Pergamon

Tetrahedron Letters 39 (1998) 4115–4118

TETRAHEDRON LETTERS

A Comparison of Histidine Protecting Groups in the Synthesis of Peptide-Oligonucleotide Conjugates

Maite Beltrán, 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 21 January 1998; revised 26 March 1998; accepted 27 March 1998

Abstract: Protection of the imidazole ring of the histidine residue is not required for the elongation of an oligonucleotide chain at the side chain hydroxyl group of an amino acid residue by the phosphite triester approach. For the assembly of the peptide chain, the 2,4-dinitrophenyl group is the best alternative when histidine is placed at the C-terminal position.

Peptide-oligonucleotide conjugates are usually synthesized for use in structural studies or as potential therapeutic agents. The ability of imidazole rings to complex metal ions and cleave nucleic acid chains makes histidine-containing conjugates particularly interesting, since they may be designed to behave as artificial nucleases[1-5]. We wish to report on our results on the synthesis of phosphodiester-linked peptide-oligonucleotide hybrids (nucleopeptides) with a histidine residue at different positions of the peptide chain, in particular nucleopeptides Ac-Gly-Ala-Hse(p$^{3'}$dACTAGT)-His-Val-OH and Phac-Hse(p$^{5'}$dG)-His-OH[6].

Peptide-oligonucleotide hybrids can be obtained using different approaches. The structural integrity of the molecule can be ascertained from the amino acid and nucleoside composition and mass spectrometry, but these analytical methods may fail to discriminate between hybrid molecules differing only in the type of peptide-oligonucleotide covalent linkage. Consequently, linking the two components while all the other functional groups are protected seems the safest alternative to prepare conjugates of unequivocally defined structure. In the stepwise solid-phase methodology developed in our laboratory for the synthesis of nucleopeptides[7,8], the oligonucleotide is elongated at the free hydroxyl group of a protected peptide-resin. Boc and DMT groups are used to temporarily block the α-amino and 5'-hydroxyl groups, respectively, and base-labile groups to protect all the other functions.

The imidazole ring must be protected during the coupling of the histidine residue, since carboxyl activation of imidazole-unprotected histidine derivatives causes extensive racemization at the α-carbon[9,10]. It has also been reported that the quality of synthetic peptides is higher if the histidine residue remains protected throughout their assembly[11]. Among the commercially available N$^{\alpha}$-Boc histidine derivatives, only those with the imidazole ring protected with 2,4-dinitrophenyl or tosyl groups are compatible with our synthesis scheme. Both protecting groups were shown to be stable to all the reagents necessary for the assembly of the hybrid molecule, with the known exception of HOBt which eliminates the tosyl group[12].

Homoserine was chosen as the linking amino acid in the two target nucleopeptides in order to circumvent problems associated with the lability of phosphodiester linkages between nucleosides and the natural proteinogenic amino acids[13]. Homoserine derivatives are obtained as salts[13] that can only be coupled in the presence of a proton donor such as HOBt[13]. Since in the two target nucleopeptides the histidine residue is incorporated onto the resin before homoserine, a DCC/HOBt-mediated coupling of the homoserine moiety would

0040-4039/98/$19.00
PII: S0040-4039(98)00669-8

only leave unaltered a 2,4-dinitrophenyl-protected histidine. We have found that tetrazole is a suitable alternative to HOBt, since the stability of tosyl-protected histidine derivatives to tetrazole allows this reagent to be used both in the homoserine coupling and in the oligonucleotide elongation without significant removal of the tosyl group (25% elimination by a 0.7 M tetrazole solution in anhydrous THF, 24 h). Efficient coupling of the homoserine residue was consequently carried out under standard conditions substituting tetrazole for HOBt.

The lability of both tosyl and dinitrophenyl protecting groups to the reaction conditions commonly used for the final deprotection of nucleopeptides was also checked. The tosyl group was shown (TLC) to be quantitatively eliminated by a 0.05 M TBAF solution in anhydrous THF in 30 min, and after 17 h at 55 °C with a 1:1 conc. aq. ammonia/dioxane treatment. The lability of the 2,4-dinitrophenyl group was tested by treating aliquots of Ac-Gly-Ala-Hse-His(Dnp)-Val-OPAM-resin (PAM=phenylacetamidomethyl) with either 0.05 M TBAF in anhydrous THF, conc. aq. NH_3/dioxane 1:1 at room temperature or at 55 °C, or 2 M Ph-SH in dimethylformamide[14]. After a certain reaction time the peptide-resin bond was cleaved with HF (1 h, 0 °C). HPLC analysis[15] of the crudes showed that the Dnp group was quantitatively removed by TBAF in 30 min and, with the 1:1 conc. aq. ammonia/dioxane solution, either after a 17 h treatment at room temperature or after 6 h at 55 °C. It is also worth noticing that the cleanest peptide crude was obtained when the ammonia solution was used to eliminate the Dnp group, and the most complex crude when the standard thiophenol deprotection conditions[11,14] were used.

Both nucleopeptides Phac-Hse(p$^{5'}$dG)-His-OH and Ac-Gly-Ala-Hse(p$^{3'}$dACTAGT)-His-Val-OH were assembled on a hydroxyl-functionalized solid matrix[16] following essentially the same steps. The synthesis of the largest one (Ac-Gly-Ala-Hse(p$^{3'}$dACTAGT)-His-Val-OH) is summarized in Figure 1.

Attachment of the histidine residue to the solid support is the key step in the synthesis of Phac-Hse(p$^{5'}$dG)-His-OH. This reaction can be carried out, with a minimum risk of racemization, by carbodiimide activation in the presence of HOBt[17]. Since the tosyl group would be removed under these conditions, the 2,4-dinitrophenyl group was chosen for the protection of the C-terminal histidine residue. After histidine incorporation, unreacted hydroxyl groups were acetylated, the Boc group was removed and Phac-Hse(DMT)-O$^-$ $HNEt_3^+$ was coupled in the presence of diisopropylcarbodiimide and HOBt. The 2'-deoxyguanosine derivative (iPr_2N(CNEO)P-5'-O-dG^{Dmf}-3'-O-DMT)[13] was incorporated after the acid treatment that removed the side chain homoserine protecting group. Crude nucleopeptide Phac-Hse(p$^{5'}$dG)-His-OH was obtained after elimination of the 3'-DMT group, imidazole and phosphate deprotection with a 30 min TBAF treatment and nucleobase deprotection with a 1:1 conc. aq. ammonia/dioxane mixture at room temperature (6 h). The product was purified by reversed-phase MPLC[18,19] and its identity was confirmed by mass spectrometry[20].

In order to elucidate whether imidazole protection is necessary or not during oligonucleotide elongation, in the synthesis of nucleopeptide Ac-Gly-Ala-Hse(p$^{3'}$dACTAGT)-His-Val-OH (Figure 1) we prepared three partially protected peptides, Ac-Gly-Ala-Hse-His(X)-Val-NPE-Leu-resin, in which X=Dnp, Tos, or H, respectively. Instead of HOBt, tetrazole was used as the proton donor in the carbodiimide-mediated coupling of Boc-Hse(DMT)-O$^-$ $HNEt_3^+$ for the synthesis of Ac-Gly-Ala-Hse-His(Tos)-Val-NPE-Leu-resin, as already stated. An aliquot of the latter was then treated with HOBt to obtain Ac-Gly-Ala-Hse-His-Val-resin. The oligonucleotide chain was assembled on the three peptide-resins using 5'-DMT-3'-phosphoramidite nucleoside derivatives (T, dA^{Bz}, dC^{Bz} and dG^{iBu}), and the 5'-terminal DMT group was removed.

The three nucleopeptide-resins were treated with TBAF (30 min) and conc.aq. ammonia/dioxane 1:1 at 55 °C (overnight) to give crude Ac-Gly-Ala-Hse(p$^{3'}$dACTAGT)-His-Val-OH[21]. The product was purified by reversed-phase MPLC[18,19], and characterized from amino acid and nucleoside composition and mass

spectrometric analysis[22]. The comparison of the HPLC traces of crude nucleopeptides obtained from the three peptide-resins shows that the product obtained from Ac-Gly-Ala-Hse-His-Val-resin is the purest (Figure 1), so we conclude that protection of the imidazole ring is not required for the oligonucleotide elongation[23].

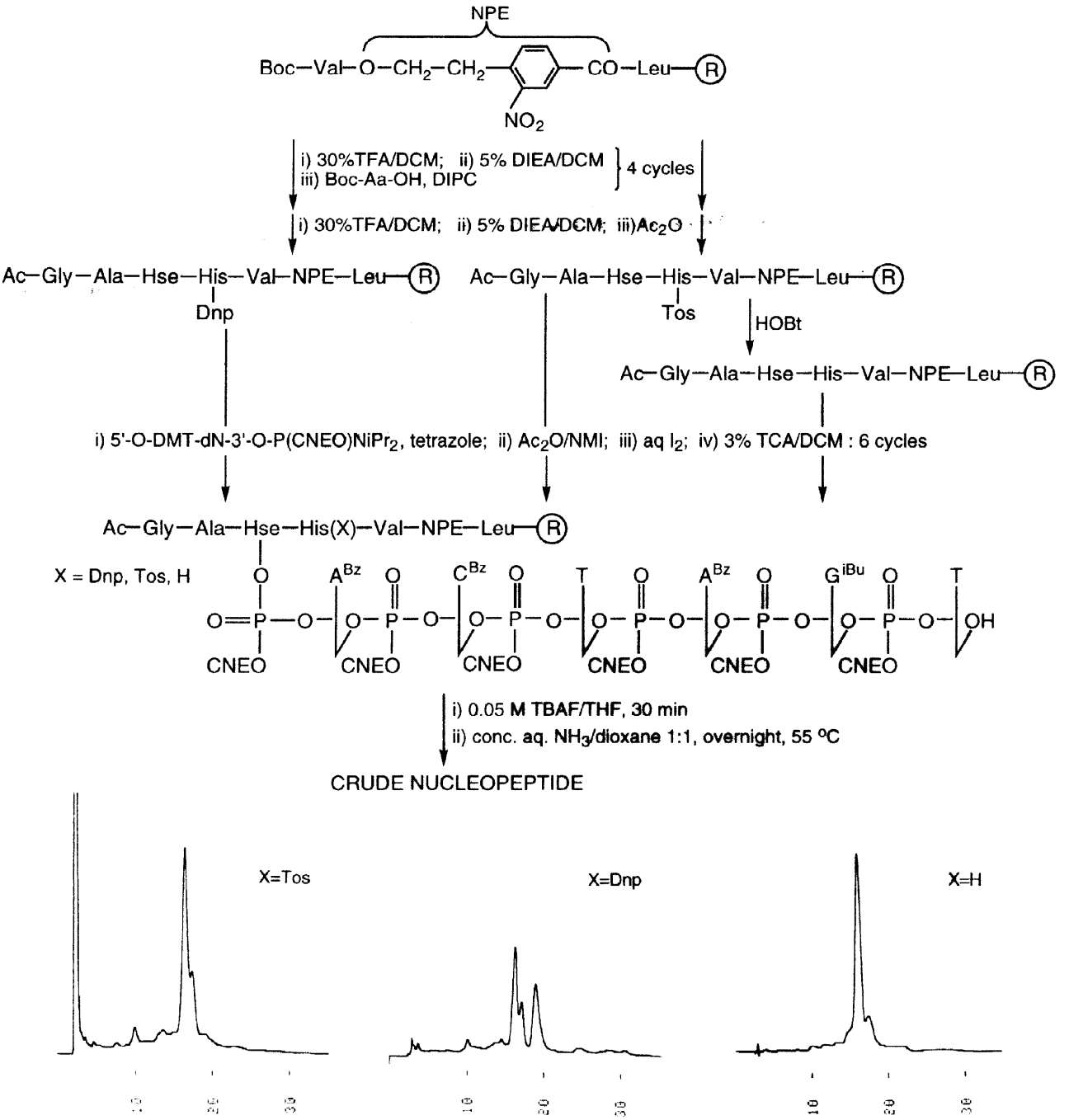

Figure 1. Synthesis of Ac-Gly-Ala-Hse(p$^{3'}$dACTAGT)-His-Val-OH and C-18 HPLC traces of crude nucleopeptides obtained from the three peptide-resins.

In summary, to minimize the risk of racemization during histidine coupling, the dinitrophenyl protecting group is the best option for the stepwise solid-phase preparation of peptide-oligonucleotide conjugates with a C-terminal histidine. When the histidine residue is placed at other positions of the peptide sequence, the imidazole

ring can be kept protected throughout the nucleopeptide assembly, using either the tosyl or the more stable 2,4-dinitrophenyl group, but the best alternative seems to be to assemble the peptide using a tosyl-protected histidine derivative and to remove the imidazole protecting group prior to oligonucleotide elongation.

Acknowledgements. This work was supported by funds from the DGICYT (PB94-0844) and the Generalitat de Catalunya (SRG96-1 and Centre de Referència de Biotecnologia).

REFERENCES AND NOTES.

1. Truffert, J-C.; Asseline, U.; Brack, A.; Thuong, N. T. *Tetrahedron* **1996**, *52*, 3005-3016.
2. Hovinen, J.; Guzaev, A.; Azhayeva, E.; Azhayev, A.; Lönnberg, H. *J. Org. Chem.* **1995**, *60*, 2205-2209.
3. Polushkin, N. N.; Chen, B-C.; Anderson, L. W.; Cohen, J. S. *J. Org. Chem.* **1993**, *58*, 4606-4613.
4. Bashkin, J. K.; McBeath, R. J.; Modak, A. S.; Sample, K. R.; Wise, W. B. *J. Org. Chem.* **1991**, *56*, 3168-3176.
5. Bashkin, J. K.; Gard, J. K.; Modak, A. S. *J. Org. Chem.* **1990**, *55*, 5125-5132.
6. Abbreviations used: Aa= amino acid; Boc=*t*-butoxycarbonyl; Bz=benzoyl; CNE=2-cyanoethyl; DCM=dichloromethane; DIEA=diisopropylethylamine; DIPC=N,N'-diisopropylcarbodiimide; Dmf=dimethylaminomethylene; DMT=4,4'-dimethoxytrityl; dN=2'-deoxynucleoside; Dnp=2,4-dinitrophenyl; HOBt=1-hydroxybenzotriazole; Hse=homoserine; iBu=isobutyryl; NMI=N-methylimidazole; NPE=3-nitro-4-(2-ethoxy)benzoyl; Phac=phenylacetyl; TBAF=tetrabutylammonium fluoride; TCA=trichloroacetic acid; TFA=trifluoroacetic acid; THF=tetrahydrofuran; Tos=tosyl.
7. Robles, J.; Pedroso, E.; Grandas, A. *J. Org. Chem.* **1994**, *59*, 2482-2486.
8. Robles, J., Pedroso, E., and Grandas, A. *Nucleic Acids Res.* **1995**, *23*, 4151-4161.
9. Barany, G.; Merrifield, R. B. Solid-Phase Peptide Synthesis. In *The Peptides. Analysis, Synthesis, Biology (volume 2, Special Methods in Peptide Synthesis, part A)* Gross, E.; Meienhofer, J., Eds. Academic Press; New York, 1980; pp. 179-189.
10. Geiger, R.; König, W. Amine Protecting Groups. In *The Peptides. Analysis, Synthesis, Biology (volume 3, Protection of Functional Groups in Peptide Synthesis)* Gross, E.; Meienhofer, J., Eds. Academic Press; New York, 1981; pp. 70-80.
11. Carreño, C.; Roig, X.; Cairó, J.; Camarero, J.; Mateu, M. G.; Domingo, E.; Giralt, E.; Andreu, A. *Int. J. Peptide Protein Res.* **1992**, *39*, 41-47.
12. Fuji, T.; Sakakibara, S. *Bull. Chem. Soc. Jpn.* **1974**, *47*, 3146-3151.
13. Beltrán, M.; Maseda, M.; Robles, J.; Pedroso, E.; Grandas, A. *Lett. Pept. Sci.* **1997**, *4*, 147-155.
14. Stewart, J. M. and Young, J. D. In *Solid Phase Peptide Synthesis* (2nd ed.) Pierce Chemical Co.; Rockford (Illinois), 1984; p. 83.
15. HF Treatment of Ac-Gly-Ala-Hse-His(Dnp)-Val-OPAM-resin afforded a sample of Ac-Gly-Ala-Hse-His(Dnp)-Val-OH. The sample of Ac-Gly-Ala-Hse-His-Val-OH was obtained after HF acidolysis of Ac-Gly-Ala-Hse-His(Tos)-Val-OPAM-resin.
16. Fmoc-leucine was first anchored to an amino-functionalized solid support (*p*-methylbenzhydrylamine resin) to yield an amino acid-resin with a substitution degree suitable for the oligonucleotide elongation, and unreacted amino groups were acetylated. A piperidine treatment eliminated the Fmoc group, and coupling of 3-nitro-4-(2-hydroxyethyl)benzoic acid afforded the resin onto which the nucleopeptides were assembled.
17. Grandas, A.; Jorba, X.; Giralt, E.; Pedroso, E. *Int. J. Peptide Protein Res.* **1989**, *33*, 386-390.
18. C-18 MPLC purification conditions: solvent A, 0.05 M triethylammonuim acetate; solvent B, acetonitrile/H_2O 1:1. Phac-Hse(p^5'dG)-His-OH: gradient from 5 to 20% of B; Ac-Gly-Ala-Hse(p^3'dACTAGT)-His-Val-OH: gradient from 5 to 25% of B.
19. A coloured organic impurity with the typical Dnp UV absorption may be found in the nucleopeptide crude, but uncoloured products are isolated after MPLC purification.
20. Electrospray-MS (negative mode): m/z 702.3, $[M-H]^-$, calculated mass for $C_{28}H_{34}N_9O_{11}P$: 703.61.
21. In the case of nucleopeptide-resin with X=Tos, previous removal of the tosyl group by reaction with HOBt afforded a slightly cleaner crude.
22. Amino acid composition after acid hydrolysis: Hse 0.14, Gly 1.49, Ala 1.13, Val 0.91, His 0.95; nucleoside composition after digestion with snake venom phosphodiesterase and alkaline phosphatase: dC 0.94, dG 1.06, T 1.99, dA 0.92. Electrospray-MS (negative mode): m/z 493.5 $[M-4H]^{4-}$, 792.2 $[M-3H]^{3-}$, 1188.1 $[M-2H]^{2-}$; experimental mass: 2378.63±0.70; calculated mass for $C_{81}H_{109}N_{29}O_{44}P_6$: 2378.76.
23. In addition, we carried out a complete nucleoside incorporation cycle on either Boc-Val-resin or Boc-His-Val-resin, and the same amount of DMT^+, virtually nil, was found in both cases (approx. 1%, probably due to some nucleoside adsorption on the polystyrene solid support).