suspended in anhydrous MeCN (5 mL) and added to H-Aib-AlaT-Aib-Lys(Boc)-NH$_2$ (obtained by catalytic hydrogenolysis of 0.33 g, 0.40 mmol, of the corresponding Z-protected peptide) and the resulting mixture was heated at gentle reflux, with gradual dissolution, for 6 days. The solvent was evaporated at reduced pressure to a reddish oil, which underwent flash chromatography (eluent: $CHCl_3$–MeOH, gradient 15 : 1 to 7 : 1), yielding crude Z-(Aib-AlaT-Aib)$_2$-Lys(Boc)-NH$_2$ (0.13 g, 28%) as a waxy solid. The protected nucleopeptide, *m/z* (ESI) 1110.55 (M + H$^+$; $C_{51}H_{76}N_{13}O_{15}$ requires 1110.56), was used in the next steps without further purification (HPLC analysis, Agilent

20–50% B′ in 30 min: 69% t_r 27.07, 12% t_r 26.45 min, 10% t_r 27.84 min). Crude Z-(Aib-AlaT-Aib)$_2$-Lys(Boc)-NH$_2$ (0.10 g, 0.09 mmol) was N$^{\alpha}$-deprotected by catalytic (Pd) hydrogenolysis in MeOH solution. After removal of the solvent, the residue was dissolved in a DCM–TFA–H_2O mixture (25 : 25 : 1, 1 mL) and stirred for 1 hour at room temperature. The solvent was stripped with N_2, the residue was taken up with DCM and stripped again with N_2 2 times, then taken up with MeOH, yielding crude **T2U4** (0.09 g, 95%), which was purified by preparative HPLC (Shimadzu, 1.5–4.8% B′ in 22 min), then lyophilized, affording 18 mg (18%) of a colorless solid. Mp: 176–178 °C (from H_2O–MeOH). $[\alpha]_D^{25} = -27$ (c = 0.18, MeOH). TLC: R_f1 0.00, R_f2 0.05, R_f3: 0.00. IR (cm^{-1}): 3418, 1674, 1532. HPLC (Agilent, 1.5-6% B′ in 30 min) t_r 18.52 min. NMR δ_H (600 MHz, d_6-DMSO) 11.33-11.31 (2H, 2 br s, 2 N(3)H thymine AlaT^2 and AlaT^5), 8.56 (1H, br s, NH Aib^3), 8.38-8.37 (1H, d, J 6.0, αNH AlaT^2), 8.14 (3H, br s, 1 NH_3^+), 7.94 (1H, br s, NH Aib^4), 7.86-7.85 (1H, d, J 7.2, αNH AlaT^5), 7.64 (4H, br s, NH Aib^6 and 1 NH_3^+), 7.42 (1H, s, C(6)H thymine AlaT^2), 7.33-7.32 (1H, d, J 6.6, αNH Lys^7), 7.31 (1H, s, C(6)H thymine AlaT^5), 7.04 (1H, br s, 1 CONH), 6.94 (1H, br s, 1 CONH), 4.68 (1H, m, αCH AlaT^2), 4.36 (1H, m, αCH AlaT^5), 4.28-4.26 (1H, m, 1 βCH AlaT^2), 4.11 (1H, m, αCH Lys^7), 4.01-3.96 (2H, m, βCH$_2$ AlaT^5), 3.73 (1H, m, 1 βCH AlaT^2), 2.73 (2H, m, εCH2 Lys^7), 1.80 (1H, m, 1 βCH Lys^7), 1.72 (3H, s, C^{5}H$_3$ thymine AlaT^2), 1.71 (3H, s, C^{5}H$_3$ thymine AlaT^5), 1.59 (1H, m, 1 βCH Lys^7), 1.48 (2H, m, δCH$_2$ Lys^7), 1.42 (3H, s, 1 CH$_3$ Aib^1), 1.40 (6H, s, 1 CH$_3$ Aib^1 and 1 CH$_3$ Aib^6), 1.36 (3H, s, 1 CH$_3$ Aib^3), 1.34-1.33 (7H, m, 1 γCH Lys^7, 1 CH$_3$ Aib^3 and 1 CH$_3$ Aib^6), 1.29-1.27 (7H, m, 1 γCH Lys^7 and 2 CH$_3$ Aib^4). m/z (ESI) 876.52 (M+H$^+$;$C_{38}H_{62}N_{13}O_{11}$ requires 876.47).

Biological tests

Prior to biological tests, purified peptides were dissolved in 1 mM HCl and re-lyophilized in order to replace TFA counter-anions with chloride.

Stability test.[25] Nucleo-heptapeptides (about 1.0 mg) were dissolved in PBS buffer at 10 mg mL^{-1} concentration, an equal volume of freshly prepared mouse serum was added and time count started. The solutions were incubated at 37 °C in a thermostatic bath and sampled immediately after serum addition (as a control) and 1, 2, 4, 8, 24 and 48 h thereafter. For each time-point, 20 µL solution were taken, added with 2 µL TFA, with precipitation of serum proteins, diluted with 80 µL phosphate buffered saline (PBS), then centrifuged (10 min at 5000 rpm). The supernatant was diluted with 120 µL H_2O milliQ and analyzed by HPLC (Varian, 0–50% B in 20 min).

MTS cytotoxicity test. Lymphoid tumor Raji cells were used for nucleopeptide cytotoxicity evaluation by the 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2*H*-tetrazolium (MTS) test.[20] Cells were seeded at a density of 50 000 cells per well in 96-well plates, incubated in 10% fetal bovine serum/Roswell Park Memorial Institute medium (FBS/RPMI) and treated for 24 h at 37 °C in 5% CO_2 with **T2U4** at concentrations up to 100 µM. Doxorubicin at the same concentrations was used as a positive control. After treatment, the cells were incubated for 2 h at 37 °C with 20 µL of MTS reagent solution (Promega, Madison, USA). Absorbance was measured at 490 nm wavelength using a Wallac Victor^2 microplate reader (Waltham, MC, USA). Results are the average of three independent experiments.

Acknowledgements

We wish to thank the French-Italian University (UIF-UFI) for the PhD scholarship of P. G.-B. and a travel grant.

References

1 J. Bonner, M. E. Dahmus, D. Fainbrough, R. C. Huang, K. Marushige and D. H. T. Tuan, *Science*, 1968, **159**, 47.
2 S. M. Beiser and B. F. Erlanger, *Cancer Res.*, 1966, **26**, 2012.
3 P. Lohse, B. Oberhauser, B. Oberhauser-Hofbauer, G. Baschang and A. Eschenmoser, *Croat. Chem. Acta*, 1996, **69**, 535.
4 (*a*) U. Diederichsen, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 445; (*b*) U. Diederichsen, *Angew. Chem., Int. Ed. Engl.*, 1997, **36**, 1886; (*c*) M. F. H. Hoffmann, A. M. Brückner, T. Hupp, B. Engels and U. Diederichsen, *Helv. Chim. Acta*, 2000, **83**, 2580; (*d*) U. Diederichsen, D. Weicherding and N. Diezemann, *Org. Biomol. Chem.*, 2005, **3**, 1058.
5 (*a*) A. M. Brückner, M. Garcia, A. Marsh, S. H. Gellman and U. Diederichsen, *Eur. J. Org. Chem.*, 2003, 3555; (*b*) P. Chakraborty and U. Diederichsen, *Chem.–Eur. J.*, 2005, **11**, 3207; (*c*) A. Weiß and U. Diederichsen, *Eur. J. Org. Chem.*, 2007, 5531.
6 (*a*) P. Garner, S. Dey and Y. Huang, *J. Am. Chem. Soc.*, 2000, **122**, 2405; (*b*) Y. Huang, S. Dey, X. Zhang, F. Sönnichsen and P. Garner, *J. Am. Chem. Soc.*, 2004, **126**, 4626.
7 Y. Ura, J. M. Beierle, L. J. Leman, L. E. Orgel and R. Ghadiri, R., *Science*, 2009, **325**, 73.
8 P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497.
9 P. Geotti-Bianchini, J. Beyrath, O. Chaloin, F. Formaggio and A. Bianco, *Org. Biomol. Chem.*, 2008, **6**, 3661.
10 *Nucleic Acid Structure*, ed. S. Neidle, Oxford University Press, Oxford, 1999.
11 P. Geotti-Bianchini, M. Crisma, C. Peggion, A. Bianco and F. Formaggio, *Chem. Commun.*, 2009, 3178.
12 C. Toniolo, M. Crisma, F. Formaggio and C. Peggion, *Biopolymers*, 2001, **60**, 396.
13 P. Chaltin, E. Lescrinier, T. Lescrinier, J. Rozanski, C. Hendrix, H. Rosemeyer, R. Busson, A. Van Aershot and T. Herdewijn, *Helv. Chim. Acta*, 2002, **85**, 2258.
14 K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, 1986.
15 C. Toniolo and E. Benedetti, *Trends Biochem. Sci.*, 1991, **16**, 350.
16 N. Greenfield and G. D. Fasman, *Biochemistry*, 1969, **8**, 4108.
17 N. Sreerama and R. W. Woody, Circular dichroism of peptides and proteins, in *Circular dichroism: principles and applications*, ed. N. Berova, K. Nakanishi and R. W. Woody, Wiley, New York, 2000, pp 601–620.
18 H. Yamaguchi, H. Kodama, S. Osada, F. Kato, M. Jelokhani-Niharaki and M. Kondo, *Biosci., Biotechnol., Biochem.*, 2003, **67**, 2269.
19 (*a*) J. D. Sadowsky, J. K. Murray, Y. Tomita and S. H. Gellman, *ChemBioChem*, 2007, **8**, 903; (*b*) L. P. Miranda, K. A. Winters, C. V. Gegg, A. Patel, J. Aral, J. Long, J. Zhank, S. Diamond, M. Guido, S. Stanislaus, M. Ma, H. Li, M. J. Rose, L. Poppe and M. M. Véniant, *J. Med. Chem.*, 2008, **51**, 2758; (*c*) M. De Zotti, B. Biondi, F. Formaggio, C. Toniolo, L. Stella, Y. Park and K.-Soo Hahm, *J. Pept. Sci.*, 2009, **15**, 615.
20 A. H. Cory, T. C. Owen, J. A. Barltrop and J. G. Cory, *Cancer Commun.*, 1991, **3**, 207.
21 M. Von Arx, M. Faupel and M. Brugger, *J. Chromatogr., A*, 1976, **120**, 224.
22 J. Neimark and J.-P. Briand, *Pept. Res.*, 1993, **6**, 219.
23 E. Kaiser, R. L. Colescott, C. D. Bossinger and P. I. Cook, *Anal. Biochem.*, 1970, **34**, 595.
24 N. Fujii, A. Otaka, O. Kamura, K. Akaji, S. Funakoshi, Y. Hayashi, Y. Kuroda and H. Yajima, *J. Chem. Soc., Chem. Commun.*, 1987, 274.
25 G. Guichard, N. Benkirane, G. Zeder-Lutz, M. H. van Regenmortel and J.-P. Briand, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 9765.