Available online at www.sciencedirect.com
ScienceDirect
Tetrahedron Letters
Tetrahedron Letters 47 (2006) 9349–9352

Chemical synthesis of the HPV16 E7 protein

Dmitri V. Filippov,[a] Marij J. P. Welters,[b] A. Rob P. M. Valentijn,[c] Cornelis J. M. Melief,[b] Gijsbert A. van der Marel,[a] Sjoerd H. van der Burg,[d] Herman S. Overkleeft[a] and Jan W. Drijfhout[b,*]

[a]*Leiden Institute of Chemistry, Leiden University, PO Box 9502, 2300 RA Leiden, The Netherlands*
[b]*Department of Immunohematology and Blood transfusion, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands*
[c]*Department of Clinical Pharmacy and Toxicology, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands*
[d]*Department of Clinical Oncology, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands*

Received 7 August 2006; revised 8 October 2006; accepted 19 October 2006
Available online 9 November 2006

Abstract—The full-length E7 oncoprotein of HPV16 was prepared from two peptide fragments using native chemical ligation and purified by a single gel-filtration step under denaturing conditions.

Human Papillomavirus 16 (HPV16) is the main causative agent of cervical cancer.[1] One of the two viral proteins responsible for malignant transformation is the E7 protein, a 98 amino acid residue polypeptide.[2] This oncoprotein is continuously expressed in HPV16 infected and transformed tissue. The finding that E7 acts as a tumor rejection antigen[3] makes it a promising target in the development of immunotherapy against cervical cancer. In this context, we have embarked on the evaluation of E7 and its truncated oligopeptides, in combination with various immunostimulatory agents, as potential synthetic vaccines.[4,5] Good manufacturing practice (GMP) protocol dictates the use of homogeneous E7 peptides in eventual clinical trials, and it is our belief that such material is best obtained by means of chemical peptide synthesis and HPLC purification.

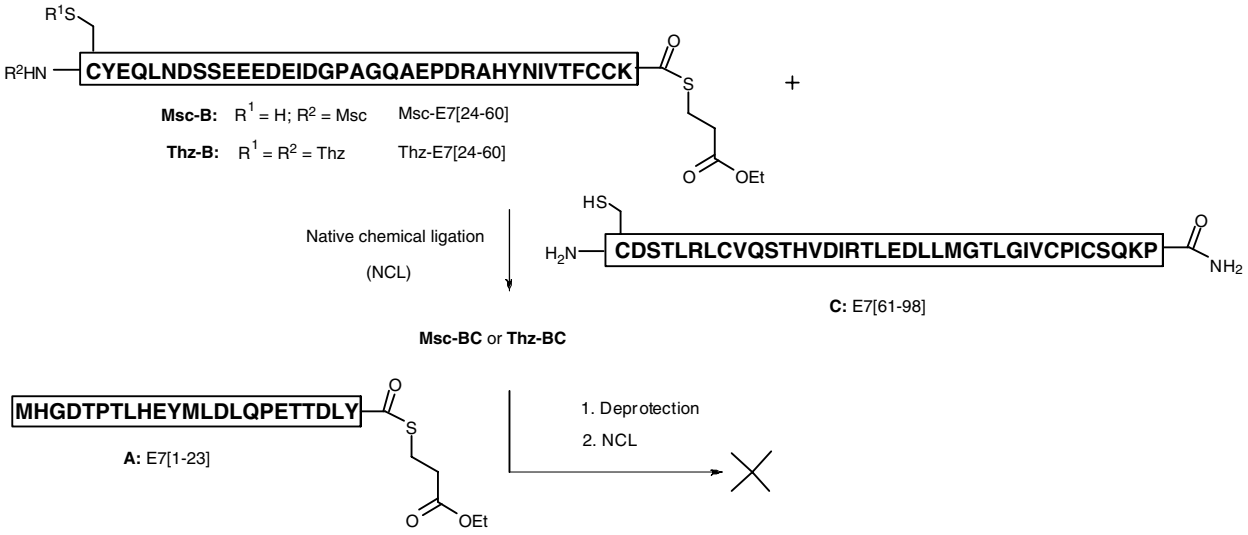
As we have shown previously, this poses no problem for the preparation of truncated E7 peptides.[6] However, the full-length E7 polypeptide is of a size that normally defies preparation using standard solid phase peptide synthesis (SPPS) protocols.[7] One paper has appeared in the literature in which the chemical synthesis of E7 using Boc/Bn based SPPS is mentioned, but experimental details on condensation procedures, purification and yield are absent.[8] We reasoned that the application of the native chemical ligation methodology, invented by Kent and co-workers,[9] would be more effective in the preparation of useful quantities of E7. In native chemical ligation (NCL), an unprotected peptide equipped with a C-terminal thioester functionality is mixed with another unprotected peptide having an N-terminal cysteine residue. After an initial thioester exchange, the peptide linkage is formed through a subsequent intramolecular N-acylation.[10] The fact that E7 contains seven cysteine residues spread throughout the polypeptide makes it highly suitable for native chemical ligation, and different retrosynthetic disconnections are possible.

In a first attempt, we studied a three-fragment strategy with disconnections at Cys24 and Cys61. The connection of more than two peptide fragments by native chemical ligation necessitates the use of a cysteine protecting group on the fragment that reacts as the thioester in the first connecting step. For our purpose we decided to compare the efficacy of the N-Msc protecting group[11] and the 1,3-thiazolidine-4-carbo (Thz) group as a masked cysteine (Scheme 1).[12]

The required fragments **A** (E7[1-23]), **Msc-B** (Msc-E7[24-60]), **Thz-B** (Thz-E7[24-60]), and **C** (E7[61-98]) were prepared and purified uneventfully employing

* Corresponding author. E-mail: J.W.Drijfhout@lumc.nl

0040-4039/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2006.10.119

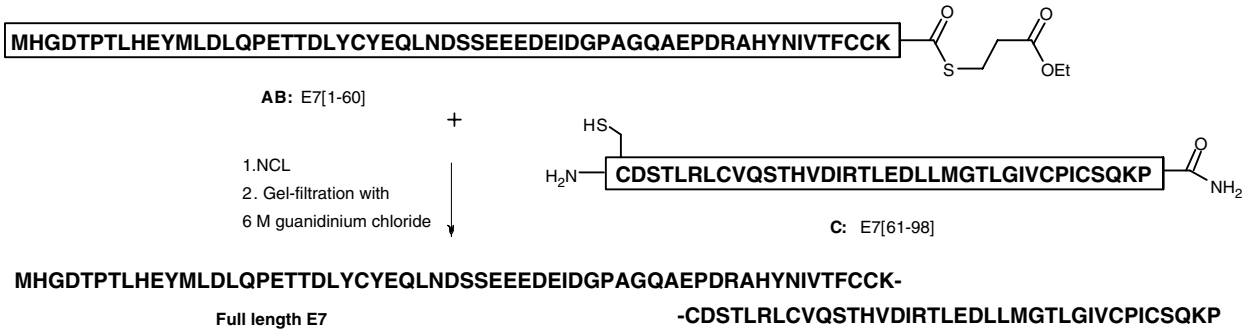


Scheme 1. Attempted synthesis of the full-length E7 protein via a three-fragment approach.

Fmoc-based SPPS using the safety catch resin[13] for peptide thioester fragments **A**, **Msc-B**, and **Thz-B**, and the RAM Tentagel resin for fragment **C**.[14] Note that we elected to synthesize E7 as its C-terminal amide, which is necessitated by the presence of a proline residue at position 98 in the natural polypeptide.[15]

Our first attempt to couple fragment **Msc-B** to **C** under standard native ligation conditions (6 M guanidine hydrochloride, 100 mM phosphate, 2% thiophenol and 2% benzyl mercaptan, pH 8) was unsuccessful. No product **Msc-BC** (Msc-E7[24-98]) was observed in the LCMS trace after either 16, 24, or 48 h. Instead, precipitation of the contents of the reaction mixture was observed, and neither addition of 20% NMP nor 20% acetonitrile altered this behavior. We reasoned that precipitation could be the result of either the formation of oligomers via disulfide bridges or the intrinsic insolubility of the E7[24-98] fragment itself. In order to exclude the former, we added 50 mM TCEP to the ligation buffer, but to no avail. The addition of 20 mM EDTA to the buffer did give the desired fragment in the N-terminal Msc protected form (**Msc-BC**), and we believe that this particular fragment is prone to aggregation under most conditions. An alternative explanation may be that thioester intermediates, formed when the side chain thiols of the internal Cys residues enter the fast thioester exchange step of NCL, are insoluble. This process then leads to sequestering of the peptide fragments from the solution precluding the productive ligation. Indeed, aggregation is a characteristic of the E7 protein itself, and has been noted in several publications[16,17] on the study of the properties of E7. The tendency of E7 to aggregate has in all likelihood prevented its structural analysis, since no NMR- or X-ray diffraction data on the protein are available. In continuation of our synthetic efforts, we encountered difficulties in the removal of the Msc protecting group.[18] The stringent conditions required for Msc removal, pH 13, proved to be detrimental to either the integrity or the solubility of the peptide, and no product was obtained. We therefore switched our attention to the preparation of fragment **BC** having the Thz protective group (**Thz-BC**). Ligating **Thz-B** to **C** to give **Thz-BC** proceeded uneventfully provided that EDTA was present in the ligation buffer to prevent aggregation. The conversion of the N-terminal thioproline residue to the corresponding cysteine with methoxyamine hydrochloride[12] went sluggishly, and it proved to be difficult to keep the ensuing deprotected fragment **BC** in solution.

We therefore decided to abandon the three-fragment approach altogether in favor of a two-fragment strategy (Scheme 2). Rather than making the E7[1-60] fragment



Scheme 2. Synthesis of the E7 protein via two-fragment ligation.

AB (corresponding to linked fragments **A** and **B** in the previous scheme) by means of native chemical ligation, we decided to investigate whether this fragment could be prepared on resin. Although the on-resin synthesis of a 60-mer oligopeptide is considered to be at the limits of standard SPPS,[7] it proved to be possible in this particular case, and fragment **AB** was synthesized as the C-terminal thioester by making use of Fmoc-based SPPS and starting from the safety catch resin. Purified fragment **AB** was added to **C** in our native ligation buffer (6 M guanidine hydrochloride, 100 mM phosphate, 2% thiophenol and 2% benzyl mercaptan, 20 mM EDTA, pH 8) in a 1:1 ratio. Productive ligation was observed, but the yield was low (2–3%) due to both unreacted **C** and extensive product aggregation leading to losses at the final RP HPLC purification stage. In a second attempt, fragment **C** was added in an 1.4-fold excess. Purification was now effected by gel filtration chromatography under denaturing conditions (6 M guanidine hydrochloride at pH 7). These strongly denaturing conditions proved to be essential for obtaining useful quantities of pure E7 (in this particular example 6.7 mg, 44% based on fragment **AB**). Other gel filtration conditions (for instance, 7 M urea at pH 7) proved abortive. Finally, we replaced the malodorous thiol catalysts benzyl mercaptan and thiophenol with odorless and water soluble sodium 2-mercaptoethane sulfonate[19] to facilitate future scale up of the synthesis.

In conclusion, we have demonstrated that native chemical ligation is an effective method for the preparation of the 98 amino acid polypeptide E7. Rather than the size of the polypeptide, it is its propensity and that of its truncated counterparts to precipitate that hampers its preparation by chemical synthesis. At least, in the particular example presented here, it pays to prepare one of the fragments for ensuing ligation as large as possible. Our two-fragment approach enables the preparation of useful quantities of E7 for vaccination studies, and in a previous paper we have already demonstrated that our synthetic material is at least as effective in eliciting immune response as the recombinant material.[4]

Synthesis of peptide thioester AB (E7[1-60]): The peptide chain was assembled starting from 4-sulfamylbutyryl AM resin (Novabiochem). The N-terminal methionine residue was introduced as Boc-Met-OH and the thioester was generated from the immobilized and protected oligopeptide essentially as described.[13] Briefly, the resin was alkylated with $TMSCHN_2$ and the peptide was cleaved with ethyl 3-mercaptopropionate in the presence of NaSPh and deprotected with TFA/ethyl 3-mercaptopropionate/TIS/*m*-cresol/H_2O 96/1.2/1.2/0.8/0.8 (v/v/v/v/v). Purification[20] with C_{18} RP HPLC gave pure peptide thioester **AB** in 5.1% overall yield. ESI MS: 1400.4 $[M+5H]^{5+}$; 1750.2 $[M+4H]^{4+}$; 2333.8 $[M+3H]^{3+}$; 2800.0 $[2M+5H]^{5+}$. MALDI MS: 3498.2 $[M+2H]^{2+}$; 6997.2 $[M+H]^{+}$. Calcd: 6997.7 (for M+H, average isotope composition).

Synthesis of peptide C: The peptide was assembled on RAM Tentagel resin (Rapp Polymere). Upon completion of the peptide chain assembly, the product was deprotected and cleaved from the resin by a 2-h treatment with the cleavage mixture TFA/EDT/TIS/*m*-cresol/H_2O (96/1.2/1.2/0.8/0.8, v/v/v/v/v). Purification[20] with C_{18} RP HPLC gave pure peptide amide **C** in a 4.8% overall yield. ESI MS: 1387.2 $[M+3H]^{3+}$; 2080.8 $[M+2H]^{2+}$; MALDI MS: 4162.2 $[M+H]^{+}$; Calcd: 4160.0 (for M+H, average isotope composition).

Optimized synthesis of the E7 protein: Peptide thioester **AB** (10.5 mg, 1.4 μmol) was mixed with peptide **C** (16 mg, 2 μmol) and dissolved in 1.8 ml of the ligation buffer (200 mM phosphate, 20 mM EDTA, 6 M guanidine hydrochloride, 75 mM MesNa, pH 8.5) and the mixture was stirred for 65 h at 22 °C. Subsequently, the reaction mixture was loaded on a Superdex 75 HL 16/60 column (Pharmacia) equilibrated in purification buffer (6 M guanidine hydrochloride, 25 mM Na_2HPO_4, 25 mM NaH_2PO_4, 5 mM EDTA, 5 mM DTT, pH 7). Elution was performed at 0.8 ml/min and the material eluting at 50 ml was collected (5.5 ml) and dialyzed against water using dialysis tubing with 3 kD cut-off (2 l water, changed 3 times over 40 h) to give 7.5 ml of a solution containing 6.7 mg (yield 44%) of HPV16-E7 protein. ESI MS: 1576.0 $[M+7H]^{7+}$; 1838.4 $[M+6H]^{6+}$; 2205.4 $[M+5H]^{5+}$; 2756.4 $[M+4H]^{4+}$; MALDI MS: 3675.78 $[M+3H]^{3+}$; 5513.15 $[M+2H]^{2+}$; 11023.61 $[M+H]^{+}$ Calcd: 11022.5 (for M+H, average isotope composition).

References and notes

1. Walboomers, J. M. M.; Jacobs, M. V.; Manos, M. M.; Bosch, F. X.; Kummer, J. A.; Shah, K. V.; Snijders, P. J. F.; Peto, J.; Meijer, C.; Munoz, N. *J. Pathol.* **1999**, *189*, 12–19.
2. Seedorf, K.; Krammer, G.; Durst, M.; Suhai, S.; Rowekamp, W. G. *Virology* **1985**, *145*, 181–185.
3. Chen, L. P.; Thomas, E. K.; Hu, S. L.; Hellstrom, I.; Hellstrom, K. E. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 110–114.
4. Welters, M. J. P.; Filippov, D. V.; van den Eeden, S. J. F.; Franken, K. L. M. C.; Nouta, J.; Valentijn, A. R. P. M.; van der Marel, G. A.; Overkleeft, H. S.; Lipford, G.; Offringa, R.; Melief, C. J. M.; van Boom, J. H.; van der Burg, S. H.; Drijfhout, J. W. *Vaccine* **2004**, *23*, 305–311.
5. Zwaveling, S.; Mota, S. C. F.; Nouta, J.; Johnson, M.; Lipford, G. B.; Offringa, R.; van der Burg, S. H.; Melief, C. J. M. *J. Immunol.* **2002**, *169*, 350–358.
6. Vambutas, A.; DeVoti, J.; Nouri, M.; Drijfhout, J. W.; Lipford, G. B.; Bonagura, V. R.; van der Burg, S. H.; Melief, C. J. M. *Vaccine* **2005**, *23*, 5271–5280.
7. Albericio, F. *Curr. Opin. Chem. Biol.* **2004**, *8*, 211–221.
8. Rawls, J. A.; Pusztai, R.; Green, M. *J. Virol.* **1990**, *64*, 6121–6129.
9. Dawson, P. E.; Muir, T. W.; Clarklewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776–779.
10. Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923–960.
11. Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068–10073.
12. Bang, D.; Kent, S. B. H. *Angew. Chem., Int. Ed.* **2004**, *43*, 2534–2538.

13. Ingenito, R.; Bianchi, E.; Fattori, D.; Pessi, A. *J. Am. Chem. Soc.* **1999**, *121*, 11369–11374.
14. Chan, W. C.; White, P. D. In *Fmoc Solid Phase Peptide Synthesis*; Hames, B. D., Ed.; Oxford University Press, 2000.
15. Chiva, C.; Vilaseca, M.; Giralt, E.; Albericio, F. *J. Pept. Sci.* **1999**, *5*, 131–140.
16. Alonso, L. G.; Garcia-Alai, M. M.; Nadra, A. D.; Lapena, A. N.; Almeida, F. L.; Gualfetti, P.; de Prat-Gay, G. *Biochemistry* **2002**, *41*, 10510–10518.
17. Clements, A.; Johnston, K.; Mazzarelli, J. M.; Ricciardi, R. P.; Marmorstein, R. *Biochemistry* **2000**, *39*, 16033–16045.
18. Tesser, G. I.; Balvertgeers, I. C. *Int. J. Pept. Protein Res.* **1975**, *7*, 295–305.
19. Flavell, R. R.; Huse, M.; Goger, M.; Trester-Zedlitz, M.; Kuriyan, J.; Muir, T. W. *Org. Lett.* **2002**, *4*, 165–168.
20. Byrnes, M. E. In *Methods in Molecular Biology: Reversed-Phase High-Performance Liquid Chromatography*. In *Peptide Analysis Protocols*; Dunn, B. M., Pennington, M. W., Eds.; Humana Press, 1994; Vol. 36.