SIMULTANEOUS SYNTHESIS OF SEQUENCE-UNRELATED PEPTIDES DERIVED FROM PROTEINS OF HUMAN PAPILLOMAVIRUSES

Viktor KRCHŇÁK[a], Josef VÁGNER[a] and Ivan HIRSCH[b]

[a] *Léčiva-Pharmaceuticals, 143 10 Prague 4 and*
[b] *Department of Experimental Virology, Intitute of Sera and Vaccines, 101 03 Prague 10*

Received April 29th, 1988
Accepted May 20th, 1988

Dedicated to the memory of Dr Karel Bláha.

Amino acid sequences derived from the early (E) and late (L) coding regions of human papillomaviruses (HPV) type 6, 11, and 16 were analysed by means of a computer program for prediction of B-cell determinants. Eight peptides ranging from 9 to 17 amino acids and containing potential B-cell determinants were selected for synthesis. The peptides were synthesized by using a configuration for solid-phase synthesis that enables the simultaneous synthesis of several peptides in flow reactors at low pressure on standard poly(styrene–1% divinylbenzene) resin. Crude peptides were obtained in good yield and purity. They were then futher purified by high-performance liquid chromatography for use as solid antigens in site-directed serology.

Relatively short synthetic peptides (10 to 20 amino acid residues) whose amino acid sequences have been derived from native proteins can maintain a certain degree of the antigenicity and immunogenicity of the parent proteins[1-3]. The antigenic properties of such peptides can be utilized for diagnostic purposes: the synthetic antigens can detect antibodies against the respective proteins (site-directed serology). Their immunogenic properties can be exploited to raise monodeterminant antipeptide antibodies that will recognize not only the peptide but also the parent protein[3]. To display parent protein antigenicity or immunogenicity, a synthetic peptide must contain a region of the protein that is co-responsible for the humoral immunological response, a so-called B-cell determinant. B-cell determinants are located on the surface of the protein, and its three-dimensional structure, if known, can help in prediction of determinants in accessible surface regions[4]. However, the tertiary structure of most proteins is not known; the primary sequence is often determined by translation of the nucleic acid sequence. Recently we have shown that B-cell determinants located in those regions of a protein molecule that have a high tendency to form a β-turn are the most promising candidates for synthetic immunogens[5].

We have applied this predictive method to select potential B-cell determinants in the early (E) and late (L) coding regions of HPV* types 6, 11, and 16. Papillomaviruses have been chosen for several reasons[6].

(i) There is no suitable system for in vitro replication of HPV and obtaining antigens in vivo is restricted to structural proteins. These are coded by late regions of DNA of only some HPV types that form warts and are not particularly important from the clinical point of view.

(ii) Genetic information arrangement is colinear in different types of HPV and many regions of their DNA possess a high degree of homology, which is even higher at the amino acid sequence level. As type-specific, noncross-reacting antigens, synthetic peptides will expectedly be more facile to prepare than long amino acid sequences obtainable by genetic engineering.

(iii) Molecular hybridization of DNA and RNA has provided evidence implicating HPV 16 in malignant changes in os uteri epithelium (c. 60% of all cases), whereas HPV 6 and 11 are associated with benign changes (c. 80% of cases). To establish during the first symptomps of a pathological process in the os uteri which type of HPV has been involved may be of great consequence for prevention and prognosis. There are no suitable hybridization techniques for routinely screening a large population in a cytologic-gynaecologic laboratory. Synthetic peptides used as antigen should provide a simple and inexpensive serological technique.

From the preparatory point of view, the synthesis of peptides the size of 10 to 20 amino acid residues is a matter of routine, thanks to Merrifield's technique of solid-phase peptide synthesis[7]. The increasing demand for synthetic peptides primarily used as synthetic antigens and immunogens has motivated the development of rapid synthetic methods. Modifications based on Merrifield's idea permit the simultaneous synthesis of multiple peptides. In 1985 Houghten described the so-called "tea bags" peptide synthesis method that allowed the preparation of several hundreds of peptides within one month[8]. Geysen devised a method for the synthesis of antigen peptides on polypropylene rods[9]. The simultaneous synthesis of peptides is based on the similarity of the individual reaction steps in the synthetic cycle. All steps that in one synthetic cycle are identical for different peptides (e.g. removal of protective groups, neutralization, washings) are carried out in common for several peptides and only the protected amino acid is condensed to each peptide separately. Recently, we described a configuration of low-pressure continuous-flow solid-phase peptide synthesis on standard resins[10] and have extended this method to the

* Standard abbreviations and symbols are used. Other abbreviations: Boc, t-butyloxycarbonyl; DCM, dichloromethane; DIEA, diisopropylethylamine; DMAP, N,N-dimethylaminopyridine; DMF, N,N-dimethylformamide; DMS, dimethylsulphide; E, early antigen; HOBt, N-hydroxybenzotriazole; HPV, human papillomavirus; L, late antigen.

simultaneous synthesis of sequence-related peptides[11]. In this contribution, the simultaneous synthesis of eight sequence-unrelated peptides derived from HPV proteins is described.

RESULTS AND DISCUSSION

Amino acid sequences derived from the early (E) and late (L) coding regions of HPV type 6, 11, and 16 (refs[12-14]) were analysed using a computer program for prediction of B-cell determinants[5]. The predictive algorithm is based on the estimation of β-turn occurrence probability, p, according to the Chou-Fasman method[15]. Peptides capable of inducing production of antipeptide antibodies that recognize the native protein are very often located in regions with a high tendency to form a β-turn[5]. Peptides comprising a tetrapeptide sequence with $p > 1.5 \times 10^{-4}$ elicit antipeptide antibodies that recognize the parent protein in 96% of cases[5]. Fig. 1 shows an example of computer output for the E6 protein of HPV 16. The predicted β-turns are denoted by stars below the amino acid sequence. Usually such prediction indicates several potential B-cell determinants. To select the most promising B-cell determinants, we applied a further criterion, high local hydrophilicity, a common approach used in the prediction of antigenic determinants[16-19]. A synoptical computer output is shown in Fig. 1, where the local hydrophilicity of a heptapeptide window moving along the protein sequence is expressed by bars centred at the middle (the fourth) amino acid. The appearance of Cys residues in predicted B-cell determinants is a complicating factor. Cysteine residues most probably form disulfide bridges in proteins and so a discontinuous determinant is formed. The second part of such a determinant cannot be located when the position of the S S bridge is not known. We therefore excluded such determinants from the selection. As a last criterion we applied similarity of protein fragments among the different types of HPV. We selected identical or similar fragments, which should detect type-common antibodies (peptides P-3-9 for E6 of HPV 11 and HPV 6B, P-4-10 for E7 of HPV 11 and HPV 6B, and P-6-15 for L2 of HPV 16, HPV 6B, and HPV 11). On the other hand, peptides that did not exhibit any mutual homology were destined for the detection of type-specific antibodies against the individual proteins.

The selected tetrapeptides with $p > 1.5 \times 10^{-4}$ are listed in Table I. For the purpose of site-directed serology they were extended at both the amino and the carboxy termini so that the β-turn sequence would find itself in the middle of the fragment. The target sequence in an immunological response does not include only the β-turn sequence, but extends to the adjacent amino acid residues, which contribute to its interaction with the immunoglobulin receptor on B-cells. By the above mentioned approach we selected eight peptides ranging from 9 to 17

amino acid residues; their sequences and locations in proteins are shown in Table II.

All the peptides were synthesized simultaneously using multiple continuous-flow solid-phase peptide synthesis (MCF SPPS) on a manually operated synthesizer of our own construction. The scheme of the synthesizer and the synthetic protocol have been published[10,11]. We used commercially available "Presep" polypropylene cartridges (Laboratorní přístroje, Prague) as flow reactors with constant inner volume. The flow reactors were concatenated and all solvents and solutions were passed through this set of reactors using moderate overpressure of innert gas charged into the reservoirs with the solvents and reagents. The synthesis was carried out under the commonly used Boc-Bzl protection strategy on *p*-methylbenzhydrylamine copoly(styrene-1% divinylbenzene) resin. The temporary protecting groups were cleaved off by TFA.

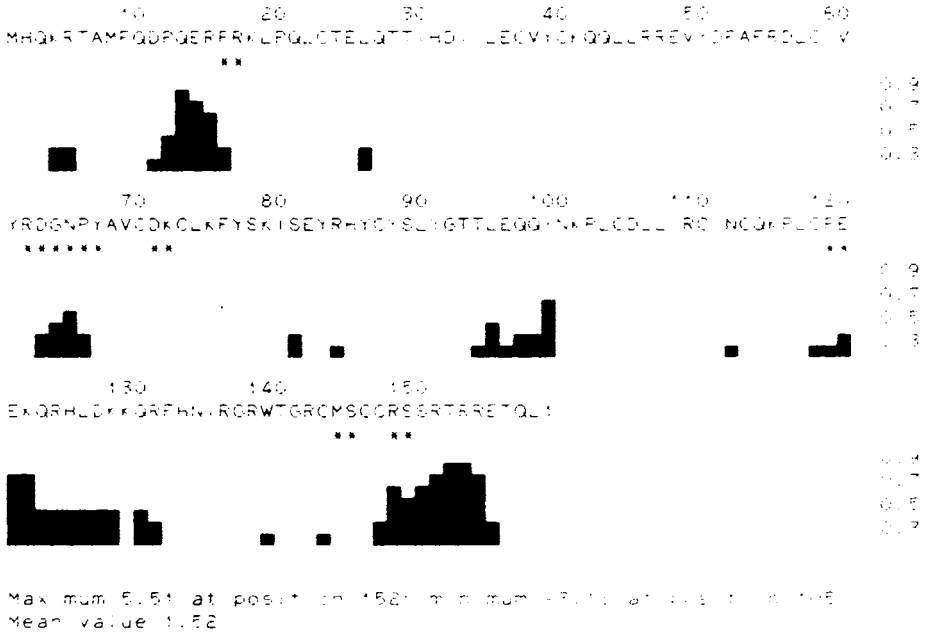

FIG. 1

Protein profile of E6 antigen of HPV 16. Asterises below the amino acid sequence indicate corner amino acids of predicted β-turns with probability of occurrence $p > 1.5 \times 10^{-4}$. The columnar graph illustrates the hydrophilicity profile of the protein in terms of the hydrophilicity coefficients of Parker at al.[19]. For comparison of hydrophilicity profiles constructed with different parameters, the hydrophilicity profile has been normalized, i.e. zero value is taken to lie on the mean hydrophilicity of the protein and the maximal and minimal values are taken as +1 or −1, respectively

Condensation of protected amino acids was carried out by 5 equivalents of a 0.2 M solution of HOBt esters in DMF. The HOBt esters were prepared in advance and stored at −20 °C. After finishing the washing cycle before the condensation, the reactors were disconnected and a solution of activated amino acid was injected into each reactor separately. The condensation was monitored by an indicative amount of 3,4-dihydro-3-hydroxy-1,2,3-benzotriazine[10,20]. After the disappearence of yellow colour a sample of peptidyl-resin was withdrawn (approximately equal to the weight gain in one synthetic cycle) and tested by ninhydrin[21]. In the case of positive test, the condensation was repeated, or the reaction was catalyzed by DMPA. After splitting off the last Boc group, the

TABLE I
β-Turn sequences of selected B-cell determinants[a]

Type HPV	No.	Sequence	p	P_h	P_t	P_s	H	
16 E6	15	Arg Pro Arg Lys	1.98	0.92	1.11	0.79	2.97	2.41
16 E7	40	Gly Pro Ala Gly	1.63	0.79	1.33	0.72	3.37	4.81
	46	Glu Pro Asp Arg	2.57	1.02	1.17	0.60	4.90	4.34
11 E6[b]	75	Asn Gln Tyr Arg	1.53	0.87	1.16	1.10	2.16	0.03
6B E6[b]								
11 E7[b]	17	Pro Pro Asp Pro	3.74	0.68	1.51	0.55	1.34	3.47
6B E7[b]	19	Asp Pro Val Gly	1.88	0.80	1.26	0.89	1.30	1.30
16 L2	96	Asp Pro Val Gly	1.88	0.80	1.26	0.89	2.53	2.71
	99	Gly Pro Ser Asp	3.11	0.73	1.49	0.65	3.54	4.17
	100	Pro Ser Asp Pro	1.73	0.73	1.48	0.60	4.17	3.56
	102	Asp Pro Ser Ile	3.10	0.86	1.22	0.86	2.21	2.84
11 L2[b]	98	Ala Pro Ser Asp	1.83	0.94	1.27	0.67	3.03	3.66
6B L2[b]	99	Pro Ser Asp Pro	1.73	0.73	1.48	0.60	3.66	3.04
	101	Asp Pro Ser Ile	3.10	0.86	1.22	0.86	2.21	2.84
6B L2	413	Thr Pro Phe Ser	1.78	0.83	1.13	0.97	1.17	1.24
	422	Leu Pro Thr Gly	1.81	0.79	1.16	0.95	1.44	0.61
16 L2	138	Ser Thr Asp Thr	1.83	0.86	1.20	0.92	6.07	5.63

[a] No., sequential number of the first amino acid in protein; p, probability of β-turn occurrence × 10^4; P_h, P_t, P_s, confomational potential of α-helix, β-turn, and sheet; H, mean hydrophilicity of heptapeptide fragment centered et corner positions of β-turn using Parker et al. hydrophilicity scale[19]; [b] sequence is identical for both HPV types.

peptidyl-resin was washed and dried. Crude peptides were obtained by "low-high" or "high" HF cleavage[22], as shown in Table III, and purified to homogeneity by HPLC.

Eight peptides were prepared by low-pressure MCF SPPS in amounts of 10–25 mg. The simultaneous synthesis of several peptides accelerated the synthesis by an order of magnitude when compared with the standard batchwise or continuous-flow methodology. Monitoring the condensation with an indicative amount of 3,4-dihydro-3-hydroxy-1,2,3-benzotriazine gives a good

TABLE II
Synthesized peptides

Peptide	HPV	Position	Amino acid sequence
P-1-12	16 E6	11–22	D P Q E R P R K L P Q L
P-2-13	16 E7	38–50	I D G P A G Q A E P D R A
P-3-9	11 E6, 6 E6	71–79	G K I N Q Y R H F
P-4-10	11 E7, 6 E7	15–24	L Q P P D P V G L H
P-5-15	16 L2	97–111	P V G P S D P S I V S L V E E
P-6-15	6B L2	96–110	P V A P S D P S I V S L I E E
	11 L2	95–109	
P-7-17	6 BL2	411–427	M G T P F S P V T P A L P T G P V
P-8-12	16 L2	134–145	S I T T S T D T T P A I

TABLE III
Yield and purity of peptides

Peptide	HF cleavage[a]	Crude peptide		Purified peptide	
		yield, mg	purity[b], %	yield, mg	R_f
P-1-12	H	41.5	73.5	13.9	0.27[c]
P-2-13	LH	34.4	62.0	10.5	0.21[c]
P-3-9	H	41.2	90.6	16.6	0.39[c]
P-4-10	H	43.5	91.4	21.5	0.38[d]
P-5-15	H	38.1	71.4	10.4	0.38[d]
P-6-15	H	45.2	69.1	12.8	0.40[d]
P-7-17	LH	49.5	69.2	18.2	0.21[e]
P-8-12	H	49.3	88.4	25.9	0.33[d]

[a] H, "high" HF cleavage; LH, "low-high" HF cleavage; [b] purity according to HPLC; [c] mobile phase 1-butyl alcohol–pyridine–AcOH–water (30:20:6:17); [d] chloroform–MeOH–AcOH–water (45:30:6:9); [e] 1-butyl alcohol–AcOH–water (4:1:1).

picture of the conversion of the reaction and in most cases gives the same result as the ninhydrin test. Nevertheless, the colour change is sometimes not pronounced enough and can be eclipsed by the colour of the peptidyl-resin. For this reason we recently developed a new method of monitoring acylation reactions with the acid-base indicator bromophenol blue[23].

A constant-volume flow reactor has, however, one shortcoming. A peptidyl-resin changes its physical properties in different solvents and also in the course of the synthesis, which is, inter alia, manifested by altered swelling capacity. If the swollen peptidyl-resin does not entirely fill up the reactor volume, cavities form, the reactants flow around the resin, and the reactor loses in efficiency. This disadvantage is nullified by using flow reactors with moving pistons, where the changes in the volume of swollen resin can be compensated. Such reactors are nowadays routinely used in our laboratory[10,11].

The described configuration of peptide synthesis in flow reactors facilitates the simultaneous synthesis of several peptides without compromising the purity or yield of the crude products. The method is not demanding for laboratory equipment and combines the advantages of multiple peptide synthesis batchwise with those of continuous-flow procedures.

The results of site-directed serology obtained with the peptides described in the present paper will be published elsewhere.

EXPERIMENTAL

Materials and General Procedures

Solvents, DIEA (Fluka) and TFA (Riedel de Haën) were redestilled before use, HOBt (Fluka) was recrystallized. Other chemicals were used as purchased: N,N-dicyclohexylcarbodiimide, *p*-cresol, *p*-thiocresol, DMS (dimethylsulfide), di-tert-butyl dicarbonate, DMAP and 3,4-dihydro-3-hydroxy-4-1,2,3-benzotriazine (Fluka); *p*-methylbenzhydrylamine copoly(styrene-1% divinylbenzene) resin was from Chemical Dynamics Co.

Side-chain functional groups were protected as follows: Arg, by the *p*-toluene sulphonyl group, Asp, Glu, Ser, Thr, and Tyr, by the benzyl group; Lys, by the *o*-bromobenzyloxycarbonyl group; and Trp, by the formyl group. Analytical and preparative HPLC was carried out isocratically on a chromatograph equipped with an HPP 5001 pump (Laboratorní přístroje, Prague), a UVM 4 detector (Vývojové dílny ČSAV), a steel column, 17 × 250 mm, and a glass column, 26 × 130 mm, for preparative purpose, and steel columns 4.6 × 250 mm for analytics. The columns were packed with stationary phase Separon SGX C18, 5 μm (analytical column) and 10 μm (preparative column). Pre-coated Silikagel 60 F-254 plates (Merck) were used for thin-layer chromatography, detection was by chlorine–tolidine. Samples for amino acid analysis were hydrolysed in 6 M HCl at 110°C for 24 h in the presence of phenol in sealed ampules and were analysed on an AAA 881 amino acid analyser (Mikrotechna, Prague).

The scheme of the synthesizer was described in a previous paper. Presep cartriges of internal volume 0.8 ml (Laboratorní přístroje, Prague) were used as flow reactors.

The HPV sequences analysed were taken from refs[12,13].

Multiple Continuous-flow Solid-phase Peptide Synthesis

Each reactor was charged with 100 mg of *p*-methylbenzhydrylamine copoly(styrene-1% divinylbenzene) resin (0.4 mmol/g). Reactors were concatenated and connected to the synthesizer. Before the first amino acid was condensed, the resin was washed with DCM, neutralized with 5% DIEA in DCM for 20 min, and again washed with DCM. The individual steps of the synthetic cycle and the preparation of HOBt esters were described in ref.[10]. Solutions of 0.2 M HOBt esters in DMF were prepared for the entire synthesis in advance and were stored at −20°C. Before the condensation of the HOBt esters the solution was tempered to ambient temperature, a polypropylene syringe was charged with 1 ml of this solution, an indicative amount of 3,4-dihydro-3-hydroxy-1,2,3-benzotriazine was added, and the solution was injected in several doses into the reactors, previously washed with DMF. After the disappearance of the yellow colour a sample of peptidyl-resin was withdrawn and tested by ninhydrin[21]. If the test was positive, fresh HOBt ester was injected into the cartridge, or the reaction was catalyzed by DMAP (10 mg). After splitting off the Boc group from the last amino acid, the peptidyl-resin was washed with DCM, MeOH, and dried in a stream of nitrogen. Crude peptides were obtained by "low-high" or "high" HF treatment[22], as shown for individual peptides in Table III. Peptidyl-resin was cleaved in 10 ml of a mixture that contained 6.5 ml DMS, 2.5 ml HF, 250 mg *p*-thiocresol, and 750 mg *p*-cresol ("low-high" HF treatment), or 9 ml HF, 250 mg *p*-thiocresol, and 750 mg *p*-cresol ("high" HF treatment). After removing DMS and/or HF, the scavengers were extracted by cold ethyl acetate, the crude peptide was dissolved in 20% AcOH, resin was filtered off, and the solution was lyophilized.

Crude peptides were dissolved in mobile phase (aqueous MeOH containing 0.1% TFA) and injected (5 mg per run) onto preparative columns. Elution profiles were monitored at 235 nm. Lyophilized chromatographically pure fractions were passed through a Biogel P4 column (2.6 × 100 cm).

TABLE IV
Amino acid analysis

Peptide	P-1-12	P-2-13	P-3-9	P-4-10	P-5-15	P-6-15	P-7-17	P-8-12
Ala	–	3.01	–	–	–	1.02	1.07	1.00
Arg	0.96	1.03	1.07	–	–	–	–	–
Asp	0.95	2.02	1.08	1.03	1.06	1.07	–	1.12
Glu	3.08	2.07	0.90	0.92	2.28	2.21	–	–
Gly	–	2.00	1.04	1.02	1.07	–	2.08	–
His	–	–	1.05	1.03	–	–	–	–
Ile	–	0.98	0.93	–	0.73	1.83	–	2.14
Leu	2.13	–	–	2.04	1.06	1.17	1.03	–
Lys	1.02	–	0.91	–	–	–	–	–
Met	–	–	–	–	–	–	0.97	–
Phe	–	–	1.04	–	–	–	1.02	–
Pro	3.85	1.90	–	2.93	3.05	3.05	4.94	1.00
Ser	–	–	–	–	2.85	2.74	1.00	2.12
Thr	–	–	–	–	–	–	–	4.63
Tyr	–	–	0.96	–	–	–	2.86	–
Val	–	–	–	1.03	2.90	1.89	2.07	–

The purity of products was controlled by analytical HPLC, TLC (mobile phases and R_F in Table III), and amino acid analysis (Table IV).

REFERENCES

1. Berzofsky J.A.: Science *229*, 932 (1985).
2. Van Regenmortel M.H.V.: Protides Biol. Fluids *34*, 81 (1986).
3. Walter G.: J. Immunol. Methods *88*, 149 (1986).
4. Novotný J., Handschumachar M., Bruccoleri R.E.: Immunol. Today *8*, 26 (1987).
5. Krchňák V., Mach O., Malý A.: Anal. Biochem. *165*, 200 (1987).
6. Gissmann L.: Cancer Surv. *3*, 161 (1984).
7. Merrifield R.B.: J. Am. Chem. Soc. *85*, 2149 (1963).
8. Houghten R.A.: Proc. Natl. Acad. Sci. U.S.A. *82*, 5131 (1985).
9. Geysen H.M., Meloen R.H., Barteling S.J.: Proc. Natl. Acad. Sci. U.S.A. *81*, 3998 (1984).
10. Krchňák V., Vágner J., Flegel M., Mach O.: Tetrahedron Lett. *28*, 4469 (1987).
11. Krchňák V., Vágner J., Mach O.: Int. J. Pep. Protein Res., submitted.
12. Schwarz E., Dürst M., Demankowski C., Lattermann O., Zech R., Wolfsperger E., Suhai S., zur Hausen H.: EMBO J. *2*, 2341 (1983).
13. Dartmann K., Schwarz E., Gissmann L., zur Hausen H : Virology *151*, 124 (1986).
14. Seedorf K., Kraemmer G., Dürst M., Suhai S., Roewekamp W.G.: Virology *145*, 181 (1985).
15. Chou P.Y., Fasman G.D.: Adv. Enzymol. *47*, 46 (1978).
16. Kyte J., Doolittle R.F.: J. Mol. Biol. *157*, 105 (1982).
17. Hopp T.P., Woods K.R.: Proc. Natl. Acad. Sci. U.S.A. *78*, 3824 (1981).
18. Fraga S.: Can. J. Chem. *60*, 2606 (1982).
19. Parker J.M.R., Guo G., Hodges R.S.: Biochemistry *25*, 5425 (1986).
20. Atherton E., Cameron L., Meldal M., Sheppard R.C.: J. Chem. Soc., Chem. Commun. *1986*, 1763.
21. Kaiser E., Colescott R.L., Bossinger C.D., Cook P.I.: Anal. Biochem. *34*, 595 (1970).
22. Tam J.P., Heath W.F., Merrifield R.B.: J. Am. Chem. Soc. *105*, 6442 (1983).
23. Krchňák V., Vágner J., Šafář J., Lebl M.: Collect. Czech. Chem. Commun. *53*, 2542 (1988).

Translated by the author (V. K.).