Pergamon

Tetrahedron 55 (1999) 13251–13264

TETRAHEDRON

TOWARDS NUCLEOPEPTIDES CONTAINING ANY TRIFUNCTIONAL AMINO ACID

Jordi Robles, Maite Beltrán, Vicente Marchán, Yolanda Pérez, Isaura Travesset, Enrique Pedroso and Anna Grandas*

Departament de Química Orgànica, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1-11, E-08028 Barcelona, Spain

Received 24 June 1999; revised 26 August 1999; accepted 9 September 1999

Abstract. Nucleopeptides incorporating, besides the linking residue, other trifunctional amino acids such as lysine, arginine, tryptophan, serine, threonine and tyrosine have been obtained following stepwise solid-phase assembly. The best alternatives for the protection of the side chains of these trifunctional amino acids are discussed. The use of different solid supports (polystyrene-*co*-1%-divinylbenzene, controlled pore glass and polyethylene glycol-polystyrene) and nucleopeptide-resin linkers has also been evaluated.

Introduction

Nucleic acids are covalently linked to proteins in two well-studied biological processes: the protein-primed replication of the genetic material of some viruses[1] and the topoisomerase-mediated modification of the topological state of DNA.[2] In both cases the covalent linkage is established through a phosphodiester bond between the side chain of a hydroxylated amino acid (serine, threonine or tyrosine) and either the 5' or the 3' end of an oligonucleotide chain.

The synthesis of phosphodiester-linked peptide-oligonucleotide hybrids, commonly called nucleopeptides, has attracted the attention of several research groups,[3-6] first because they can be used as models to study the replication and topoisomerisation processes referred to above. Information on the structure of these peptide-oligonucleotide hybrids may also be of value in the search for new drugs that might interfere with those biological processes and thus be used as antiviral or anti-cancer drugs. Covalently linked peptide-oligonucleotide hybrids, either nucleopeptides or conjugates with any other kind of linkage between the two components, are also attractive because of their potential use in antisense or antigene therapy.[7-15]

Nucleopeptides have been prepared using either convergent[3,4,6,16-19] or stepwise approaches,[5,20,21] and the formation of the amino acid-nucleoside covalent linkage has always been carried out using maximal protection tactics. The choice of protecting groups is the most difficult point in the overall strategy, since the deprotection conditions must be compatible with the presence of acid-labile glycosidic bonds and a base-labile amino acid-nucleoside linkage in the target molecule.[22,23] Various protection schemes have been explored, using both chemically-[3-5,16-21] and enzymatically-cleavable[6] protecting groups.

Nucleopeptides have been synthesized in our laboratory using stepwise solid-phase procedures[20] in which the assembly of the peptide moiety on a solid matrix is followed by elongation of the oligonucleotide at the side chain of the desired hydroxylated amino acid. With respect to the protection scheme, the acid lability of glycosidic linkages precludes the use of strong acid final deprotection conditions, so we prefer to use a set of base-labile groups for permanent protection and acid-labile groups for the temporary protection of α-amine and

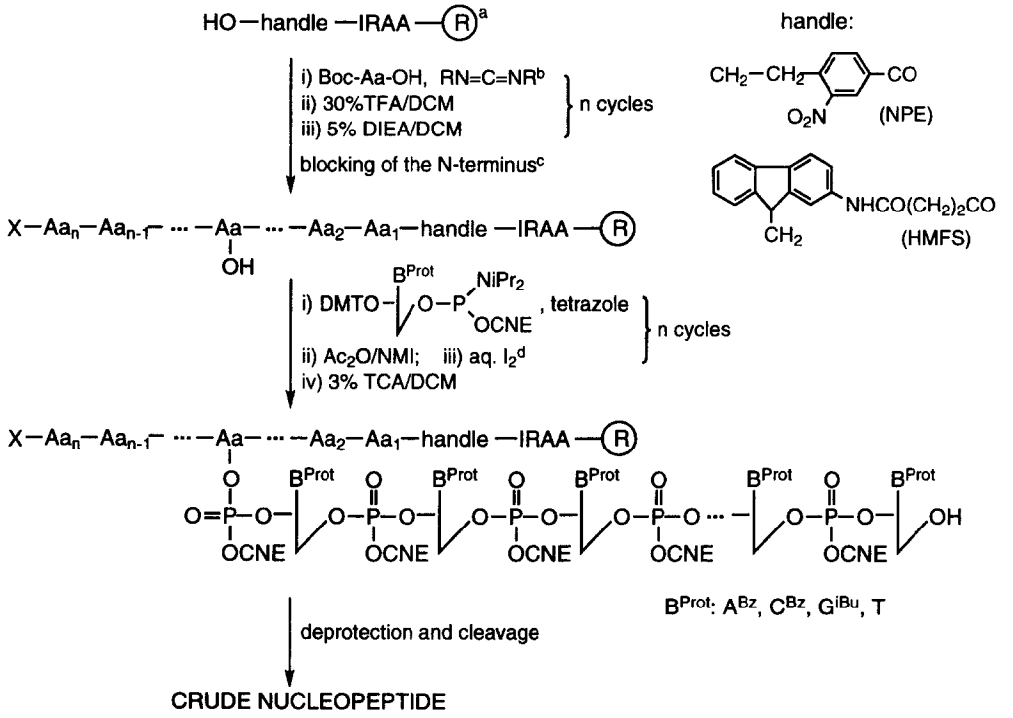
0040-4020/99/$ - see front matter
PII: S0040-4020(99)00815-7

5' (or 3', depending on the synthons used) hydroxyl groups.[5,20] This combination of protecting groups has the advantage that most deoxynucleosides and amino acid derivatives can be obtained from commercial sources.

Here we report on our results on the preparation of nucleopeptides containing lysine, arginine, tryptophan, and hydroxylated amino acids besides the linking residue, and we also compare different solid supports and handles.

Results and discussion

The main purpose of this study was to add lysine, arginine, tryptophan, serine, threonine and tyrosine[24] to the list of trifunctional amino acids (aspartic acid,[3,5] histidine[25]) so far introduced in nucleopeptides besides the linking residue. The molecules were fully assembled on a solid support (Scheme 1), most often with homoserine as the linking amino acid,[26] polystyrene-1%-divinylbenzene as the solid matrix, and either the NPE or HMFS



Scheme 1. General synthetic scheme for the stepwise solid-phase assembly of nucleopeptides. [a]IRAA: internal reference amino acid (Leu); R: *p*-methylbenzhydrylamine resin (polystyrene); [b]dicyclohexylcarbodiimide or diisopropylcarbodiimide; [c]N-terminus blocking was effected by acylation with either acetic anhydride or phenylacetic anhydride; if the product was to be obtained unblocked N^{α}-Fmoc-protected N-terminal amino acids were used; [d]aqueous iodine was replaced by *t*-BuOOH in some cases (see experimental part); sulfurization with 3*H*-1,2-benzodithiole-3-one 1,1-dioxide (Beaucage reagent) was used instead of oxidation to obtain the phosphorothioated analogues.

handles.[27-29] The amino acid side chain protecting groups chosen were always shown to be stable to the reagents used in oligonucleotide synthesis (phosphite triester approach). All protecting groups were labile to at least one of the reaction conditions used for the final deprotection (see experimental section).

Lysine-containing nucleopeptides

The commercially available N^{α}-Boc-N^{ε}-Fmoc- and N^{α}-Boc-N^{ε}-trifluoroacetyl-lysine derivatives both seemed appropriate for use in nucleopeptide synthesis. They had already been used in the preparation of peptide-oligonucleotide conjugates,[8,30,31] and side reactions had been observed during the final cleavage of N^{ε}-Fmoc-containing conjugates.[30]

No essential differences between the two protecting groups were found during the assembly of Ac-Gly-Ala-Hse(p$^{3'}$dACTAGT)-Lys-Val-OH (**1**), and both nucleopeptide crudes showed essentially the same HPLC profile. We then synthesized a more complex lysine-rich molecule, Boc-Hse(p$^{3'}$dATATTTGTTACTCTGT)-Lys-Lys-Lys-Lys-Lys-OH (**2**). The oligonucleotide moiety is complementary to the HIV-1 reverse trancriptase region coding for the 138 glutamic acid that mutates to lysine in HIV-1-infected cells treated with TSAO derivatives,[32] and its transport into cells might be facilitated by covalent linkage to the cationic peptide.[33-36] We found that, surprisingly, incorporation of the first (C-terminal) lysine residue onto the resin was more difficult in the case of the N^{ε}-Tfa-protected derivative, whereas the pentapeptide elongation proceeded more easily with the Tfa than with the Fmoc side chain protecting group (coupling of the fourth and fifth Lys(Fmoc) had to be repeated). The coupling yield of the first nucleoside phosphoramidite onto Boc-Hse-[Lys(Tfa)]$_5$-resin was slightly better than onto Boc-Hse-[Lys(Fmoc)]$_5$-resin. Even though these small differences did not discriminate between the two protecting groups, when a complex base-stable hybrid is to be obtained, the trifluoroacetyl protecting group may be more appropriate, since the impurities formed by elimination of the Fmoc group will not accompany the target molecule.

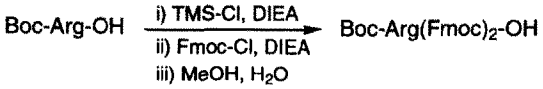
When nucleopeptide **2** was hybridized to $^{5'}$TATAAACAATGAGACA$^{3'}$ in the absence of magnesium salts, the melting temperature (Tm) was 2 °C higher than the Tm of the duplex formed by oligonucleotides $^{5'}$TGTCTCATTGTTTATA$^{3'}$ and $^{5'}$TATAAACAATGAGACA^{3} (50 °C and 48 °C, respectively). Thus, attachment of pentalysine to the oligonucleotide had a relatively small effect on its hybridization to a complementary sequence, the average value being 0.4 °C increase in Tm per lysine residue. Other authors have described similar results (Tm~0.4-0.7 °C per lysine or ornithine residue)[33,37-39] and, to our knowledge, only in one case was the effect of the polycationic peptide higher (2.1 °C/lysine residue).[40] In the presence of magnesium salts, both the nucleopeptide-oligonucleotide and the oligonucleotides duplex had the same Tm value (55 °C).

Arginine-containing nucleopeptides

Even though the strongly basic guanidine group of the arginine residue remains protonated in most reaction conditions, the side chain of arginine is almost always protected in peptide synthesis.[41,42] This precludes side reactions such as the acylation of free guanidine groups, which may account for the replacement of arginine by ornithine, or intramolecular reactions that may give rise to 1,3-diazacycloheptanes or δ-lactams.[41,42]

Commercially available Boc-arginine derivatives are unsuitable for preparing nucleopeptides using stepwise solid-phase synthesis, since the side chain protecting groups commonly used in peptide synthesis are removed in acidic conditions that would cause depurination at the oligonucleotide chain. Consequently, we synthesized a new N^{α}-Boc-arginine derivative, Boc-Arg(Fmoc)$_2$-OH. This product was prepared following

essentially previously described methodology,[43,44] by temporary silylation with trimethylsilyl chloride and diisopropylethylamine and reaction with an excess of Fmoc-Cl in the presence of the same base (Scheme 2). Slight modifications in the synthesis procedure (see experimental section) lead to failure in obtaining the N^{α}-Boc-N^{G}-bisFmoc-arginine derivative. To evaluate the suitability of using this derivative we prepared nucleopeptide Phac-Ala-Hse(p[3']dACTAGT)-Gly-Arg-Val-OH (**3**).



Scheme 2. Synthesis of the arginine derivative.

Once Phac-Ala-Hse-Gly-Arg(Fmoc)$_2$-Val-O-NPE-Leu-resin had been assembled, an amino acid analysis was carried out, as usual, to check that all amino acids had been incorporated correctly. Surprisingly, the relative proportion of arginine was much lower than expected (0.3-0.5 instead of 1 depending on the hydrolysis conditions), and a certain proportion of ornithine (0.2-0.3) was also present. Acid hydrolysis and amino acid analysis of Boc-Arg(Fmoc)$_2$-OH afforded a mixture of arginine and ornithine, which indicated that decomposition of Arg(Fmoc)$_2$ in the acidic peptide hydrolysis mixture was the origin of the presence of ornithine. An aliquot of Phac-Ala-Hse-Gly-Arg(Fmoc)$_2$-Val-O-NPE-Leu-resin was treated with tetrabutylammonium fluoride, and the peptide crude was compared with the Phac-Ala-Hse-Gly-Arg-Val-OH obtained from HF treatment of Phac-Ala-Hse-Gly-Arg(Tos)-Val-O-Pam-resin. Reversed-phase HPLC analysis showed that the main peaks of both samples coeluted, further supporting the correct incorporation of Arg(Fmoc)$_2$ in the peptide sequence.

Even though Boc-Arg(Fmoc)$_2$-OH had been shown to be stable to all the reagents used in nucleotide incorporation cycles, for the preparation of **3** oxidation with *t*-BuOOH gave better results than oxidation with aqueous iodine.

Tryptophan-containing nucleopeptides

Protection of the indole ring prevents both oxidation and alkylation during elimination of N^{α}-Boc protecting groups, and the latter side reaction can also be circumvented by the addition of suitable scavengers.[45] Boc-tryptophan was shown not to be stable to iodine, but *t*-BuOOH had no effect on the unprotected indole system. The indole moiety, even unprotected, did not react with dimethoxytrityl cations.

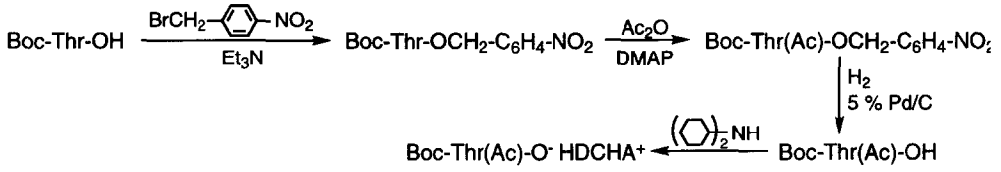
To synthesize nucleopeptide H-Trp-Val-Hse(p[3']dACTAGT)-Gly-OH (**4**) we prepared two peptide-resins, Fmoc-Trp-Val-Hse-Gly-HMFS-Leu-resin and Fmoc-Trp(For)-Val-Hse-Gly-HMFS-Leu-resin. Even though the iodine oxidizing reagent was replaced by 1 M *t*-BuOOH during the oligonucleotide chain assembly, the comparison of the HPLC traces of crude nucleopeptide **4** obtained using either unprotected or N^{In}-formylated tryptophan derivatives shows that protection of the indole ring affords better results (both the number and the proportion of impurities, with respect to the target product, were lower), and is thus recommended. A nucleopeptide lacking the tryptophan and valine N-terminal residues was isolated from all nucleopeptide crudes (**4** and **5**). We were not able to explain its formation, but we realized that the amount of truncated product was smaller when the oligonucleotide was assembled on the N^{In}-formyl protected peptide-resin.

The fully phosphorothioated analog H-Trp-Val-Hse(S-p$^{3'}$dACTAGT)-Gly-OH (**5**) was obtained after assembly of the hexanucleotide on Fmoc-Trp-Val-Hse-Gly-HMFS-Leu-resin replacing the oxidation reaction by a sulfurization step (the unprotected indole ring was shown not to react with the Beaucage sulfurizing reagent,[46] 3*H*-1,2-benzodithiol-3-one 1,1-dioxide).

Before oligonucleotide elongation, an aliquot of the Fmoc-Trp-Val-Hse-Gly-HMFS-Leu-resin was treated with concentrated aqueous ammonia/dioxane 1:1 at 55 °C for 17 h and the reaction crude was analyzed by HPLC-electrospray MS. Surprisingly, an impurity lacking the glycine residue was detected (20%) besides the expected tetrapeptide. Since treatment of pure peptide H-Trp-Val-Hse-Gly-NH_2 with the ammonia solution gave rise to about 20% of H-Trp-Val-Hse-OH, we believe that the glycine-lacking molecule was produced at the peptide level during the cleavage and deprotection treatment, through formation of homoserinelactone and cleavage of the homoserine-glycine peptide bond. Obviously, this reaction cannot take place when the whole nucleopeptide has been assembled, since the homoserine hydroxyl group is no longer free. This result, which once again shows that homoserine is very prone to lactonisation, leads us to disregard the possibility of preparing nucleopeptides with a C-terminal homoserine, since in this case an ester linkage would have to be established between the homoserine carboxyl function and hydroxyl groups on the solid support. This ester group could be cleaved much more easily than amide bonds and produce the undesired detachment of peptide chains from the solid support even under milder conditions.

Nucleopeptides with hydroxylated amino acids besides the linking residue

The O-acetyl derivatives of Boc-serine and Boc-tyrosine are commercially available, and Boc-Thr(Ac)-OH was prepared (Scheme 3) as follows: i) reaction of Boc-threonine with *p*-nitrobenzyl bromide and triethylamine; ii) protection of the side chain hydroxyl group by acetylation, and iii) deprotection of the carboxyl group by hydrogenolysis in the presence of 5% Pd/C. Reaction with dicyclohexylamine afforded a solid dicyclohexylammonium salt. From these three amino acid derivatives we synthesized nucleopeptide Phac-Hse(p$^{3'}$dACTAGT)-Ala-Thr-Ser-Val-Tyr-Gly-OH (**6**).



Scheme 3. Synthesis of the threonine derivative.

Since the phenolic hydroxyl is a good leaving group, we checked for migration of the phosphate linkage from the tyrosine residue to the side chain hydroxyl group of another amino acid in the hybrid molecule by comparing the ^{31}P-NMR spectra of nucleopeptides Phac-Ser(p$^{3'}$dTATAT)-Ala-Ser-Val-OH (**7**), Phac-Tyr(p$^{3'}$dTATAT)-Ala-Tyr-Val-OH (**8**) and Phac-Tyr(p$^{3'}$dTATAT)-Ala-Ser-Val-OH (**9**) (Table 2). Such a side reaction would afford, in the case of **7** and **8**, constitutional isomers with the same linking residues, whereas migration of the phosphodiester bond in **9** would convert a tyrosine-nucleoside into a serine-nucleoside linkage.

The two signals observed in the ^{31}P-NMR spectrum of **8** (δ -1.3 and -5.7 ppm, relative proportion ~4:1 in both cases) correspond to internucleoside dialkyl phosphates and to the aryl-alkyl tyrosine-nucleoside

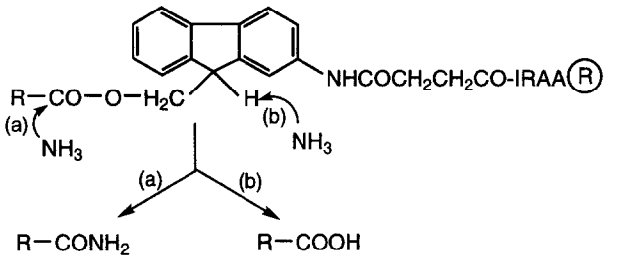
phosphodiester bond, respectively. The detection of a single broad signal in the case of **7** (δ -3.7 ppm) is consistent with an all dialkyl phosphate-containing molecule. In the case of **9**, the ^{31}P-NMR spectrum showed two signals with the same chemical shifts and in the same relative proportion as in the case of the **8**, thus indicating that the molecule contained dialkyl phosphates as well as an aryl-alkyl phosphate linkage. Had there been migration, we should have obtained a serine-nucleopeptide and observed a single peak instead of two. This result indicates that tyrosine-nucleopeptides remain stable in the presence of free hydroxyl groups.

The choice of the linker and the solid support

In this study, both 2-(2-nitrophenyl)ethyl and 9-fluorenylmethyl esters (NPE and HMFS handles, respectively[27,28]) were used to link the nucleopeptide to the solid matrix.

The greater stability of the nitrophenylethyl ester linkage offers the possibility of elimination of Fm and Fmoc protecting groups before the molecule is detached from the resin, which would reduce the amount of non-nucleopeptide UV-absorbing products in the crude product.[30] Yet, we believe that the amount of piperidine used in this treatment should be reduced to *ca.* 2-5 % to prevent undesired cleavage of the nitrophenylethyl ester bond, with the concomitant loss of nucleopeptide chains.

The fluorenylmethyl ester bond is more labile than the nitrophenylethyl ester bond, but sufficiently stable for the stepwise solid-phase synthesis of nucleopeptides. We have tried, on many occasions, to use a single ammonia treament for the final deprotection and cleavage of HMFS-linked nucleopeptides, and found that two products are often obtained, the C-terminal acid and the carboxamide. Cleavage of the peptide-resin bond may take place through a β-elimination mechanism, which affords the C-terminal acid, or by nucleophilic addition of ammonia to the carbonyl ester group leading to the carboxamide (Scheme 4). So far, the nucleopeptide detached from the resin by β-elimination (C-terminal acid) has always been the main product in the crude, the acid/carboxamide ratio ranging between 2:1 and 1.25:1 (this study, ref. 47). Nevertheless, we feel that the ratio of nucleopeptide acid and amide also depends on the nature of the side chain of the C-terminal amino acid.



Scheme 4. Base- or nucleophile-promoted cleavage of the peptide-resin bond yields the target product as the C-terminal acid or C-terminal carboxamide, respectively.

Detachment from the solid support by reaction with tetrabutylammonium fluoride affords exclusively the C-terminal acid nucleopeptide, irrespective of the handle used, but this treatment may degrade base-labile nucleopeptides to a greater extent than ammonia treatment at room temperature (which is only appropriate if nucleobase protecting groups are more labile than the standard ones[48]). On the other hand, the tyrosine-nucleoside phosphodiester bond may be cleaved during the fluoride treatment. It should be evaluated, in those cases, under which deprotection conditions (tetrabutylammonium fluoride+ammonia, or ammonia only) the

largest amount of target product is obtained, taking into account that the presence of the C-terminal carboxamide nucleopeptide accompanying the target product may not interfere with its subsequent use.[49]

Still concerning the degree of purity of the nucleopeptides, we have found that apparently highly heterogeneous crudes are sometimes obtained when the molecule has been treated with tetrabutylammonium fluoride, and that the HPLC profile improves considerably when the product is subjected to cation exchange chromatography through Dowex resins, and tetrabutylammonium cations are replaced, for instance, by sodium or ammonium cations.

Three types of solid support were used for the assembly of the pentalysine sequence of nucleopeptide **2**: our standard polystyrene-*co*-1%-divinylbenzene resin, the glass beads commonly used in oligonucleotide synthesis and a polyethylene glycol-polystyrene copolymer (both the PerSeptive Biosystems PEG-PS support and the Rapp Polymere TentaGel resin were tested). In this section we only refer to the results of the syntheses in which Boc-Lys(Tfa)-OH was used in combination with any of the three solid matrices.

Significant difficulties in coupling all lysine residues (based on ninhydrin tests) were found when the synthesis of the pentalysine sequence was attempted on polyethylene glycol-polystyrene copolymers. This, and the fact that only a tiny amount of peptide could be detached from the resin under different conditions, led us to abandon this type of solid support.

Assembly of the nucleopeptide on the glass beads proceeded smoothly, suggesting that both this support and polystyrene are suitable for assembling the hybrid molecule. Detachment of the nucleopeptide from the fluoride-incompatible glass beads was carried out by treatment with LiOH, and complete deprotection was achieved after an overnight ammonia treatment. PAGE analysis of the crudes discriminated between the two solid supports. A main band corresponding to the target product was observed for the polystyrene-assembled nucleopeptide, whilst two main bands appeared in the lane of the product synthesized on controlled pore glass. One was the target nucleopeptide, and the impurity contained four lysines instead of five.

In summary, with regard to the solid matrix we rule out the use of polyethylene glycol-polystyrene copolymers and CPG. We recommend that nucleopeptides should be assembled on polystyrene solid supports.

Concluding remarks

Phosphodiester-linked peptide-oligonucleotide hybrids can be easily obtained by stepwise solid-phase assembly on polystyrene-divinylbenzene solid supports, and most nucleoside and amino acid derivatives can be obtained from commercial sources (only Boc-Arg(Fmoc)$_2$-OH and Boc-Thr(Ac)-OH were synthesized in our laboratory). Acid residues (aspartic and glutamic acids), all the basic residues (histidine, lysine and arginine), tryptophan and the hydroxylated amino acids (serine, threonine and tyrosine) can be incorporated into nucleopeptides. Either serine, threonine, tyrosine or homoserine may be used as the linking residue, although to avoid possible side reactions an exception is that homoserine should not be placed at the C-terminal position.

Tyrosine- and homoserine-nucleopeptides are stable to bases, so there is no restriction in the choice of permanent protecting groups for the nucleobases, the side chain of trifunctional amino acids or the carboxyl group of the C-terminal residue. For the preparation of serine- and threonine-nucleopeptides, protecting groups that are removable under mild basic conditions such as isobutyryl (or acetyl) and dimethylaminomethylene are appropriate to block the exocyclic amines of cytosine and guanine, respectively, and the lysine side chain is best protected with the Fmoc group and not the trifluoroacetyl group. These few changes allow the final deprotection

and cleavage to be carried out under conditions in which the extent of base-promoted degradation of the target molecule is kept to a minimum.

We are currently evaluating the best alternatives for the introduction of asparagine, glutamine, methionine and cysteine in nucleopeptides.

Experimental section

Fmoc-L-amino acids, Boc-L-amino acids and resins for peptide synthesis (MBHA resin and Boc-Val-Pam-resin) were obtained from Bachem and Novabiochem. The 3'-phosphoramidite derivatives of DMT-dA^{Bz}, DMT-dC^{Bz}, DMT-dG^{iBu} and DMT-T, and 5'DMT-nucleoside-CPG were purchased from Glen Research Corporation. Solid-phase syntheses were performed manually in a polypropylene syringe fitted with a polyethylene disc (peptides), or on the 380B Applied Biosystems synthesizer (oligonucleotides and nucleopeptides).

Reversed-phase HPLC analyses were performed on either Spherisorb or Nucleosil C18 columns (25 x 0.5 cm, 10 μm), using linear gradients of 0.045% TFA in H_2O and 0.036% TFA in ACN for peptide analyses, and 0.01 M aq triethylammonium acetate and ACN/H_2O (1:1) for oligonucleotide and nucleopeptide analyses. For the anion exchange HPLC a Partisphere SAX column (10 x 0.5 cm, Whatman) was used, eluting with linear gradients of NaH_2PO_4, pH=6.4 in ACN/H_2O 3:7 (solvent A: 0.01 M NaH_2PO_4, solvent B: 0.2 M NaH_2PO_4). Octadecylsyloxane-Vydac resin was used for purification by reversed-phase MPLC.

Electrophoreses were run under denaturing conditions (7 M urea) on 20 % polyacrylamide gels, at 500-750 V for 3-4 h. At analytical scale, oligonucleotides and nucleopeptides were detected by reaction with "stains-all" dye (Sigma). In preparative gels products were visualised by UV-shadowing, and isolated by the "crush and soak" method.[50] The target products were desalted by elution through Sep-Pak cartridges.

Amino acid analyses were performed on a Beckman System 6300 analyzer. Hydrolyses of peptide-resins, peptides and nucleopeptides were carried out in 12 M HCl/propionic acid (1:1) either at 150 °C for 60-90 min or at 110 °C for 24 h (homoserine, serine and threonine are to some extent degraded under these hydrolysis conditions). The tryptophan content was determined after hydrolysis with 3 M *p*-toluenesulfonic acid (12 h, 110 °C) in the presence of 0.2 % of 3-(2-aminoethyl)indole (~70 % of homoserine may be recovered after this hydrolysis procedure). Unusually high values of glycine in nucleopeptides have been attributed to contamination from nucleobase degradation during acid hydrolysis (see, for example, reference 8).

Digestions with snake venom phosphodiesterase (EC 3.1.4.1, Sigma) and alkaline phosphatase (EC 3.1.3.1, Sigma) were carried out on 0.5-1.0 OD_{260} of oligonucleotides or nucleopeptides. The digestion mixtures were analyzed by reversed-phase HPLC using oligonucleotide conditions (10 min 5% B, linear gradient from 5 to 60% B in 10 min).

Unless otherwise indicated, ^1H, ^{13}C and ^{31}P NMR spectra were recorded at 200, 50 and 81 MHz, respectively, in $CDCl_3$. Mass spectrometric analysis was carried out using Hewlett-Packard 5988 A (FAB), VG-Quattro (ES) or Bruker Biflex™III (MALDI-TOF) instruments. Calculated mass values for neutral compounds (M) are indicated in all casses.

General procedure for the preparation of nucleopeptides

Incorporation of the internal reference amino acid (leucine, unless otherwise indicated) onto *p*-methylbenzhydrylamine-resin (f=0.5-0.6 mmol/g) was carried out by reaction with 0.5 equivalents of both the Fmoc-amino acid derivative and DCC (15 min) in order to achieve partial acylation of polymeric amine groups (when the lowly substituted H_2N-LCAA-CPG was the solid matrix, to achieve maximum coupling yield an excess of the internal reference amino acid was used). The substitution degree was determined on an aliquot by UV spectrophotometric quantitation of the amount of N-(9-fluorenylmethyl)piperidine formed upon deprotection with piperidine, and, if suitable for oligonucleotide synthesis,[51] unreacted amine groups were capped by reaction with Ac_2O and DIEA (substitution degree values ranged between 0.19 and 0.27 mmol/g). The Fmoc group was eliminated and the bifunctional linker (either 4-(2-hydroxyethyl)-3-nitrobenzoic acid, H-NPE-OH,[27] or N-2-(9-hydroxymethylfluorenyl)succinamic acid, H-HMFS-OH[28]) coupled in the presence of DCC or DIPC (3 equivalents of each reagent, 2-4 h). Hydroxyl groups were esterified by reaction with the C-terminal Boc-amino acid and DCC or DIPC (10 equiv) in the presence of 0.5 equiv of DMAP for 2 h. The coupling yield was determined from the relative proportion [C-terminal residue]/[internal reference amino acid] after acid hydrolysis and amino acid analysis (typical values: 0.91-0.95), and unreacted hydroxyl groups were capped by reaction with benzoyl chloride (or acetic anhydride) and pyridine.

The remaining amino acids of the peptide sequence were incorporated following the standard procedures for

solid-phase peptide assembly (3 equiv of Boc-amino acid and DCC or DIPC and 1-1.5 h reaction time were used for the coupling step), with the addition of an equimolar amount of HOBt in the incorporation of the homoserine[26] and arginine derivatives. Except in the case of nucleopeptide **4**, acylation of the N-terminal amine group was carried out by reaction with acetic or phenylacetic anhydride, as desired. Homoserine derivatives (Boc-Hsc(DMT)-O$^-$ HTEA$^+$ and Phac-Hse(DMT)-O$^-$ HTEA$^+$) were prepared as previously described.[26]
Oligonucleotide elongation was carried out (4-10 μmol-scale) following standard phosphite triester procedures with the minor modifications that render this methodology compatible with the use of polystyrene supports.[51] 0.15-0.2 M solutions of commercially available 3'-phosphoramidite derivatives (5'-DMT-N-P(OCNE)NiPr$_2$, N=T, A^{Bz}, C^{Bz}, G^{iBu}) in anhydrous DCM and 0.7-0.8 M solutions of tetrazole in anhydrous THF were used for the coupling step (30 min coupling time for the first nucleoside and 15 min for the others). Coupling yields for the incorporation of the first nucleoside derivative typically ranged between 80 and 95 %, and the average coupling yield was 95-99 %. The 5'-DMT group was removed from the nucleopeptide-resin.
Nucleopeptides were generally detached from the solid support by treatment with 0.05 M TBAF in anhydrous THF (30 min, 10-20 equivalents of fluoride per reactive group). The resin was subsequently washed with THF and MeOH (the filtrate may be neutralized by adding glacial acetic acid), the solvent was eliminated under reduced pressure, and the remaining protecting groups were removed by reaction with concd aq. ammonia/dioxane 1:1 at 55 °C. Alternatively, a single cleavage and deprotection step was carried out by reaction with the 1:1 ammonia/dioxane solution at 55 °C, or at room temperature if sufficiently labile nucleobase protecting groups were used.

Stability of the amino acid side chain protecting groups
Trifunctional amino acid derivatives (Boc-Lys(Tfa)-OH, Boc-Lys(Fmoc)-OH, Boc-Arg(Fmoc)$_2$-OH, Boc-Trp-OH, Boc-Trp(For)-OH, Boc-Ser(Ac)-OH, Boc-Thr(Ac)-OH and Boc-Tyr(Ac)-OH) were treated with the reagents used in nucleotide incorporation cycles by the phosphite triester approach. TLC or HPLC analysis showed that all of these derivatives remained undegraded after a period of 2 h, with the exception of Boc-Trp-OH which was not stable to the aqueous iodine oxidizing reagent. This amino acid was stable to a 1 M solution of *t*-BuOOH and did not react with dimethoxytrityl cations. The lability of the side chain protecting groups to the basic conditions used for the final deprotection of nucleopeptides (0.05 M TBAF in anhyd. THF, 30 min, rt; concd. aq. NH_3/dioxane 1:1, rt or 55 °C) was as follows:
-Lys(Fmoc)- : quantitatively deprotected both by TBAF (5 min) or NH_3 (6h, rt)
-Lys(Tfa)- : stable to TBAF (2 h). Elimination by NH_3 : 60 % at rt, 24 h; quantitative at 55 °C, 7 h
-Arg(Fmoc)$_2$- : quantitatively deprotected both by TBAF (10 min) or NH_3 (6h, rt)
-Trp(For)- : quantitatively deprotected both by TBAF (2 h) or NH_3 (6 h, rt) treatments
-Ser(Ac)- : stable to TBAF (5 h). Quantitative removal by NH_3 at rt, 17 h, or at 55 °C, 4 h
-Thr(Ac)- : stable to TBAF (5 h). Quantitative removal by NH_3 at rt, 17 h, or at 55 °C, 4 h
-Tyr(Ac)- : stable to TBAF (5 h). Quantitative removal by NH_3 at rt, 1 h.

Ac-Gly-Ala-Hse(p^3'dACTAGT)-Lys-Val-OH, 1
The oligonucleotide elongation was carried out on both Ac-Gly-Ala-Hse-Lys(Fmoc)-Val-O-NPE-Leu-resin and Ac-Gly-Ala-Hse-Lys(Tfa)-Val-O-NPE-Leu-resin. Final deprotection: TBAF + NH_3.
1 was purified by reversed-phase MPLC (gradient from 0 to 30 % of B, A=0.05 M TEAAc (pH=7), B=ACN/H_2O 1:1) and anion exchange through DEAE-Sephadex (gradient from 0.1 to 1.5 M triethylammonium bicarbonate). Overall synthesis, deprotection and purification yield: 10 %.
HPLC analysis: C18, gradient from 5 to 45 % of B in 30 min: t_R=16.8 min; SAX, gradient from 0 to 30 % of B in 30 min, t_R=21.7 min. Amino acid composition: Gly 1.49, Ala 1.08, Val 0.97, Lys 0.95; nucleoside composition: dC 0.94, dG 1.04, dT 2.01, dA 0.85. Electrospray-MS (negative mode): m/z 591.49 $[M-4H]^{4-}$; 788.80 $[M-3H]^{3-}$, 1183.65 $[M-2H]^{2-}$; calculated mass for $C_{81}H_{114}N_{28}O_{44}P_6$: 2369.80, mass found: 2369.58±0.30. MALDI-TOF MS (negative mode, THAP): 2367.4, $[M-H]^-$.

Boc-Hse(p^3'dATATTTGTTACTCTGT)-Lys-Lys-Lys-Lys-Lys-OH, 2
Synthesis on a polystyrene-co-1%-divinylbenzene matrix. Boc-Lys(Fmoc)-O-NPE-Leu-resin: Lys/Leu=0.86; Boc-Lys(Tfa)-O-NPE-Leu-resin: Lys/Leu=0.57. Coupling yield of the first nucleoside: on Boc-Hse-[Lys(Fmoc)]$_5$-resin, 80 %, on Boc-Hse-[Lys(Tfa)]$_5$-resin: 94 %. Final deprotection: TBAF + NH_3.

The purest batch of crude **2**, corresponding to oligonucleotide assembly on Boc-Hse(DMT)-[Lys(Tfa)]$_5$-NPE-Leu-resin, was purified by preparative PAGE and analytical reversed-phase HPLC (gradient from 5 to 30 % of B in 30 min). Overall synthesis, deprotection and purification yield: 7 %.

Synthesis on CPG. All couplings were carried out using excess of acylating reagent to obtain quantitative yields. Oligonucleotide elongation on Boc-[Lys(Tfa)]$_5$-O-NPE-Ala-resin was carried out following standard methodology, with 0.1 M and 0.5 M solutions of phosphoramidites and tetrazole, respectively. Cleavage of the nucleopeptide-resin bond and partial elimination of permanent protecting groups was accomplished by treatment with 0.05 M LiOH in MeOH/dioxane/H_2O 1.5:1.5:1 for 4 h at room temperature. After addition of acetic acid to the filtrate and evaporation to dryness, complete deprotection was achieved by overnight reaction with concd aq. ammonia at 55 °C.

Nucleoside composition: dC 2.21, dG 1.76, dT 9.02, dA 1.68. Electrospray-MS (negative mode): m/z 824.17 $[M-7H]^{7-}$, 961.43 $[M-6H]^{6-}$ (different peaks corresponding to $[M-xH+yNa]^{z-}$ were also detected), calculated mass for $C_{197}H_{278}N_{60}O_{112}P_{16}$: 5774.28, mass found: 5775.68 ±2.22. MALDI-TOF MS (negative mode, THAP): m/z 5771.8.

$^{5'}$TGTCTCATTGTTTATA$^{3'}$ and $^{5'}$TATAAACAATGAGACA$^{3'}$

Both oligonucleotides were assembled on CPG at the 1 μmol-scale following standard procedures (phosphite triester approach), and deprotected by reaction with concd aqueous ammonia at 55 °C. Both were purified by preparative PAGE.

$^{5'}$TGTCTCATTGTTTATA$^{3'}$: dC 2.17, dG 1.83, dT 9.37, dA 2.62. MALDI-TOF MS (negative mode, THAP): m/z 2426.32 $[M-2H]^{2-}$, 4852.55 $[M-H]^-$; calculated mass for $C_{158}H_{202}N_{49}O_{100}P_{15}$: 4852.21.

$^{5'}$TATAAACAATGAGACA$^{3'}$: dC 2.31, dG 1.91, dT 3.30, dA 8.46. MALDI-TOF MS (negative mode, THAP): m/z 2451.93 $[M-2H]^{2-}$, 4904.63 $[M-H]^-$, calculated mass for $C_{158}H_{196}N_{67}O_{88}P_{15}$: 4906.30.

Thermal denaturation of nucleopeptide-oligonucleotide and oligonucleotide duplexes

Equimolar amounts of Boc-Hse(p$^{3'}$dATATTTGTTACTCTGT)-Lys-Lys-Lys-Lys-Lys-OH and $^{5'}$TATAAACAATGAGACA$^{3'}$, or $^{5'}$TATAAACAATGAGACA$^{3'}$ and $^{5'}$TGTCTCATTGTTTATA$^{3'}$ were mixed, lyophilised, and dissolved to obtain a 2 μM duplex concentration. Buffers: [10 mM PIPES, 100 mM NaCl and 10 mM $MgCl_2$, pH=7] or [10 mM PIPES and 100 mM NaCl, pH=7]. Duplex-containing solutions were rapidly heated to 90 °C, and after 5 min at this temperature were slowly cooled at room temperature and kept in a refrigerator (4 °C) until they were used. The melting profile was obtained by measuring the absorbance of the mixture every 0.5 °C, from 5 to 90 °C (Cary 5E spectrophotometer (Varian); cuvettes cell was kept under N_2 to prevent H_2O condensation). Melting temperatures were obtained from the first derivative.

Boc-Arg(Fmoc)$_2$-OH

1 g of Boc-Arg-OH.H_2O.HCl (3.04 mmol) was coevaporated three times with dry acetonitrile. The resulting solid and 2.4 g of Fmoc-Cl were dried in a dissicator. Under argon atmosphere, 25 mL of dry DCM was added to the amino acid, followed by 1.6 mL (9.4 mmol) of dry DIEA and 1.5 mL (11.9 mmol) of TMS-Cl. The mixture was heated at 40 °C for 90 min, and the original suspension slowly became a solution. After cooling in an ice bath, 1.6 mL of DIEA and 2.4 g (9.28 mmol) of Fmoc-Cl (dissolved in the minimal amount of dry DCM) was added. The mixture was stirred for 20 min in the ice bath and 6.5 h at rt. The reaction mixture (pH=2) was washed with 50 mL of H_2O, and the aqueous phase reextracted with DCM (3 x). The organic extract was dried over Na_2SO_4 and filtered, and the solvent eliminated under reduced pressure to yield an orange oil.

Elimination of the less polar contaminants was partially achieved by dissolving the oil in 1.5-2 mL of DCM and pouring the solution onto 30 mL of hexanes. The yellowish precipitate obtained after filtration was further purified by silica gel column chromatography eluting with solvents of increasing polarity: hexanes/DCM 1:1, hexanes/DCM 25:75, DCM, and DCM/MeOH mixtures (99:1, 98:2, 97:3, 96:4, 95:5). Synthesis and purification yield: 60 %.

R_f (DCM/MeOH/AcOH 93:5:2)=0.54 ; mp 106-109 °C; IR (film): 3800, 1720, 1620, 1515, 1460, 1260, 1110, cm^{-1}; ^{1}H-NMR: δ 7.4-7.3 (m, 12 H), 5.0 (1H, d), 4.7 (m, 2H), 4.3 (m, 5H), 3.5 (m, 2H), 1.8 (m, 4H), 1.5 (s, 9H); ^{13}C-NMR: δ 161, 159, 157, 128-120, 78, 68, 47, 44, 29, 28, 25; FAB-MS (positive mode, magic bullet): m/z 497.1 [(Boc-Arg(Fmoc)-OH)+H]$^+$, 719.4 $[M+H]^+$, calculated mass for $C_{41}H_{42}N_4O_8$: 718.81.

Phac-Ala-Hse(p$^{3'}$dACTAGT)-Gly-Arg-Val-OH, 3

Phac-Ala-Hse-Gly-Arg(Fmoc)$_2$-Val-O-NPE-Leu-resin: Gly 0.96, Ala 0.86, Val 1.04, Orn 0.34, Arg 0.54. A sample of Phac-Ala-Hse-Gly-Arg-Val OH was obtained by treatment of an aliquot of peptide-resin with 0.05 M TBAF in anhydrous THF for 30 min. Pentapeptide Phac-Ala-Hse-Gly-Arg-Val-OH was also obtained after HF

treatment of Phac-Ala-Hse-Gly-Arg(Ts)-Val-O-Pam-resin. The main peak of the two peptide crudes coeluted (t_R: 12 min, peptide analysis conditions: gradient from 10 to 50 % of B, 20 min).
Oligonucleotide elongation was carried on two Phac-Ala-Hse-Gly-Arg(Fmoc)$_2$-Val-resin batches, either oxidizing with the standard iodine solution or with 1 M *t*-BuOOH (in DCM/toluene 2:1). The protected nucleopeptide-resin was first treated with 20 % piperidine/DMF (2 x 1 min), and the nucleopeptide deprotection was then carried out (TBAF + NH_3).
3 was purified by reversed-phase MPLC (gradient from 15 to 40 % of B, A=[0.05 M TEAAC/(ACN/H_2O 1:1) 85:15, B=[0.05 M TEAAC/(ACN/H_2O 1:1) 60:40). Overall synthesis, deprotection and purification yield: 2 %.
HPLC analysis: C18, gradient from 20 to 60 % of B in 30 min: t_R=8.1 min. Nucleoside composition: dC 1.07, dG 1.01, dT 1.92, dA 1.13. MALDI-TOF MS (THAP, negative mode): 2474.27, $[M-H]^-$, calculated mass for $C_{87}H_{118}N_{30}O_{44}P_6$: 2473.91.

H-Trp-Val-Hse(p$^{3'}$dACTAGT)-Gly-OH, 4

Tripeptide-resin Boc-Val-Hse-Gly-O-HMFS-Leu-resin was divided in two different batches. In the first batch the peptide was completed by coupling Fmoc-Trp-OH; in the second, the incorporation of Boc-Trp(For)-OH was followed by elimination of the Boc group and introduction of the Fmoc protecting group by reaction with Fmoc-Cl (3 equiv) and DIEA (1.1 equiv). HPLC-MS analysis of the peptide crude after ammonia treatment (overnight, 55 °C): gradient from 10 to 50 % of B in 30 min, t_R of the tetrapeptide=10.4 min, t_R of H-Trp-Val-Hse-NH_2=9.6 min.
After oligonucleotide elongation (oxidizing with 1 M *t*-BuOOH) and deprotection with ammonia, **4** was purified by reversed-phase MPLC (gradient from 0 to 30 % of B, A=0.05 M TEAAc (pH=7), B=ACN/0.05 M TEAAc 3:7). Overall synthesis, deprotection and purification yield: 22 %.
HPLC analysis: C18, gradient from 10 to 40 % of B in 30 min: t_R=11.1 min; SAX, gradient from 10 to 50 % of B in 30 min, t_R=13.1 min. Amino acid composition: Gly 1.50, Val 0.70, Hse 0.90, Trp 0.95; nucleoside composition: dC 0.96, dG 0.84, dT 2.20. Electrospray-MS (negative mode): m/z 577.5 $[M-4H]^{4-}$; 770.4 $[M-3H]^{3-}$, calculated mass for $C_{81}H_{105}N_{27}O_{42}P_6$: 2314.72, mass found: 2314.3±0.09.

H-Trp-Val-Hse(S-p$^{3'}$dACTAGT)-Gly-OH, 5

The oligonucleotide was assembled on Fmoc-Trp-Val-Hse-Gly-O-HMFS-Leu-resin replacing the oxidizing reagent by a 0.05 M solution of 3*H*-1,2-benzodithiole-3-one 1,1-dioxide (Beaucage sulfurizing reagent) in anhyd ACN. The sulfurization was carried out before the capping step. Deprotection and cleavage: NH_3.
The target nucleopeptide (mixture of diastereomers) was isolated after reversed-phase MPLC purification (gradient from 0 to 35 % of B, A=0.05 M TEAAc (pH=7), B=ACN/0.05 M TEAAc 35:65). A single spot was visualized after PAGE analysis under denaturing conditions. Overall synthesis, deprotection and purification yield: 19 %.
Amino acid composition: Gly 1.71, Val 0.80, Hse 0.90, Trp 0.65. Electrospray-MS (negative mode): m/z 601.7 $[M-4H]^{4-}$; 802.6 $[M-3H]^{3-}$, calculated mass for $C_{81}H_{105}N_{27}O_{36}P_6S_6$: 2411.09, mass found: 2410.71±0.02.

Boc-Thr(Ac)-OH.DCHA

Boc-Thr-ONbn. 3.4 g (15.7 mmol) of 4-nitrobenzyl bromide and 1.6 g (15.7 mmol) of triethylamine was added to a solution of 2.5 g (12.5 mmol) of Boc-Thr-OH in AcOEt (20 mL). The mixture was refluxed for 6 h, cooled and filtered. The filtrate was washed with H_2O, 1 M HCl and 5 % $NaHCO_3$ (3x20 mL), dried over Na_2SO_4, and evaporated to dryness. Boc-Thr-ONbzl was obtained pure (white solid) after silica gel column chromatography eluting with DCM and increasing amounts of MeOH (0 to 5 %). 70 % yield; mp 95-98 °C; R_f ($CHCl_3$/MeOH 95:5)=0.60; $[\alpha]_D$=-9.8 (*c* 0.62, MeOH); IR (KBr): 3450, 1760, 1670, 1520, 1350 cm^{-1}; ^{1}H-NMR: δ 8.2 (d, 2H, J=8.8 Hz), 7.5 (d, 2H, J=8.8 Hz), 5.4 (d, 1H), 5.3 (s, 2H), 4.4 (m, 2H), 1.5 (s, 9H), 1.3 (d, 3H); ^{13}C-NMR: δ 171, 156, 152, 143, 128, 124, 80, 68, 66, 59, 28, 20; FAB-MS (positive mode, Xe, magic bullet): m/z 355.1 $[M+H]^+$, 377.0 $[M+Na]^+$, calculated mass for $C_{16}H_{22}N_2O_7$: 354.36.
Boc-Thr(Ac)-ONBn. 4.3 mL (45 mmol) of Ac_2O and 0.55 g (4.5 mmol) of DMAP was added to a solution of 3.2 g of Boc-Thr-ONbn (9 mmol) in AcOEt. The mixture was stirred for 2 h at room temperature. 10 mL MeOH was added to the reaction mixture and it was stirred for 10 min more. After elimination of the solvent, the resulting oil was purified by silica gel column chromatography eluting with dichloromethane and increasing amounts of MeOH (0 to 5 %). 95 % yield (oil); R_f ($CHCl_3$/MeOH 95:5)=0.84; $[\alpha]_D$=-7.2 (*c* 0.23, MeOH); IR (film): 1750, 1710, 1610, 1530, 1350 cm^{-1}; ^{1}H-NMR: δ 8.2 (d, 2H, J=8.8 Hz), 7.5 (d, 2H, J=8.8 Hz), 5.4 (m, 1H), 5.3 (s, 2H), 5.2 (d, 1H), 4.5 (m, 1H), 2.1 (s, 3H), 1.5 (s, 9H), 1.3 (d, 3H); ^{13}C-NMR: δ 177, 170, 156,

148, 142, 129, 124, 81, 70, 66, 57, 28, 21, 17; FAB-MS (positive mode, Xe, NBA): m/z 397.0 $[M+H]^+$, calculated mass for $C_{18}H_{24}N_2O_8$: 396.40.

Boc-Thr(Ac)-OH. 0.18 mg of 5 % Pd/C was added to a solution of 0.95 g (2.4 mmol) of Boc-Thr(Ac)-ONbn in AcOEt/MeOH 1:1. The system was first purged with Ar, and H_2 was bubbled into the reaction mixture for 3 h. The catalyst was eliminated by filtration and washed with MeOH, and the filtrate was evaporated to dryness. The resulting oil was dissolved in AcOEt and washed with 1 M HCl (6 x) and H_2O (3 x). The organic solution was dried over Na_2SO_4 and filtered, and the solvent eliminated. The product (oil) was purified by silica gel column chromatography eluting with dichloromethane and increasing amounts of MeOH (0 to 5 %). 85 % yield (oil); R_f ($CHCl_3$/MeOH 95:5)=0.71; IR (film): 3400, 1750, 1710, 1520 cm^{-1}; ^{1}H-NMR: δ 5.5 (m, 1H), 5.2 (d, 1H), 4.5 (m, 1H), 2.1 (s, 3H), 1.5 (s, 9H), 1.3 (d, 3H); ^{13}C-NMR: δ 174, 170, 156, 81, 71, 57, 28, 21, 17.

Boc-Thr(Ac)-O⁻HDCHA⁺. A solid derivative was obtained by dissolving Boc-Thr(Ac)-OH in hexanes/ether 1:1 and adding an equimolar amount of dicyclohexylamine. After one night at 0 °C, the white solid was filtered and washed. 93 % yield; R_f ($CHCl_3$/MeOH/AcOH 85:10:5)=0.47; ^{13}C-NMR: δ 174, 170, 156, 78, 73, 59, 53, 29, 28, 25-24, 21, 17; FAB-MS (positive mode, Xe, magic bullet): m/z 443.0 $[M+H]^+$, 465.0 $[M+Na]^+$, calculated mass for $C_{23}H_{42}N_2O_6$: 442.60.

Phac-Hse(p$^{3'}$dACTAGT)-Ala-Thr-Ser-Val-Tyr-Gly-OH, 6

Boc-Ser(Ac)-OH and and Boc-Thr(Ac)-OH derivatives were liberated from their dicyclohexylammonium salts by treatment with 1 M H_2SO_4. Oligonucleotide elongation was carried out on Phac-Hse-Ala-Thr(Ac)-Ser(Ac)-Val-Tyr(Ac)-Gly-O-HMFS-Leu-resin. Deprotection: NH_3, 6 h at room temperature and overnight at 55 °C.

Purification was accomplished by reversed-phase MPLC (gradient from 15 to 25 % of B, A=0.05 M TEAAc (pH=7), B=ACN/H_2O 1:1). Overall synthesis, deprotection and purification yield: 4 %.

HPLC analysis: C18, gradient from 16 to 30 % of B in 30 min: t_R=13.8 min. Amino acid composition: Thr 0.71, Ser 0.43, Gly 2.26, Ala 1.11, Val 0.89, Tyr 0.90; nucleoside composition: dC 1.2, dG 1.1, dT 1.9, dA 1.0. Electrospray-MS (negative mode): m/z 663.3 $[M-4H]^{4-}$; 672.0 $[M-5H+Na]^{4-}$, 677.5 $[M-6H+2Na]^{4-}$, 888.5 $[M-3H]^{3-}$, 896.4 $[M-4H+Na]^{3-}$, 903.5 $[M-5H+2Na]^{3-}$, 1366.5 $[M-5H+3Na]^{2-}$; calculated mass for $C_{97}H_{127}N_{29}O_{49}P_6$: 2667.08, mass found: 2669.6±0.7.

Phac-Ser(p$^{3'}$dTATAT)-Ala-Ser-Val-OH (7), Phac-Tyr(p$^{3'}$dTATAT)-Ala-Tyr-Val-OH (8) and Phac-Tyr(p$^{3'}$dTATAT)-Ala-Ser-Val-OH (9)

For the synthesis of **7**, Boc-Ser-OH was coupled onto H-Val-O-HMFS-Leu-resin, and, before the Boc group was removed, the side chain hydroxyl group of the serine residue was acetylated by reaction with Ac_2O (0.15 mL), pyridine (0.20 mL) and DMAP (2 mg). The NH_2 group was deprotected and Boc-Ala-OH incorporated. Boc-Ala-Ser(Ac)-Val-resin was divided in two batches, on which peptide assembly was continued to yield either Phac-Tyr-Ala-Ser(Ac)-Val-resin or Phac-Ser-Ala-Ser(Ac)-Val-resin. Peptide-resin Phac-Tyr-Ala-Tyr(Ac)-Val-O-HMFS-Leu-MBHA was obtained from incorporation of Boc-Tyr-OH onto H-Val-resin, acetylation of the phenolic hydroxyl group as described above, and subsequent coupling of Boc-Ala-OH, Boc-Tyr-OH and phenylacetic acid. The peptide-resins composition was checked before the oligonucleotide elongation to ensure that acetylation of the side chain hydroxyl group of the serine or tyrosine residues had been complete and that no peptide chain had grown at these hydroxyl groups: Phac-Tyr-Ala-Ser(Ac)-Val-resin, Ser 0.51, Ala 0.97, Val 1.03, Tyr 0.70; Phac-Ser-Ala-Ser(Ac)-Val-resin, Ser 0.49, Ala 0.81, Val, 1.19; Phac-Tyr-Ala-Tyr(Ac)-Val-resin, Ala 0.93, Val 1.07, Tyr 0.74.

7 was obtained by treating the corresponding nucleopeptide-resin with ammonia (18 h, rt) and reversed-phase MPLC purification (gradient from 17 to 30 % of B, A=0.05 M TEAAc (pH=7), B=ACN/H_2O 1:1).

After deprotection with ammonia (4 h at 55 °C), **8** and **9** were and purified by reversed-phase MPLC (gradients from 15 to 35 % of B and from 16 to 30 % of B, respectively; A=0.05 M TEAAc (pH=7), B=ACN/H_2O 1:1).

7. Overall synthesis, deprotection and purification yield: 2 %. HPLC analysis: C18, gradient from 17 to 30 % of B in 30 min: t_R=15.9 min; amino acid composition: Ser 0.55, Ala 1.02, Val 0.98; nucleoside composition: dT 3.1, dA 1.9; electrospray-MS (negative mode): m/z 504.4 $[M-4H]^{4-}$; 672.1 $[M-3H]^{3-}$, 1008.4 $[M-2H]^{2-}$; calculated mass for $C_{72}H_{95}N_{20}O_{39}P_5$: 2019.52, mass found: 2019.93±0.7.

8. Overall synthesis, deprotection and purification yield: 13 %. HPLC analysis: C18, gradient from 15 to 35 % of B in 30 min: t_R=14.5 min; amino acid composition: Ala 1.02, Val 0.98, Tyr 1.41; nucleoside composition: dT 3.0, dA 2.0; electrospray-MS (negative mode): m/z 541.6 $[M-4H]^{4-}$; 722.6 $[M-3H]^{3-}$, 729.7 $[M-4H+Na]^{3-}$, 1084.3 $[M-3H+Na]^{2-}$, 1095.6 $[M-2H]^{2-}$; calculated mass for $C_{84}H_{103}N_{20}O_{39}P_5$: 2171.72, mass found: 2170.7±0.5.

9. Overall synthesis, deprotection and purification yield: 11 %. HPLC analysis: C18, gradient from 16 to 30 % of B in 30 min: t_R=19.1 min; amino acid composition: Ser 0.52, Ala 1.02, Val 0.98, Tyr 0.85; nucleoside composition: dT 3.1, dA 1.9; electrospray-MS (negative mode): m/z 418.3 $[M-5H]^{5-}$, 522.8 $[M-4H]^{4-}$; 528.4 $[M-5H+Na]^{4-}$, 697.5 $[M-3H]^{3-}$, 704.9 $[M-4H+Na]^{3-}$; 712.1 $[M-5H+2Na]^{3-}$; calculated mass for $C_{78}H_{99}N_{20}O_{39}P_5$: 2095.62, mass found: 2095.7±0.4.

Acknowledgements. This work was supported by funds from the Spanish MEC (grant PB97-0941-C02-01) and the Generalitat de Catalunya (SGR98-0001 and Centre de Referència de Biotecnologia). The authors thank Dr. Irene Fernández and Dr. Marta Vilaseca for the electrospray mass spectrometric analysis.

References and notes

1 Salas, M. *Annu. Rev. Biochem.* **1991**, *60*, 39-71.
2 Wang, J. C. *Annu. Rev. Biochem.* **1996**, *65*, 635-692.
3 Dreef-Tromp, C. M.; van der Maarel, J. C. M.; van den Elst, H.; van der Marel, G. A.; van Boom, J. H. *Nucleic Acids Res.* **1992**, *20*, 4015-4020.
4 Ueno, Y.; Saito, R.; Hata, T. *Nucleic Acids Res.* **1993**, *21*, 4451-4457.
5 Robles, J.; Pedroso, E.; Grandas, A. *Nucleic Acids Res.* **1995**, *23*, 4151-4161.
6 Flohr, S.; Jungmann, V.; Waldmann, H. *Chem. Eur. J.* **1999**, *5*, 669-681.
7 Bongartz, J-P.; Aubertin, A-M.; Milhaud, P. G.; Lebleu, B. *Nucleic Acids Res.* **1994**, *22*, 4681-4688.
8 G. de la Torre, B.; Aviñó, A.; Tarrasón, G.; Piulats, J.; Albericio, F.; Eritja, R. *Tetrahedron Lett.* **1994**, *35*, 2733-2736.
9 Basu, S.; Wickstrom, E. *Tetrahedron Lett.* **1995**, *36*, 4943-4946.
10 Souckchareun, S.; Tregear, G. W.; Haralambidis, J. *Bioconjugate Chem.* **1995**, *6*, 43-53.
11 Reed, M. W.; Fraga, D.; Schwartz, D. E.; Scholler, J.; Hinrichsen, R. D. *Bioconjugate Chem.* **1995**, *6*, 101-108.
12 Arar, K.; Aubertin, A-M.; Roche, A-C.; Monsigny, M.; Mayer, R. *Bioconjugate Chem.* **1995**, *6*, 573-577.
13 Wei, Z.; Tung, C-H.; Zhu, T.; Dickerhof, W. A.; Breslauer, K. J.; Georgopoulos, D. E.; Leibowitz, M. J.; Stein, S. *Nucleic Acids Res.* **1996**, *24*, 655-661.
14 Soukchareun, S.; Haralambidis, J.; Tregear, G. W. *Bioconjugate Chem.* **1998**, *9*, 466-475.
15 Peyrottes, S.; Mestre, B.; Burlina, F.; Gait, M. J. *Tetrahedron* **1998**, *54*, 12513-12522.
16 Kuyl-Yeheskiely, E.; Tromp, C. M.; Lefeber, A. W. M.; van der Marel, G. A.; van Boom, J. H. *Tetrahedron* **1988**, *44*, 6515-6523.
17 Dreef-Tromp, C. M.; van Dam, E. M. A.; van den Elst, H.; van der Marel, G. A.; van Boom, J. H. *Nucleic Acids Res.* **1990**, *18*, 6491-6495.
18 Dreef-Tromp, C. M.; van den Elst, H.; van den Bogaart, J. E.; van der Marel, G. A.; van Boom, J. H. *Nucleic Acids Res.* **1992**, *20*, 2435-2439.
19 Robles, J.; Pedroso, E.; Grandas, A. *Tetrahedron Lett.* **1991**, *32*, 4389-4392.
20 Robles, J.; Pedroso, E.; Grandas, A. *J. Org. Chem.* **1994**, *59*, 2482-2486.
21 Robles, J.; Maseda, M.; Beltrán, M.; Concernau, M.; Pedroso, E.; Grandas, A. *Bioconjugate Chem.* **1997**, *8*, 785-788.
22 Shabarova, Z. A. In *Progress in Nucleic Acid Research and Molecular Biology*; Vol. 10; Davidson, J. N. and Colum, W. E. Eds.; Academic Press: London, 1970; pp. 145-182.
23 Juodka, B. A. *Nucleosides and Nucleotides* **1984**, *3*, 445-483.
24 Beltrán, M.; Maseda, M.; Pérez, Y.; Robles, J.; Pedroso, E.; Grandas, A. *Nucleosides and Nucleotides* **1997**, *16*, 1487-1488 (preliminary communication).
25 Beltrán, M.; Pedroso, E.; Grandas, A. *Tetrahedron Lett.* **1998**, *39*, 4115-4118.
26 Beltrán, M.; Maseda, M.; Robles, J.; Pedroso, E.; Grandas, A. *Lett. Pept. Sci.* **1997**, *4*, 147-155.
27 Eritja, R.; Robles, J.; Fernández-Forner, D.; Albericio, F.; Giralt, E.; Pedroso, E. *Tetrahedron Lett.* **1991**, *32*, 1511-1514.
28 Rabanal, F.; Giralt, E.; Albericio, F. *Tetrahedron Lett.* **1992**, *33*, 1775-1778.
29 Abbreviations: Ac=acetyl, ACN=acetonitrile, Boc=*t*-butoxycarbonyl, Bz=benzoyl, CNE=2-cyanoethyl, CPG=controlled pore glass, DBU=1,8-diazabicyclo[5.4.0]undec-7-ene, DCC=N,N'-dicyclohexylcarbodiimide, DCHA=dicyclohexylamine, DCM=dichloromethane, DIEA=N-ethyl-N,N-diisopropylamine, DIPC=N,N'-diisopropylcarbodiimide, DMT=4,4'-dimethoxytrityl, DMAP=4-N,N-dimethylaminopyridine Dnp=2,4-dinitrophenyl, Fm=9-fluorenylmethyl, Fmoc=9-fluorenylmethoxycarbonyl, For=formyl,

HMFS=2-(9-oxymethylfluorenyl)succinyl, HOBt=1-hydroxybenzotriazole, Hse=homoserine, iBu=isobutyryl, IRAA=internal reference amino acid, MBHA=*p*-methylbenzhydrylamine, MPLC=medium pressure liquid chromatography, N=nucleoside, NBA=nitrobenzyl alcohol, Nbn=4-nitrobenzyl, NMI=N-methylimidazole, NPE=4-(2-oxyethyl)-3-nitrobenzoyl, PAGE=polyacrylamide gel electrophoresis, Pam=phenylacetamidomethyl, Phac=phenylacetyl, R =resin (solid matrix), SAX=strong anion exchange, TBAF=tetrabutylammonium fluoride, S-pdNNN...: phosphorothioated oligonucleotide, TEA=triethylamine, Tfa=trifluoroacetyl, TFA=trifluoroacetic acid, THAP=trihydroxyacetophenone, Tm=melting temperature, TMS=trimethylsilyl, Tos=tosyl.

30 Truffert, J. C.; Lorthioir, O.; Asseline, U.; Thuong, N. T.; Brack, A. *Tetrahedron Lett.* **1994**, *35*, 2353-2356.

31 Bashkin, J. K.; McBeath, R. J.; Modak, A. S.; Sample, K. R.; Wise, W. B. *J. Org. Chem.* **1991**, *56*, 3168-3176.

32 Balzarini, J.; Camarasa, M. J.; Karlsson, A. *Drugs of the Future* **1993**, *18*, 1043-1055.

33 Degols, G.; Leonetti, J-P.; Gagnor, C.; Lemaître, M.; Lebleu, B. *Nucleic Acids Res.* **1989**, *17*, 9341-9350.

34 Zhu, T.; Wei, Z.; Tung, C-H.; Dickerhof, W. A.; Breslauer, K. J.; Georgopoulos, D. E.; Leibowitz, M. J.; Stein, S. *Antisense Research and Development* **1993**, *3*, 265-275.

35 Wadhwa, M. S.; Collard, W. T.; Adami, R. C.; McKenzie, D. L.; Rice, K. G. *Bioconjugate Chem.* **1997**, *8*, 81-88.

36 Vivès, E.; Lebleu, B. *Tetrahedron Lett.* **1997**, *38*, 1183-1186.

37 Eritja, R.; Pons, A.; Escarceller, M.; Giralt, E.; Albericio, F. *Tetrahedron* **1991**, *47*, 4113-4120.

38 Zhu, T.; Pooyan, S.; Wei, Z.; Leibowitz, M. J.; Stein, S. *Antisense and Nucleic Acid Drug Development* **1996**, *6*, 69-74.

39 Wei, Z.; Tung, C-H.; Zhu, T.; Dickerhof, W. A.; Breslauer, K. J.; Georgopoulos, D. E.; Leibowitz, M. J.; Stein, S. *Nucleic Acids Res.* **1996**, *24*, 655-661.

40 Harrison, J. G.; Balasubramanian, S. *Nucleic Acids Res.* **1998**, *26*, 3136-3145.

41 Barany, G.; Merrifield, R. B. In *The Peptides. Analysis, Synthesis, Biology*, Vol. 2 (Special Methods in Peptide Synthesis, Part A); Gross, E. and Meienhofer, J. Eds.; Academic Press: New York, 1980; pp. 169-175.

42 Geiger, R.; König, W. In *The Peptides. Analysis, Synthesis, Biology*, Vol. 3 (Protection of Functional Groups in Peptide Synthesis); Gross, E. and Meienhofer, J. Eds.; Academic Press: New York, 1981; pp. 60-70.

43 Jetten, M.; Peters, A. M.; van Nispen, J.; Ottenheim, H. *Tetrahedron Lett.* **1991**, *32*, 6025-6028.

44 Filippov, D.; van der Marel, G. A.; Kuyl-Yeheskiely, E.; and van Boom, J. H. *Synlett* **1994**, 922-924.

45 Barany, G.; Merrifield, R. B. In *The Peptides. Analysis, Synthesis, Biology*, Vol. 2 (Special Methods in Peptide Synthesis, Part A); Gross, E. and Meienhofer, J. Eds.; Academic Press: New York, 1980; pp. 217-223.

46 Iyer, R. P.; Phillips, L. R.; Egan, W.; Regan, J. B.; Beaucage, S. L. *J. Org. Chem.* **1990**, *55*, 4693-4699.

47 Robles, J.; Beltrán, M.; Pedroso, E.; Grandas, A. unpublished results.

48 Vu, H.; McCollum, C.; Jacobson, K.; Theisen, P.; Vinayak, R.; Speiss, E.; Andrus, A. *Tetrahedron Lett.* **1990**, *31*, 7269-7272.

49 Detachment of tyrosine-nucleopeptides from the resin by treatment with the non-nucleophile base DBU, followed by ammonia nucleobase deprotection, should afford a single product (C-terminal acid) without cleavage of the linking phosphodiester bond.

50 Efcavitch, J. W. In *Gel Electrophoresis of Nucleic Acids. A Practical Approach* (The Electrophoresis of Synthetic Oligonucleotides), 2nd ed.; Dickwood, D. and Hames, B. D. Eds.; IRL Press: Oxford, 1990; pp. 125-149.

51 Montserrat, F. X.; Grandas, A.; Eritja, R.; Pedroso, E. *Tetrahedron* **1994**, *50*, 2617-2622.