Pergamon

Tetrahedron Letters 39 (1998) 3597–3600

TETRAHEDRON LETTERS

Synthesis of a Nucleopeptide Fragment from Poliovirus Genome

Dmitri Filippov, Esther Kuyl-Yeheskiely, Gijs A. van der Marel, Godefridus I. Tesser[§], Jacques H. van Boom*

Leiden Institute of Chemistry, Gorlaeus Laboratories, P.O. Box 9502, 2300 RA Leiden, The Netherlands

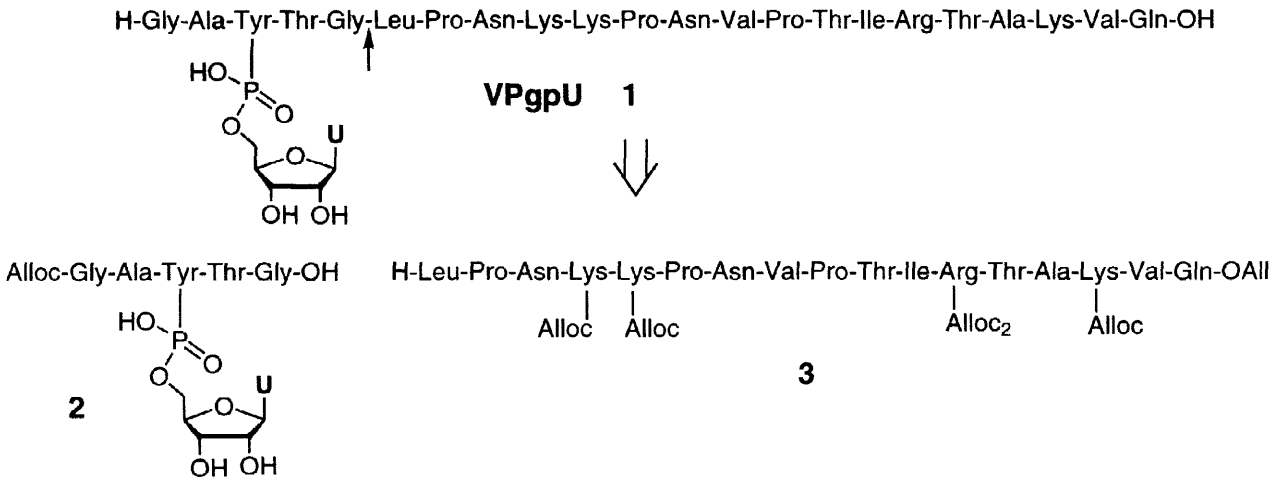
[§]Laboratory of Organic Chemistry, Catholic University, Toernooiveld, 6525 ED Nijmegen, The Netherlands

Received 13 February 1998; revised 5 March 1998; accepted 6 March 1998

Abstract. Condensation of a properly protected uridylylated nucleopeptide with a heptadecapeptide gives after one-step deprotection the target nucleopeptide H-Gly-Ala-Tyr(5'pU)-Thr-Gly-Leu-Pro-Asn-Lys-Lys-Pro-Asn-Val-Pro-Thr-Ile-Arg-Thr-Ala-Lys-Val-Gln-OH.

Thus far, the precise function of genome-linked proteins[1] which are a signature of many mammalian and plant viruses, has eluded clarification. The terminal protein (so-called VPg) of poliovirus, discovered two decades ago,[2] is a short basic polypeptide consisting of 22 amino acids whose single tyrosine unit is covalently attached to the 5'-end of the genomic RNA *via* a O^4-(5'-uridylyl)tyrosine bond. It has been suggested[3] that the uridylylated form of VPg [*i.e.* VPgpU (**1** in Fig. 1)] is involved in priming of RNA

Figure 1



synthesis by the viral primer dependent RNA polymerase $3D^{pol}$. It was expected that synthetically prepared VPgpU (**1**) would be a useful tool in probing the validity[4] of the proposed mechanism. Consequently, synthetically prepared VPgpU (**1**) would be an invaluable tool to study in detail the validity of the proposed mechanism.

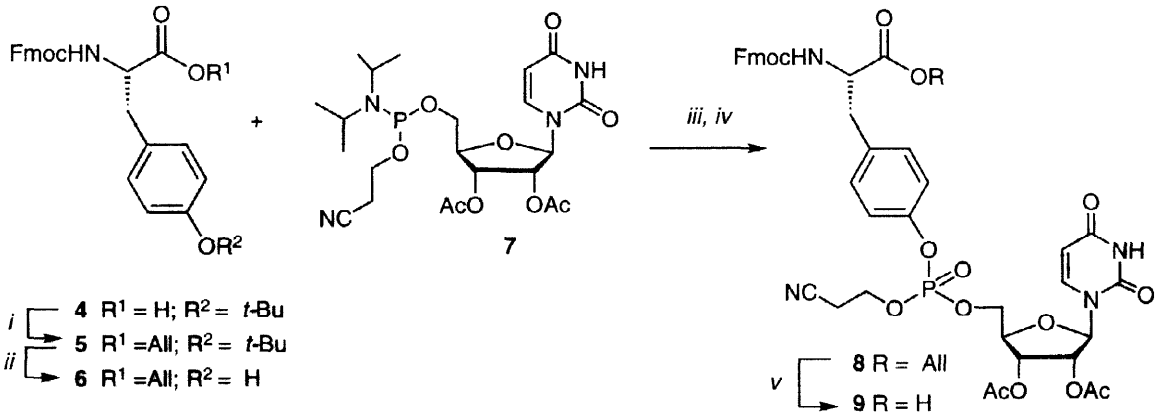
We here report that the construction of VPgpU (**1**) can be achieved by block-condensation (see Fig. 1) of the uridylylated pentapeptide **2** with the partially protected heptadecameric fragment **3**.

A standard step-wise solid support building up of the VPg peptide followed by introduction of uridine 5'-phosphate would present an ideal and straightforward route to the target molecule VPgpU (**1**). However, previous studies on the synthesis of nucleopeptides in solution[5] and on a solid support[6] indicated that a

0040-4039/98/$19.00
PII: S0040-4039(98)00558-9

successful introduction of the requisite phosphate function strongly depends on the length and primary structure of the peptide substrate. The latter uncertainty can in principle be forestalled using the suitably protected uridylylated tyrosine building unit **9**. On the other hand, block-condensation (see Fig. 1) of the uridylylated pentapeptide **2** with the partially protected heptadecameric peptide **3**, both of which are accessible *via* an on-line solid support approach, seemed to be most appropriate. Thus, coupling of **2** with **3** will not only proceed without racemization, but can also be monitored effectively by HPLC. Moreover, purification of VPgpU (**1**) was expected to be a relatively simple procedure.

Scheme 1

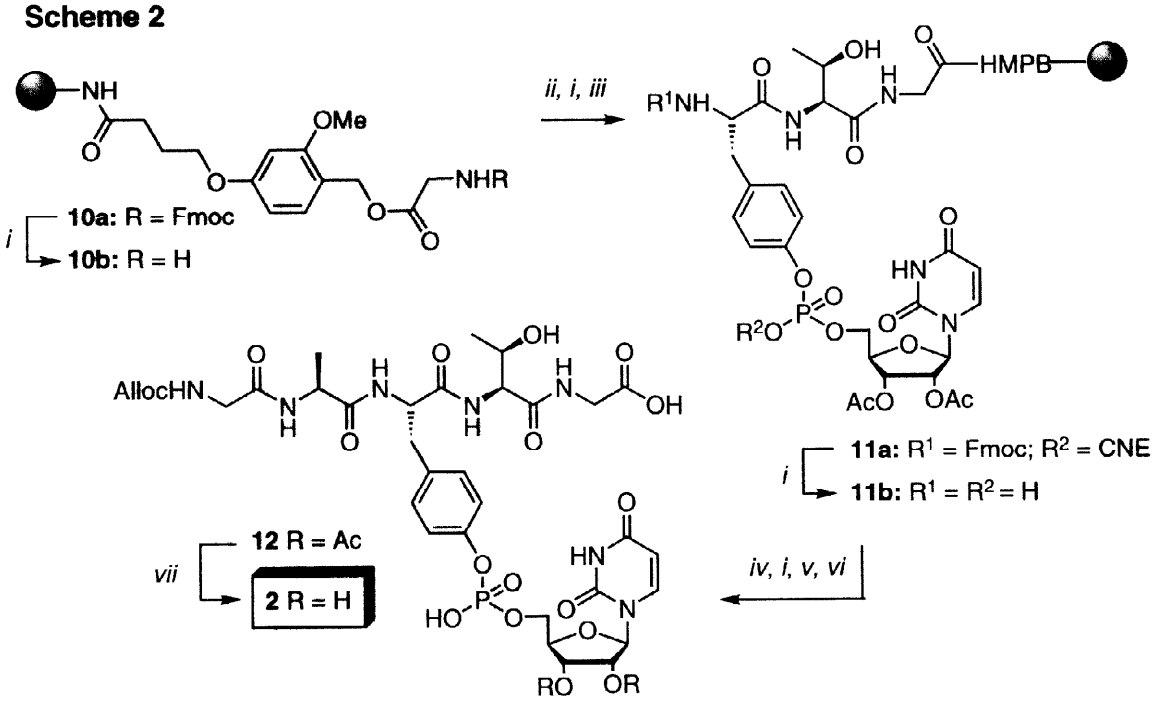


i. (a) Cs_2CO_3 (1 eq), DMF/H_2O, 10 min, (100%); (b) All-Br (3 eq), DMF, 2 h, (90%); *ii.* 70% TFA in CH_2Cl_2, 15 min, (100 %); *iii.* tetrazole (3 eq), CH_2Cl_2/CH_3CN, 20 min; *iv.* *m*-CPBA (2 eq), 0 °C, 20 min ; *v.* $Pd(PPh_3)_4$ (3 mol.%), $HSnBu_3$ (1.6 eq), AcOH (3.2 eq), THF/CH_2Cl_2, 1 h, 69% (based on **6**).

The preparation of the uridylylated derivative **9** started, as depicted in Scheme 1, with the conversion of commercially available Fmoc-Tyr(*t*-Bu)-OH (**4**) into the uridylylated tyrosine derivative **9** having a free carboxylic acid group and a β-cyanoethyl protected phosphate function. Treatment of the cesium salt of **4** with allyl bromide and subsequent removal of the *tert*-butyl protective group of **5** with TFA gave Fmoc-Tyr-OAll (**6**) in a yield of 90%. Phosphitylation of **6** with easily accessible uridine-5'-phosphoramidite **7** afforded, after *in situ* oxidation of the intermediate phosphite triester with *m*-chloroperbenzoic acid, the fully protected uridylylated tyrosine derivative **8**. Palladium(0)-catalysed hydrostannolysis[7] of **8** led, after trituration of the deallylated product with ether and subsequent purification, to the isolation of homogeneous **9**[8] in an overall yield of 70% based on **6**.

The solid phase synthesis of the nucleopeptide fragment **2** was executed (see Scheme 2) using commercially available HMPB-TentaGel® preloaded with Fmoc-Gly-OH (**10a**) as a solid support. Removal of the Fmoc-group of **10a** with DBU and subsequent elongation of **10b** were performed on an ABI 433A peptide synthesizer following the Fmoc-strategy. In the first elongation step partially protected Fmoc-Thr-OH was preactivated with TBTU/HOBt for 1 min in the presence of DIPEA and subsequently coupled (45 min) with **10b**. Removal of the Fmoc-group from the resulting immobilized dipeptide followed by condensation with the preactivated uridylylated building unit **9** gave the partially protected trimer **11a**. It is evident that the ensuing removal of the Fmoc-group in **11a** will be accompanied by β-elimination of the cyanoethyl group (CNE). However, it was expected that the presence in **11b** of the phosphodiester function would not trigger the formation of unwanted side products.[9] Indeed, sequential elongation of **11b** with preactivated Fmoc-Ala-OH and Fmoc-Gly-OH, followed by mild acidolytic cleavage from the solid support gave the partially

Scheme 2

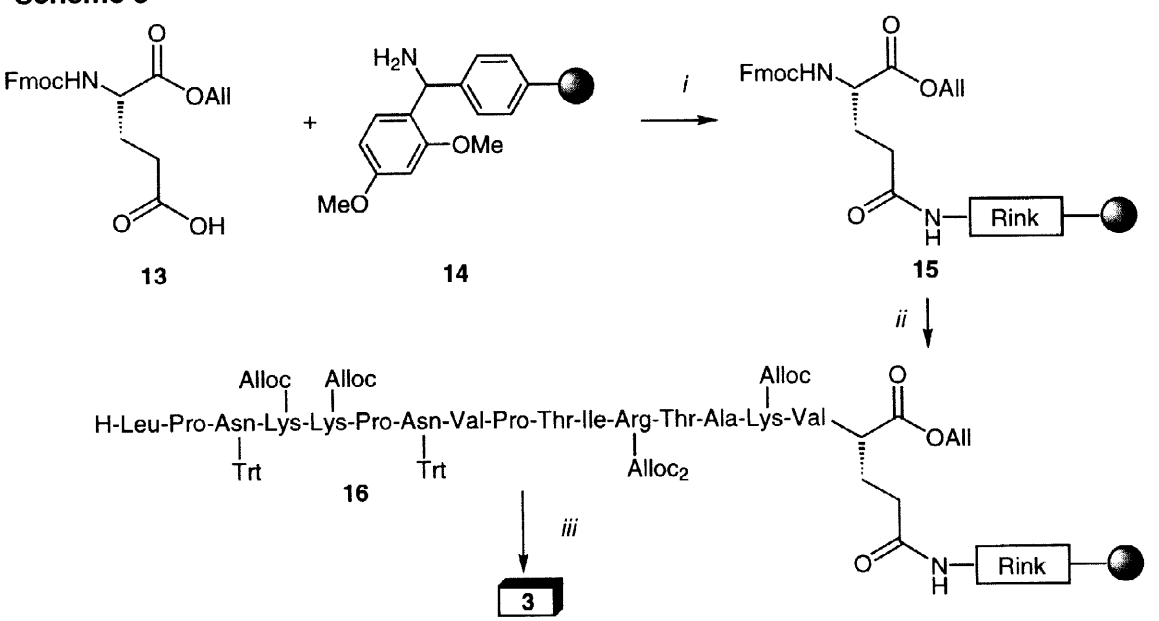


i. DBU (2.2%) in DMF, 8 min ; *ii.* **Fmoc-Thr-OH**, TBTU / HOBt; *iii.* **9**, TBTU / HOBt; *iv.* **Fmoc-Ala-OH**, TBTU / HOBt; *v.* **Alloc-Gly-OH**, TBTU / HOBt; *vi.* 2% TFA in CH_2Cl_2; *vii.* 13% aqueous NH_3, 1 h, (17%).

protected nucleopeptide fragment **12** as the main product. Mild deacetylation of purified (RP-HPLC) **12** afforded, after purification (Fractogel® HW-40) homogeneous **2**, the structure of which was firmly established by NMR as well as mass-spectroscopy.[10]

The assembly of the partially protected heptadecamer **3** (see Figure 1) could be readily accomplished (see Scheme 3) using the Rink amide MBHA resin (Novabiochem) as the solid support. Thus, condensation of commercially available Fmoc-Glu-OAll (**13**) with **14** under the influence of BOP/HOBt in the presence of

Scheme 3



i. BOP(1.2 eq.), HOBt (1.5 eq.), DIPEA (3 eq.), DMF, 20 h; *ii.* sequential elongation *via* Fmoc strategy; *iii.* (a) TFA/CH_2CH_2 (1/1, v/v),TIS, 5 h; (b) RP HPLC purification (26% overall).

DIPEA gave the immobilized glutamine derivative **15** (loading: 0.43 mmol/g), the side chain of which was attached to the solid support. Elongation of **15** with the appropriately protected amino acids, following the same Fmoc-protocol as described for the preparation of fragment **2**, led to the immobilized heptadecamer **16**. Subjection of the latter compound to TFA (50 % in CH_2Cl_2) in the presence of the cation scavenger triisopropylsilane (TIS) resulted in the removal of the trityl (Trt) groups from the side chains of the two asparagine residues and cleavage of the peptide from the solid support. Purification of the resulting partially protected heptadecamer by RP-HPLC furnished the homogeneous fragment **3**, the mass-spectrum of which was in full accordance with the proposed structure.[11]

At this stage, attention was focused on the introduction of the amide bond between fragments **2** and **3**. To this end, a mixture of BOP, HOBt and DIPEA was added to a solution of **3** and a slight excess of **2**, and the progress of the condensation was monitored by RP HPLC. Workup and purification (Sephadex LH-20, DMF/H_2O), after 20 h at 20 °C, and subsequent Pd(0)-catalysed hydrostannolysis of the allyloxycarbonyl (Alloc) and allyl group, gave after purification (RP HPLC) the target molecule VPgpU (**1**) in a yield of 17%. The homogeneity and identity of **1** was unambiguously established by capillary free-zone electrophoresis (CFZE), Edman degradation, mass-spectrometry and ^{31}P-NMR spectroscopy.[12]

A noteworthy feature of our approach to the synthesis of VPgpU (**1**) is the fact that the protecting groups (*i.e.* Alloc, All) in fragments **2** and **3** are fully compatible with the synthesis of DNA as well as RNA. The latter aspect opens the way to a future synthesis of nucleopeptides of viral origin containing defined DNA(RNA) sequences.

References and Notes

1. Salas, M. *Annu. Rev. Biochem.* **1991**, *60*, 39-71.
2. Lee, Y. F.; Nomoto, A.; Detjen, B.M.; Wimmer, E. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 59-63.
3. Wimmer, E. *Cell* **1982**, *28*, 199-201.
4. Paul, A. V.; van Boom, J.H.; Filippov, D.; Wimmer, E. *Nature*, submitted for publication.
5. Filippov, D.; Meeuwenoord, N. J.; Kuyl-Yeheskiely, E.; Wijkmans, J.C.H.M.; van der Marel, G. A.; van Boom, J. H. in Innovation and Perspectives in Solid Phase synthesis & combinatorial libraries; Epton, R. Ed. Mayflower Scientific Limited: Birmingham, 1996; pp. 367-368.
6. Dreef-Tromp, C. M.; van der Maarel, J. C. M.; van den Elst, H. *Nucleic Acids Res.* **1992**, *20*, 4015-4020. Robles,J.; Pedroso, E.; Grandas, A. *Nucleic Acids Res.* **1995**, *23*, 4151-4161 and the references cited therein.
7. Dangles, O.; Guibé, F.; Balavoine, G.; Lavielle, S.; Marquet, A. *J. Org. Chem.* **1987**, *52*, 4984-4993.
8. ^{31}P NMR ($CDCl_3$/MeOD) δ -6.9, -6.5; ^{13}C NMR (50.1 MHz, $CDCl_3$/MeOD/D_2O 45/50/5) δ 172.9 (C=O, Tyr); 169.5, 169.4 (C=O, Ac); 163.6 (C-4 uracil); 155.7 (C=O Fmoc); 150.0 (C-2 uracil); 148.4, 148.3 (C-4 Tyr); 143.3, 143.2, 140.7 (C_q Fmoc); 139.9 (C-6 uracil); 133.9 (C-1 Tyr); 130.6 (C-2, C-6 Tyr); 128.5, 127.7, 127.2, 126.6 124.8, 124.6 (CH Fmoc); 119.4 (C-3, C-5 Tyr); 116.4 (CN); 102.6 (C-5 uracil); 87.2 (C-1'); 79.7, 79.9 (C-4'); 72.1 (C-3'); 69.2 (C-2'); 66.8, 66.3, 63.0, 62.8 (OCH_2 Fmoc, C-5'); 54.3 (αC Tyr); 46.6 (CH, Fmoc); 36.5 (βC Tyr); 19.8, 19.7 (CH_3, Ac); 18.8, 19.0 ($\underline{C}H_2CN$). ESI-MS (m/z): 847 $(M+H)^+$, 864 $(M+NH_4)^+$; 869 $(M+Na)^+$; 886 $(M+K)^+$.
9. Wakamiya, T.; Saruta, K.; Yasuoka, J.; Kusumoto, S. *Chem. Lett.* **1994**, 1099-1102.
10. ^{1}H NMR (600 MHz, D_2O) δ 7.64 (d H-6 uracil *J* 8.6); 7.17 (AA'BB' 2H H-2, H-6 Tyr *J* 8.6); 7.08 (AA'BB'X 2H, H-3 H-5 J_{HH} 8.7, $^4J_{HP}$ 1); 5.94 (m 2H H-1', C$\underline{H}$ Alloc); 5.64 (d 2H, H-5 uracil, *J* 8); 5.30 (d 1H CH=C$\underline{H}$H' Alloc *J* 17.4); 5.21 (br. d 1H CH=CH$\underline{H}$' Alloc *J* 10.5); 4.64 (dd 1H αCH Tyr J_1 6.3, J_2 9); 4.55 (br. d 1H OCH_2 Alloc *J* 5.2); 4.30 (d 1H αCH Thr *J* 3.7); 4.23 (m 6H H-5'a, H-2', H-3', H-4', αCH Ala, βCH-Thr); 4.14 (m 1H H-5'b); 3.79 (AB 2H αCH_2 Gly); 3.71 (AB 2H αCH_2 Gly); 3.13 (dd 1H βCHa Tyr $J_{\beta\alpha}$ 6.3, J_{gem} -14.0); 2.96 (dd 1H βCHb Tyr $J_{\beta\alpha}$ 9.1, J_{gem} -14.0); 1.21 (d 3H, βCH_3 Ala $J_{\beta\alpha}$ 7.2); 1.13 (d 3H, γCH_3 Thr $J_{\gamma\beta}$ 6.5). ESI-MS (m/z): 902.3 $(M+2Na-H)^+$, 924.3 $(M+3Na-2H)^+$, 946.3 $(M+4Na-3H)^+$; negative mode: 427.7 $(M-2H)^{2-}$, 448.1 $(M-3H+Na)^{2-}$, 856.2 $(M-H)^-$, 878.2 $(M+Na-2H)^-$, 900.3 $(M+2Na-3H)^-$.
11. ESI-MS 2364 $(M + H)^+$, 2386 $(M + Na)^+$ (calculated for M+H 2364.3).
12. CFZE: 10.8 min (20 mM sodium citrate, 20 kV, 72 cm capillary), ^{31}P NMR (D_2O) -4.3. ESI-MS (m/z): 665.6 $(M+4H)^{4+}$, 887.2 $(M+3H)^{3+}$, 1329.3 $(M+2H)^{2+}$. MALDI-TOF (m/z): 2660.6 $(M+H)^+$ (calculated M+H 2659.5).