International Journal of Medicine and Medical Sciences. Vol. 1(9) pp. 383-387, September, 2009.
Available online at http://www.academicjournals.org/ijmms

Full Length Research Paper

Penta- and hexapeptide sharing between HPV16 and *Homo sapiens* proteomes

Darja Kanduc

Department of Biochemistry and Molecular Biology, University of Bari, Italy. E-mail: d.kanduc@biologia.uniba.it, dkanduc@gmail.com. Tel: +39 080 544 3321. Fax: +39 080 544 3321.

Accepted 21 August, 2009

The primary sequence of the HPV16 proteome was analyzed for penta- and hexa-peptide sequences shared with human proteins. The following data were obtained: 1) HPV16 and *Homo sapiens* proteomes share thousands of identical peptide motifs; 2) the overlaps are inter-dispersed among human proteins involved in fundamental housekeeping functions as well as crucial processes such as cell growth, differentiation and neurosensory regulation. The data are discussed in relationship to the potential cross-reactivity risk of an HPV16 vaccine.

Key words: Similarity analysis, sequence-sequence peptide matching, viral versus human proteome overlapping, vaccine-related cross-reactions.

INTRODUCTION

Notwithstanding current innovative research, and improved diagnostic and detection methods, infectious diseases remain a leading cause of death worldwide. The re-emergence of new infectious agents, uncontrolled use of antimicrobial drugs and pesticides with the consequent development of resistant pathogens and sexual habits have contributed to the recrudescence of diseases such as cholera, malaria, tuberculosis and cervical cancer (Church, 2004; Centurioni et al., 2005; Alavez-Ramírez et al., 2007).

This scenario underscores the need for vaccines able to eradicate disease-associated pathogens. However, the complex and intricate relationships, still largely unknown, between the human host and viral/bacterial pathogens continue to pose considerable challenges to the reaching of effective immunotherapies. One main obstacle in this regard is represented by vaccine-related adverse effects, including autoimmunity (Gilboa, 2001; Naicker et al., 2007; Maina et al., 2009).

Here, we examine the potential HPV16 immune cross-reactivity by carrying out a comparative study of HPV16 and *Homo sapiens* proteomes. Using similarity analysis, we define the level of penta- and hexa- mer overlapping between the HPV16 polyprotein and human proteins, and document that the human proteome shares a high number of peptide motifs with the HPV16 proteome. The matches are inter-dispersed among human proteins with relevant biological functions. The data are discussed in relationship to potential cross-reacting autoimmune reactions of an HPV16 vaccine.

METHODS

The HPV16 oncoprotein primary sequence (taxonomy ID: 333760) was dissected into penta- or hexa- mers that were analyzed for sequence similarity to the human proteome using PIR perfect match (pir.georgetown.edu/pirwww/search/fasta). The pentamer (or hexamer) sequences were offset by one residue that is, overlapping by four (or five) residues: MHQKR, HQKRT, QKRTA, etc. or MHQKRT, HQKRTA, QKRTAM, etc). The function of the human proteins and potential disease associations were analyzed using the universal protein resource (UniProt; www.uniprot.org/uniprot).

RESULTS

Scientific literature indicates 5 to 6 amino acids as sufficient minimal antigenic determinants (Niman et al., 1983; Oldstone, 1998; Lucchese et al., 2007; Kanduc, 2008), so that an immune peptide block can be defined as having a minimal epitopic length of 5 amino acids (Lucchese et al., 2007; Kanduc, 2008). Accordingly, the entire HPV16 polyprotein sequence was dissected into antigenic 5- or 6-mer blocks that were used as probes to measure the HPV16 similarity level to human proteins. The results of the similarity analysis are reported in Figure 1, where the

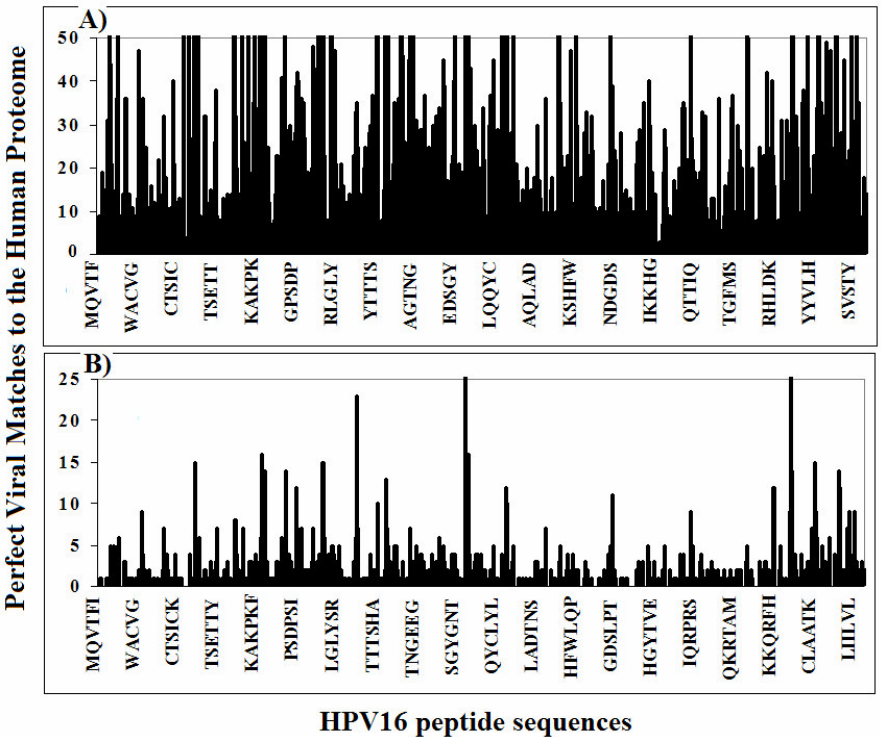

Figure 1. Similarity profile of HPV16 polyprotein to the human proteome by using (A) 5-mer or (B) 6-mer peptide sequences as probes. Penta- or hexamer position along the HPV16 polyprotein sequence is reported.

Table 1. Sharing of 5- and 6-mer motifs between HPV polyprotein and human proteome.

Motif	Similarity level[a]	Human proteins sharing viral 5-mers[b]	
	N° of matches	Total n°	%
5-mers	28,948	15,361	42.5
6-mers	2,010	1,703	4.7

[a]Similarity level refers to identical matches occurring between HPV16 polyprotein and human proteins, and includes the count of repeated/multiple matches within human proteins. [b]Human proteome is taken equal to 36,103 proteins at the time of the analysis.

profile represents the number of matches to the human proteome over the sequential viral 5- or 6-mer peptides. Figure 1, panel A, clearly documents that pentamer peptide blocks from the HPV16 polyprotein are widely and repeatedly spread in the human proteome for thousands of identical pentapeptide overlaps. Likewise, Figure 1, panel B illustrates a high level of hexapeptide sharing between the viral and human proteomes.

The numerical details are given in Table 1. Importantly, Table 1 documents that more than 40% of the 36,103 human proteins forming the *Homo sapiens* proteome contain viral HPV16 pentamers, and more than 4% of the human proteome hosts a perfect viral hexamer match.

Qualitatively, the peptide overlapping between HPV and *Homo sapiens* proteomes involves essential proteins endowed with essential functions such as cell-adhesion molecules, leukocyte differentiation antigens, enzymes, proteins associated with spermatogenesis, transcription factors and neuronal antigens. Table 2 reports examples of the human proteins having hexameric matches in common with HPV16 late major capsid protein L1. It can be seen that the crucial human proteins share hexa-peptide sequences with HPV16 L1 protein. To mention only a few: LAMA1 or laminin subunit alpha-1, involved in the positive regulation of epithelial cell proliferation (Olsen et al., 1989); GHR or growth hormone receptor, defects of which are a cause of Laron dwarfism characterized by dysmorphic facial features, truncal obesity, subnormal

Table 2. Examples of hexapeptide sharing between HPV16 late major capsid protein L1 and human proteins.

Sequence[1]	Pos[2]	Human protein(s) hosting the viral sequence[3]
LPPVPV	39	GHR
PVPVSK	41	EPS8 Q86W54
VPVSKV	42	EPS8 Q59FK2 Q96I15
TDEYVA	50	Q5T4E0 Q5T4E1 Q9HBD4 SMCA4
AGTSRL	63	K0141 Q96RF9 Q9C0C1
GTSRLL	64	Q5T630 Q5T631 Q6P2M3 Q9C0C1
RLLAVG	67	Q3LID0 Q4G186 Q6ZRH9 Q8N239 Q8NAC2 Q9NXC2
KILVPK	85	CDYL1 Q32NC5 Q9BWZ2
GVEVGR	130	COBA1 Q5VT31
GVGISG	140	AT10A AT8A1 AT8A2 AT8B1 AT8B2 AT8B4 Q6ZR02 Q6ZSP3
GISGHP	142	CO4A6 EMID1 HXC13 Q5JYH8
PGDCPP	208	Q6NXT9 Q6PKN4 Q8N5E1 Q8NEP7 Q9BQ72
PPLELI	212	Q4VXG5 Q9P225
LELINT	214	IC1 Q5UGI6 ZN37A
ELINTV	215	DYN2 DYN3 Q5W128 Q6P2G1
TLQANK	238	LAMA1
LQANKS	239	LAMA1
KYPDYI	256	SMS2
KSEVPL	243	ELMO3 MYO1B
SEVPLD	244	DMBT1 ELMO3 Q5JR20 Q5JR25
RAGAVG	289	CP007 EMID1 Q8NDR2 SPC21
AGAVGE	290	CO1A2 Q15177 Q8N5U5 Q8TDN0
AVGENV	292	BACH2
VGENVP	293	BACH2
GSGSTA	305	CADH2 MMRN2 Q6S377 Q6S378 Q6S379 Q6S380 Q6S381 Q6S382 Q8NB64 Q9ULG1 RCBT1 PLEC1
PTPSGS	319	Q6P996 Q6ZVH1 Q8IYP8 Q8N9V6 Q8TBS5 Q9HA91
TPSGSM	320	KCNQ3 Q8IYP8 Q8N9V6 Q9HA91 WDR35
AAISTS	372	ANR11 EDG2 Q9NVU7
AISTSE	373	ANR11 Q9H2F7 Q9UK61
ISTSET	374	ATS5 Q499Z3 Q49AG8 Q8N7V7
STSETT	375	Q6ZTK3 Q7RTT7 RBP22 RBP23 RBP26 RBP2 RGPD8
TSETTY	376	O43418 O43419 O43421 Q6W763
FQLCKI	402	CBLB CBL
TLTADV	408	CENPE Q4LE75 TBCE
GLQPPP	431	MICA1 Q5BLQ2
LQPPPG	432	BRSK1 PBX4 Q13577 Q8IXF5 Q8IXF6 Q8N2E7 Q8TB25 Q99581
QPPPGG	433	AKIB1 PBX4 Q5VWW5 Q8IXF5 Q8IXF6 Q8N2E7
PPPGGT	434	Q9BU61 Q9Y3Z0 ZIM10
PGGTLE	436	Q96SF2 Q9UJS3
GGTLED	437	NIBA Q96SF2 Q9HC80 Q9UJS3
TPPAPK	458	NOTC4 Q53YB1 Q5CZI7 TAU Q5SPL1 Q5STG5 Q5XWF0
PPAPKE	459	ANTR2 SAFB1 SAFB2 TAF6
VNLKEK	475	FIBG LSD1 Q7Z664
NLKEKF	476	MNDA
EKFSAD	479	SPT6H
KFSADL	480	SPT6H
ADLDQF	483	NCOA1 Q2T9G5 Q53SX3
FLLQAG	494	Q5T9V9 Q6ZNS7
LLQAGL	495	IFT80 MLC1 Q6X960 Q86YF4

Table 2: Contd.

KATPTT	512	C1QR1 CATW RBM15
ATPTTS	513	C1QR1 Q8N1F8 Q96G50 Q96KG9 Q9HAW5
TPTTSS	514	O15054 Q2NKQ4 Q8WXI7 TCRG1
PTTSST	515	Q59EI7 Q5JT55 Q7Z679 Q96NQ2 SPTN4 TBX15
TTSSTS	516	DYR1A Q59EI7 Q59G41 Q5SW77 Q5SW78 Q5SW79 Q5W067 Q5W073 Q5W074 Q8WXI7 RBM23 SIX5
TSSTST	517	MUC5B Q14879 Q6ZN80 Q6ZRA4 Q8N9T9 Q9NYE4
SSTSTT	518	MYPT1 Q6ZRG8 Q86Y92 Q9BSR0 Q9H065 TOP2A
STSTTA	519	O43418 Q7Z704 Q9H3Q7 Q9UKW9
KRKKRK	525	BRD7 DHX16 K0553 MBB1A PCDH9 Q5JP45 Q5VT82 Q5VTZ1 Q5W0G2 Q8N3K7 SIRT1 UB7I1

[1]Viral sequence, [2]position in the viral protein, [3]accession number of the human protein hosting the viral sequence.

rate of growth (Goddard et al., 1995); Q5XWF0 or neurofibrillary tangle protein filament-tau involved in Alzheimer disease (Rademakers et al., 2004); SMS2 or sphingomyelin synthase 2, that can increase atherogenic poten-tial (Dong et al., 2006); COBA1 or collagen alpha-1(XI) chain, defects of which is the cause of syndromes characterized by bone disorders, sensorineural deafness, ocular disorders including juvenile cataract, myopia, strabismus, vitreo-retinal or chorio-retinal degeneration, retinal detachment, and chronic uveitis (Annunen et al., 1999); NOTC4 or neurogenic locus notch homolog prot-ein 4, that affects the implementation of differentiation, proliferation and apoptotic programs (Sugaya et al., 1997); MNDA or myeloid cell nuclear differentiation antigen, found in promyelocyte stage and also appears in myeloblast cells in some cases of acute myeloid leukaemia (Briggs et al., 1992); PO3F2 or brain-specific transcription factor 2, with a role in positive regulation of cell proliferation (Xe et al., 1989).

DISCUSSION

The last decades have seen large-scale implementation of preventive/therapeutical immunization procedures and, in parallel, concerns about possible adverse effects of vaccines seem to rise. Unfortunately, vaccine-caused adverse reactions cannot be easily separated from vacine-independent events occurring by chance in temporal association. Actually, vaccine-caused adverse reactions generally may occur within months/years from the immunization time. A crucial example is the mass immunization with hepatitis B vaccines (HBV) in the early 1990s (Denis et al., 1998), followed years later by reports of association between HBV immunization and the onset of multiple sclerosis (Herroelen et al., 1991; Marshall,1998).

In this context, analysing the cross-reactivity potential associated with vaccine antigens is mandatory. This study seems to support the view that anti-HPV immunization may be ineluctably associated with adverse side effects. Theoretically, the cross-reactivity potential in using viral antigens is equal to zero, by being 1 in 20^5 (that is, one out 3,200,000), the mathematical probability of 20 amino acids occurring in five identical residues be-tween two proteins. Likewise, the theoretical probability of 20 amino acids occurring in six identical residues between two proteins is 1 in 20^6, that is equal to 1 out 64,000,000. However, conflicting with theoretical data, the numbers reported in Table 1 show the highest probability of cross-reactions, given the highest number of perfect exact matches between the viral and human proteomes. This contrast between theoretical versus actual values in the number of peptide overlaps is a powerful warning incum-bent on the future of vaccine development and delivery. Analysis of the quantitative results reported in Table 1 and the qualitative data exposed in Table 2 indicate that the logical consequence of cross-reactions following anti-HPV vaccine administration may be possibly represented by alterations in epithelial cell proliferation (Olsen et al., 1989), obesity and subnormal rate of growth (Goddard et al., 1995); Alzheimer's disease (Rademakers et al., 2004); increased atherogenic potential (Dong et al., 2006); bone disorders; sensorineural deafness; ocular disorders including juvenile cataract, myopia, strabismus, retinal degeneration and detachment, and chronic uveitis (Annunen et al., 1999); alteration of differentiation, proli-feration and apoptosis (Sugaya et al., 1997); alterations in myeloid cell nuclear differrentiation (Briggs et al., 1992); alteration in brain-specific regulation of cell proliferation (Xe et al., 1989), to cite only a few.

In conclusion, it seems that vaccine safety monitoring becomes more and more important with new vaccines, intensive vaccine recommendations, and new expanded immunization initiatives. In this scenario, the molecular antigen dissection described in the present paper may be the basic platform for avoiding possible cross-reactive hot

spots and achieving high standards of safety (Kanduc, 2009).

REFERENCES

Alavez-Ramírez J, Castellanos JR, Esteva L, Flores JA, Fuentes-Allen JL, García-Ramos G, Gómez G, López-Estrada J (2007). Within-host population dynamics of antibiotic-resistant M. tuberculosis. Math. Med. Biol. 24: 35-56.

Annunen S, Körkkö J, Czarny M, Warman ML, Brunner HG, Kääriäinen H, Mulliken JB, Tranebjaerg L, Brooks DG, Cox GF, Cruysberg JR, Curtis MA, Davenport SL, Friedrich CA, Kaitila I, Krawczynski MR, Latos-Bielenska A, Mukai S, Olsen BR, Shinno N, Somer M, Vikkula M, Zlotogora J, Prockop DJ, Ala-Kokko L (1999). Splicing mutations of 54-bp exons in the COL11A1 gene cause Marshall syndrome, but other mutations cause overlapping Marshall/Stickler phenotypes. Am. J. Hum. Genet. 65: 974-983.

Briggs JA, Burrus GR, Stickney BD, Briggs RC (1992). Cloning and expression of the human myeloid cell nuclear differentiation antigen: regulation by interferon alpha. J. Cell. Biochem. 49: 82-92.

Centurioni MG, Puppo A, Merlo DF, Pasciucco G, Cusimano ER, Sirito R, Gustavino CA (2005). Prevalence of human papillomavirus cervical infection in an Italian asymptomatic population. BMC Infect. Dis. 5: 77.

Church D (2004). Major factors affecting the emergence and re-emergence of infectious diseases. Clin. Lab. Med. 24: 559-586.

Denis F, Goudeau A, Aufrère A (1998). Vaccination against hepatitis B in France in 1996. Bull. Soc. Pathol. Exot. 91: 37-40.

Dong J, Liu J, Lou B, Li Z, Ye X, Wu M, Jiang XC (2006). Adenovirus-mediated over-expression of sphingomyelin synthases 1 and 2 increases the atherogenic potential in mice. J. Lipid Res. 47: 1307-1314.

Gilboa E (2001). The risk of autoimmunity associated with tumor immunotherapy. Nat. Immunol. 2: 789-792.

Goddard AD, Covello R, Luoh SM, Clackson T, Attie KM, Gesundheit N, Rundle AC, Wells JA, Carlsson LM (1995). Mutations of the growth hormone receptor in children with idiopathic short stature. The Growth Hormone Insensitivity Study Group. N. Engl. J. Med. 333: 1093-1098.

Herroelen L, De Keyser J, Ebinger G (1991). Central-nervous-system demyelination after immunisation with recombinant hepatitis B vaccine. Lancet. 338: 1174-1175.

Kanduc D (2008). Immunogenicity in peptide-immunotherapy: from self/nonself to similar/dissimilar sequences. In: Sigalov A, (ed.) Multichain Immune Recognition Receptor Signaling: From Spatiotemporal Organization to Human Disease, Austin, TX, USA, Landes Biosci. pp. 198-207.

Kanduc D (2009). Epitopic peptides with low similarity to the host proteome: towards biological therapies without side effects. Expert Opin. Biol. Ther. 9: 45-53.

Lucchese G, Stufano A, Trost B, Kusalik A, Kanduc D (2007). Peptidology: short amino acid modules in cell biology and immunology. Amino Acids 33: 703–707.

Maina G, Albert U, Bogetto F, Borghese C, Berro AC, Mutani R, Rossi F, Vigliani MC (2009). Anti-brain antibodies in adult patients with obsessive-compulsive disorder. J. Affect. Disord. 116: 192-200.

Marshall E (1998). A shadow falls on hepatitis B vaccination effort. Science 281: 630-631.

Naicker S, Fabian J, Naidoo S, Wadee S, Paget G, Goetsch S (2007). Infection and glomerulonephritis. Semin. Immunopathol. 29: 397-414.

Niman HL, Houghten RA, Walker LE, Reisfeld RA, Wilson IA, Hogle JM, Lerner RA (1983). Generation of protein-reactive antibodies by short peptides is an event of high frequency: implications for the structural basis of immune recognition. Proc. Nat. Acad. Sci. USA. 80: 4949–4953.

Oldstone MB (1998). Molecular mimicry and immune-mediated diseases. FASEB J. 12: 1255-1265.

Olsen D, Nagayoshi T, Fazio M, Peltonen J, Jaakkola S, Sanborn D, Sasaki T, Kuivaniemi H, Chu ML, Deützmann R, Timpl R, Uitto J (1989). Human laminin: cloning and sequence analysis of cDNAs encoding A, B1 and B2 chains, and expression of the corresponding genes in human skin and cultured cells. Lab. Invest. 60: 772-782.

Rademakers R, Cruts M, van Broeckhoven C (2004). The role of tau (MAPT) in frontotemporal dementia and related tauopathies. Hum. Mutat. 24: 277-295.

Sugaya K, Sasanuma S, Nohata J, Kimura T, Fukagawa T, Nakamura Y, Ando A, Inoko H, Ikemura T, Mita K (1997). Gene organization of human NOTCH4 and (CTG)n polymorphism in this human counterpart gene of mouse proto-oncogene Int3. Gene 189: 235-244.

Xe X, Treacy MN, Simmons DM, Ingraham HA, Swanson LW, Rosenfeld MG (1989). Expression of a large family of POU-domain regulatory genes in mammalian brain development. Nature 340:35-42.