

VIRUS AND BACTERIAL MEMBRANE PROTEINS

Two viruses and seven bacteria have been chosen as examples to illustrate the structures of membrane proteins.

Influenza

All the pathogenic subtypes and mutations of the influenza virus so far identified have an abundance of particular prolyl peptides (**XPY**, where **P** is proline and one or both of the adjacent amino acids, **X** and **Y**, are hydrophilic amino acids – love water) in their hemagglutinin and neuraminidase surface membrane proteins. (see Figure 1: the hemagglutinin and neuraminidase proteins are the hair-like structures on the surfaces of the viruses) In addition, these ‘active’ prolyl peptides are largely conserved in the hemagglutinins and neuraminidase proteins - essential for host cell invasion and proliferation and in the former, for causing an immune response that damages the host’s organs - in each variant of the influenza virus for which the structures are known.

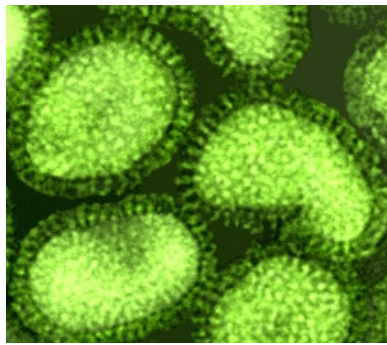


Figure 1: influenza virus

For example, the sequence of N1 neuraminidase in the Japan/China H5N1 avian influenza (Mase, M., Eto, M., Tanimura, N., Imai, K., Tsukamoto, K., Horimoto, T., Kawaoka, Y., Yamaguchi, S. "Isolation of a genotypically unique H5N1 influenza virus from duck meat imported into Japan from China" *Virology* 339 (1), 101-109 (2005)) has multiple potential sites for ginger enzyme hydrolysis (**bolded**): these sites are adjacent to P-3, 48, **93**, **120**, 154, **167**, **169**, **198**, **246**, **272**, **283**, **302**, **326**, 328, 340, **377**, **410**, **420**, **431**, and **458**, and the blue highlighted prolines are conserved over the three H5N1, the H1N1 swine flu of 1918 and 2009, and H9N7 bird flu, with the pink highlighted prolines are not conserved in the H9N7 bird flu:

```

1  MNPNQKITTI GSICMVIGIV SLMLQIGNII SIWVSHSIQT GNQHQAEPCN QSIITYENNT
61 WVNQTYVNIS NTNFLTEKAV NLVTLAGNSS LCPISGWAVY SKDNGIRIGS KGDVVFIREEP
121FISCSHLECR TFFLTQGALL NDKHSNGTVK DRSPHRTLMS CPVGEAPSPYNSRFESVAWS
181ASACHDGTSW LTIGISGPDN GAVAVLKYDG IITDTIKSWR NNILRTQESE CACVNGSCFT
241VMTDGPSNGQ ASYKIFRIEK GKVVKSAELNAPNYHYEECSYPDAGEITCVCRDNWHGSN
301RPWVSFNQNL EYRIGYICSG VFGDNPRPNDGTGSCGPVSPKGAYGIKGFSFRYGNGVWIG
361RTKSTNSRSG FEMIWDPNGW TGTDSNFSVK QDIVAITDWSGYSGSFVQHPELTGLDCIRP
421CFWVELIRGR PKESTIWTSG SSISFCGVNS DTVGWSWPDG AELPFTIDK
    
```

Each letter in the sequence represents an amino acid. The important hydrophilic amino acids adjacent to prolines are: asparagine, **N**; glutamate, **E**; cysteine, **C**; serine, **S**; histidine, **H**; aspartate, **D**; arginine, **R**; lysine, **K**; tyrosine, **Y**.

Cleavage of the neuraminidase proteins by the ginger enzyme will prevent the virus invading the host cells and prevent proliferation of the virus in the host.

The identified epitopes for hemagglutinin are not linear peptides but are what is called “conformational” epitopes in which the participating amino acids are brought together via the three-dimensional structure of the protein. Seven of the H1N1 2009 (Swine flu) active prolyl peptides are included in the conformational epitopes and six of these are conserved prolines in H1 and H5. Numerous other non-proline amino acids in the epitopes are not conserved. (Deem, M.W., Pan, K. “The epitope regions of H1-subtype influenza A, with application to vaccine efficacy” *Protein Eng., Design & Selection*, 1-4 (2009, July 3))

EPITOPES (CONFORMATIONAL)

```

A:   RQLSSFERFPKSWPNHDKGTWGD
B:   VSCPHAGAFKDKGKE LVL GIHH
C:   DTVLENVVTH AFAMER   AGSSSHTQPKNTLPFQNI
D:   AYIVDLLVKKGNSYPLSSSDQSLYQNADTYVFV SKKFKPVDERNYY
E:   VNLEKHNLLGKCNIAG LGNPETFEATGLR
    
```

Underlined letters are three or more amino acid linear peptides.

Yellow highlighted letters are part of an active proline peptide group which is cleaved by the ginger enzyme.

The inhibitors currently on the market function differently to the ginger enzyme: they are specifically designed to 'plug' the active site of neuraminidase where the neuraminidase opens the virus' surface membrane so the virus can enter the host's cells. They rely on the amino acids that surround the active site (glutamate, arginine and aspartate) to stabilize the binding of the plug in the hole. (see Figures 2). The above N1 variant in H5N1 bird flu does not have arginine, R, in positions 92 and 371 and subsequently the binding and efficacy of the 'plug' could be significantly reduced.

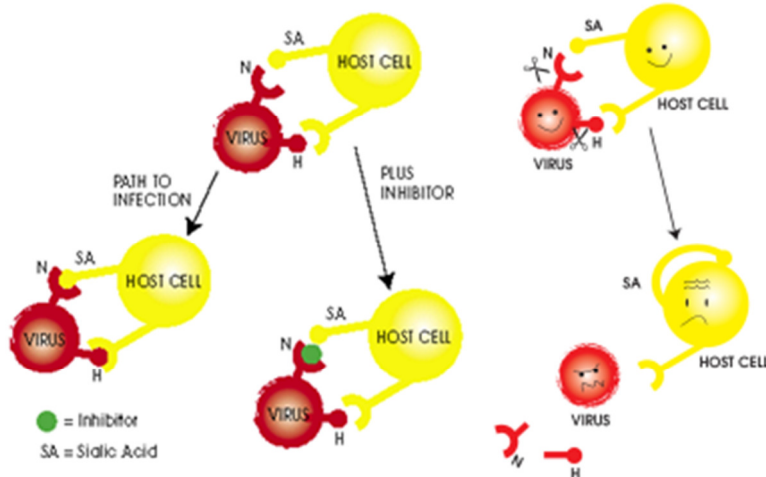


Figure 2: The influenza virus invades a host cell by initially binding to receptors on the host cell, one of which binds to the hemagglutinin (H) on the virus, and one with sialic acid (SA) which binds to the neuraminidase (N) protein on the virus. Current inhibitors such as Relenza are designed to mimic the sialic acid and to bind to the neuraminidase blocking the link to the host cell. The Biohawk ginger product acts (on the right) like a pair of scissors and specifically cuts off the hemagglutinin and neuraminidase proteins from the surface of the virus completely preventing the virus infecting the cell and replicating itself.

Importantly, although the various H5 structures show significant mutations, the potential sites for ginger enzyme hydrolysis are largely conserved and number at least 15. In the H5-hemagglutinins recently found in the Vietnam (Nguyen T.D., Hanh T.H., Puthavathana P., Long H.T., Buranathai C., Lim W., Webster R.G., Hoffmann E. "Lethality to Ferrets of H5N1 Influenza Viruses Isolated from Humans and Poultry in 2004" *J. Virol.* **79**, 2191-2198 (2005)) and Japan/China variants, and previously identified in Singapore H5N1 avian influenza (Ha, Y., Stevens, D.J., Skehel, J.J., Wiley, D.C. "Structure of Avian H5 Haemagglutinin Complexed with LSTA" *Proc. Nat. Acad. Sci. USA*, **98**, 11181 (2001)), the prolines are conserved. In addition there is an extra proline, P-233, in the Singapore H5N1, which has the hydrophilic arginine and lysine adjacent to it, but in the Vietnam and Japan/China variants, proline is replaced by serine. The hydrophilic amino acids adjacent to the prolines are also conserved, except for those adjacent to P-101 with asparagine (N-100) for the Vietnam and Singapore variants and serine (S-100) for the Japan/China protein, P-108 with the non-hydrophilic glycine for the Vietnam and Japan/China variants and the hydrophilic glutamate adjacent to the proline for the Singapore variant - giving an additional site for hydrolysis, P-134 with lysine for the Vietnam and Japan/China proteins and with arginine for the Singapore protein, and for P-337, which has glutamine following it for all three but with serine before the proline for the Vietnam and Japan/China variants and valine for the Singapore protein. The target prolines for ginger enzyme hydrolysis are: P-65, 81, 90, 101, 108 (Singapore only), 134, 174, 197, 210, 227, 233 (Singapore only) 251, 266, 297, 312, 319, 337, 506.

This gives the ginger enzyme excellent opportunity to hydrolyze the H5 at multiple sites. The conservation of the prolyl residues in the hemagglutinin structures is suggestive of these having a specific role in the function of the protein. Hydrolysis by Biohawk's ginger at these specific sites in the neuraminidase and hemagglutinin proteins, would cleave them from the virus inhibiting its ability to invade host cells and to proliferate. Further, hydrolysis of the hemagglutinin structure would prevent this viral protein from stimulating damaging cytokine fluxes. Independent studies have confirmed the ginger enzyme inhibits H5N1 bird flu (Selleck, P "Efficacy of Zingibain in inactivating H5N1 Avian Influenza Virus" Report July 2007- A/chicken/Vietnam/8/2004 **H5N1**)

Papilloma Virus

All capsid proteins of papilloma viruses are proline-rich with a high degree of conservation of the proline peptides. The known structures of the proteins associated with common warts and with anogenital papilloma infections are as follows. The proline peptides are highlighted:

PV L1A Sequences-warts common

CPV1	. MWR RPS DNKLYV PPP APVSKVLTDDAYVTRTKIFYHASSSRLLAVG NPYFPIR K.....ANKTIV PK VSGFQF	67
RHPV1R	MSMWR RPS DSKVYLL PP ...PVSKVSTDEYVSRTSIYYHAGSSRLLAVG HPY YAVKK..GNNKVS V PKVSGLQY	68
HPV29	MALWRSSDNLVYL PP .TPVSKVISTDD....YVTRTNIYYAGSSRLLTVG HPHY YSI PK SSGNKVDV PK VSAFQY	70
HPV2a	MALW RPN ESKVYLL PP .TPVSKVI..STDVYVTRTNVYYHGGSSRLLTVG HPY YSIKK....SNNKVAV PK VSGYQY	69
HPV27	MALW RPN ESKVYLL PP .TPVSKVI..STDVYVTRTNVYYHGGSSRLLTVG HPY YSIKK..GSNNRLAV PK VSGYQY	70
HPV57	MAMW RPN ESKVYLL PP .TPVSKVL..STDVYVTRTNVYYHGGSSRLLTVG HPY YSIKK SGNNKVS V PKVSGYQY	70
HPV26	MALWRTSDSKVYLL PP .TPVSRVVNTDE...YVTRTGIIYYAGSSRLLTLG HPY FSI PKTGQKAEI PK VSAYQY	69

CPV1	RVFKIVL.. PD PNKFAL P DTSFIDSTSQRLLWACI...GLEVGRG QPL GVGYCG HP CLNKFDDEVENSASYAV NP GQDNR	141
RHPV1R	RVFRVRL PD PNKFGL P DANFY DP NTQRLVWACLGVVEVGRG QPL GVGTSG HP LLNKLDDTENG PK VAGGQGDNR	142
HPV29	RVFRVRL PD PNKFGL P DARIYN PE AERLWVWACTGVVEVGRG QPL GVGLSG HP LYNKLNDDTENSINIAHAENGQDSR	144
HPV2a	RVFHVKL PD PNKFGL P DADLY DP DQRLWACVGVVEVGRG QPL GVGVSG HP YYNRLDDTENAHT PD ..ADDGR	141
HPV27	RVFHVKL PD PNKFGL P DADLY DP DQRLWACVGVVEVGRG QPL GVGVSG HP YYNRQDDTENAHTLDS..AEDGR	142
HPV57	RVFHVKL PD PNKFGL P DANLY DP DQRLWACVGVVEVGRG QPL GVGISG HP YYNKDDTENSHN PD ..ADDGR	142
HPV26	RVFRVHL PD PNKFGL P DPQLY NP DTERLVWACVGVVEVGRG QPL GLIGLSG HP LFNKLDDTENSHLATVNADTDNR	143

CPV1	VNVAMDYKQTQLCLVGC AP PLGEHWGKGKQCSGVSVQDGD CP PLELVTSVIQDGMVDTGFGAMDFAELQSNKS215	
RHPV1R	ECVSMYKQTQLCMLGC BP PVGEHWGKGN NP C..TTGAAGD CP ALELVNSVIQDGMVDTGYGAMDFNALQANKS	214
HPV29	DNIADVYKQTQLCILG CP PMGEHWGKGTVCARTSSAAGD CP PLELMTTHIEDGDMVDTGYGAMDFALQVNSK	218
HPV2a	ENISMDYKQTQLFILG CP PIGEHWSKGTTC.NGSSAAGD CP PLQFTNTTIEDGDMVETGFGALDFATLQSNKS	214
HPV27	ENISMDYKQTQLFILG CP SIGEHWSKGTTC.NGSSAAGD CP PLQFTNSTIEDGDMVETGFGALDFATLQSNRS	215
HPV57	EYISMDYKQTQLFILG CP PIGEHWSKGTTC.SGSSAVGD CP PLQFTNTTIEDGDMVETGFGALDFALQSNKS	215
HPV26	DNVSDVNKQTQLCIIG CP PLGEHWGIGTICKTNTQTRGD CP PLELISIIEDGDMIDTGFAMDFALQANKS	217

CPV1	DVPLDICTSTCKY PD YLQMAAD PY GDRLFFYLREKQMFARHFFNRAGTVGEQ IP DELVFKGTT...SRATVSSN	286
RHPV1R	DV PI DICTSVCKY PD YLKMASD PY GDLSFFYLRRQMFVRHLFNRAGTMGDSV PD DLYIKGSG...SNVKLASH	285
HPV29	DVPLDICQSTCKY PD YLGMAAD PY GDSMFFFLRREQLFARHFFNRAGVVGDK IP DSL YLKGNN...GRE TP GS	289
HPV2a	DVPLDICTNTCKY PD YLKMAAE PY GDSMFFSLRREQMFRHFFNLGGKMGDT IP DELYIKSTS...V PT PS	284
HPV27	DVPLDICTNVCKY PD YLKMAAE PY GDSMFFSLRREQMFRHFFNRAGKMGDT IP DELYIKSTT...IS DP GSH	285
HPV57	DVPLDICTNICKY PD YLKMAAD PY GDSMFFSLRREQMFRHFFNRGGSMGDAL PD ELYKSST...VQ TP GSY	285
HPV26	DV PI DISQSTCKY PD YLKMSADTYGNSMFFFLRREQLFARHFYFNKAGAVGDAI PT TLYIKGAES..GRE PT SS	289

CPV1	IYFN TP SGSLVSSEAQLFN KPY WLRHAQGHNNGICWGNLTVVDTTRSTNMTVCASTTSS SP ...SATYASE	357
RHPV1R	VFY TP SGSMVTSDAQLFN KPY WLQKAQGHNNGICWGNQVFLTVVDTTRSTNMTLCASTASTV... TP YNNES	356
HPV29	IY S TP SGSMVTSQAIFN KPY WLQQAQGHNNGICWANQVFLTVVDTTRSTNMTLCASTTSS QP ...LTTYDATK	360
HPV2a	VYT ST PSGSMVSSEQQLFN KPY WLRRAQGHNNGMCWGNRVFLTVVDTTRSTNVSLCATEASDT...NYKATN	353
HPV27	VYT ST PSGSMVSSEQQLFN KPY WLRRAQGHNNGMCWGNRIFLTVVDTTRSTNVSLCAAEVSDN...TNYKATN	355
HPV57	VYT ST PSGSMVSSEQQLFN KPY WLRRAQGHNNGMCWGNRIFLTVVDTTRSTNVSLCATVTTET...NYKASN	354
HPV26	IYSA TP SGSMVTSDAQLFN KPY WLRQAQGHNNGICWGNLQVFTVVDTRSTNLTISTLSAASA... ST PK PS D	360

CPV1	YKQYMRHVEEFDLQFIFQLCKITLTAELMAYIHT NP TVLEEWNFGL SP PPNGTLEDYRYVQSQAITCQ K.P .	429
RHPV1R	FKEYLRHVEEFDLQFIFQLCKVTLNTEVMAYIHSMDASILEDWNFGL Q PPPSGSLQDTRYFVTSAAITCQ K.P .	428
HPV29	IKEYLRHGEEYDLQFIFQLCKVTL TE IMAYLHTMNSALLEDWNFGL TL PPSTLEDTRYFVTSAAITCQ K.D .	432
HPV2a	FKEYLRHMEEYDLQFIFQLCKITL TE IMAYIHNMD DP QLLEDWNFV PP PPSASLQDTRYRYLQSQAITCQ K.P .	425
HPV27	FKEYLRHMEEYDLQFIFQLCKITL TE IMAYIHNMD DP QLLEDWNFV PP PPSASLQDTRYRYLQSQAITCQ K.P .	427
HPV57	YKEYLRHMEEYDLQFIFQLCKITL TE IMAYIHNMDARLLEDWNFV PP PPSASLQDTRYRYLQSQAITCQ K.P .	426
HPV26	YKQFIRHGEEYELQFIFQLCKITLTDVMAYIHLMNASILEDWNFGL TL PP TASLEDAYRFIKNSATT CQR.N .	432

CPV1	T.PD KEK QDP YAGLSFWEVNLKEKFSSELEQ YPL GRKFLLQTVGQSTSLARAG...TKRAA.....STST.AT P	493
RHPV1R	A.PPK EKED PL AKYTFWEVDLKEKFSADLDQ FPL GRKFLLQAGMRAR PT LRA P ...KRTAS...STSS SPR	493
HPV29	L.A PT EKQ DPY AKLNFWDVLDKDRFTLDLSQ FPL GRKFLLQIGARRRSV PSR ...KRRT.....TTT APT PA	496
HPV2a	T.PPK T PTD PYASLTFWDVLDSEFSMDLDQ FPL GRKFLLQAGM PT VSRK...AAVS...GT TP . PTS	487
HPV27	T.PPK T PTD PYANMTFWDVLDRESFSMDLDQ FPL GRKFLLQAGT... P P TV SR.KRT	479
HPV57	T.PPK T PTD PYATMTFWDVLDSEFSMDLDQ FPL GRKFLLQAG T .. PT VSRK...RAAA...TAAA. PTA	487
HPV26	A.PP PK ED PF QFKFKFDVLDLKEKFSIDLQ FPL GRKFMLQAGIQRR PK LGTK... RPL S.....STS..SST	494

CPV1	TR .KKVKRK.....	501
RHPV1R	KR.KRTKR.....	500
HPV29	KR.KRSKK.....	503
HPV2a	KR.KRVRR.....	494
HPV27	AV.GRGH.....	485
HPV57	KR.KKVVRR.....	494
HPV26	KR.KKRKLT.....	503
HPV52	KK.KVKR.....	503
HPV58	KR.KKVKK.....	498
HPV67	RK.KVKR.....	500

HPV L1A Sequences-Anogenital

HPV32	MSVW RPS DNKVYLL PP .PPVSKVSTDEYVQRTNYFYHASSSRLLAVG HPY YTIKK....TPNRTS IP KVSGLQY	69
HPV11R	.. MWR PS DSSTVYV PP PN P VSKVATDAYVTRTNIFYHASSSRLLAVG HPY YSIKK....VNKT V PKVSGYQY	67
HPV6bR	.. MWR PS DSSTVYV PP PN P VSKVATDAYVTRTNIFYHASSSRLLAVG HPY YSIKR....ANKT V PKVSGYQY	67
HPV18R	MALW RPS DNTVYLL PP .PSVARVVNTDD...YVTRTSIFYHAGSSRLLTVGN NPY FRV P AGGGNKQD IP KVSAQYQ	70
HPV16R	MSLWL P SEATVYLL PP .VPVSKVSTDEYVARTNIYYHAGTSRLLAVG HPY F PK IKK..PNNNKIL V PKVSGLQY	70
HPV31	MSLW RPS EATVYLL PP .VPVSKVSTDEYVTRTNIFYHAGSRLLTVG HPY YSI PK SDN P KKIV V PKVSGLQY	71
HPV33	MSVW RPS EATVYLL PP .VPVSKVSTDEYVSRTSIYYAGSSRLLAVG HPY FSIK NP TNAKLL V PKVSGLQY	71

HPV32	RVFRVRL	PDFN	KFTL	PETN	LYNP	ETQRM	VWAC	VGLE	VGRG	QPL	LG	VLSG	HPL	LN	RLLD	DTENG	PR	YAAG	PG	TDNR	143																																																			
HPV11R	RVFKVVL	PDFN	KFAL	PDSS	LFDP	TTQRL	VWACT	..GLE	VGRG	QPL	LG	VVSG	HPL	LN	KYDD	VNSG	GGY	GGN	PG	QDNR	141																																																			
HPV6bR	RVFKVVL	PDFN	KFAL	PDSS	LFDP	TTQRL	VWACT	..GLE	VGRG	QPL	LG	VVSG	HPL	LN	KYDD	VNSG	GGY	GGN	PG	QDNR	140																																																			
HPV18R	RVFRVRL	PDFN	KFGL	PD	TSIYN	PETQ	RLVW	ACAG	VEIG	RGQ	PL	GV	LSG	HPL	FYN	KLDD	TESS	HAATS	SNV	SE	DVR	144																																																		
HPV16R	RVFRVRL	PDFN	KFGF	PD	TSFYN	PETQ	RLVW	ACAG	VEIG	RGQ	PL	GV	LSG	HPL	LN	KLDD	TENAS	AYA	ANAG	V	DNR	144																																																		
HPV31	RVFRVRL	PDFN	KFGF	PD	TSFYN	PETQ	RLVW	ACAG	VEIG	RGQ	PL	GV	LSG	HPL	LN	KFDD	TENS	NR	YAG	PG	TDNR	145																																																		
HPV33	RVFRVRL	PDFN	KFGF	PD	TSFYN	PETQ	RLVW	ACAG	VEIG	RGQ	PL	GV	LSG	HPL	LN	KFDD	TETG	NKY	PG	Q	GADNR	145																																																		
HPV32	ENV	SMDCK	QTL	CL	VGC	KPA	IGE	HWG	KGA	ACSA	..Q	SNGD	CP	PLE	LQNS	VIQD	GDM	AD	VFG	AM	FSA	LQ	TSKA	215																																																
HPV11R	VNV	GMDY	KQ	TQ	LC	MV	GC	AP	PL	GE	HWG	KGT	QCS	NTS	VQ	NGD	CP	PLE	LITS	VIQD	GDM	VDT	G	FAM	NFADL	LQ	TNKS	215																																												
HPV6bR	VNV	GMDY	KQ	TQ	LC	MV	GC	AP	PL	GE	HWG	KGT	QCS	NTS	VQ	NGD	CP	PLE	LITS	VIQD	GDM	VDT	G	FAM	NFADL	LQ	TNKS	214																																												
HPV18R	DNV	SVDY	KQ	TQ	LC	IG	CA	PA	IGE	HWA	KGT	ACK	SR	LS	QGD	CP	PLE	LKNT	VLED	GDM	VDT	G	YG	AM	FSTL	QD	TKC	218																																												
HPV16R	ECIS	M	DY	KQ	TQ	LC	IG	CA	PA	IGE	HWG	KGS	SP	C	T	N	AV	NP	GD	CP	PLE	LINT	VIQD	GDM	VDT	G	FAM	DFTL	LQ	ANKS	218																																									
HPV31	ECIS	M	DY	KQ	TQ	LC	IG	CA	PA	IGE	HWG	KGS	SP	C	T	N	AV	NP	GD	CP	PLE	LKNS	VIQD	GDM	VDT	G	FAM	DFTL	LQ	TKS	219																																									
HPV33	ECL	S	M	DY	KQ	TQ	LC	IG	CA	PA	IGE	HWG	KGS	SP	C	T	N	AV	NP	GD	CP	PLE	LINT	VIQD	GDM	VDT	G	FAM	DFTL	LQ	ANKS	218																																								
HPV32	EV	PL	D	I	M	S	I	S	KY	PD	Y	L	K	M	S	A	E	A	Y	G	D	N	M	F	F	L	R	R	E	Q	M	F	V	R	H	L	F	N	R	A	G	T	L	G	E	P	V	P	E	D	M	I	K	A	S	N	G	A	S	R	N	L	A	S	289							
HPV11R	DV	PL	D	I	C	G	T	V	C	KY	PD	Y	L	Q	M	A	A	D	PY	G	D	R	L	F	F	L	R	K	E	Q	M	F	A	R	H	F	F	N	R	A	G	T	V	G	E	P	V	P	D	L	L	V	K	G	G	N	...	N	R	S	S	V	A	S	286							
HPV6bR	DV	PL	D	I	C	G	T	T	C	KY	PD	Y	L	Q	M	A	A	D	PY	G	D	R	L	F	F	L	R	K	E	Q	M	F	A	R	H	F	F	N	R	A	G	E	V	E	P	V	P	D	L	L	I	K	G	S	G	...	N	R	T	S	V	G	S	285								
HPV18R	EV	PL	D	I	C	Q	S	I	C	KY	PD	Y	L	Q	M	S	A	D	PY	G	D	S	M	F	F	L	R	R	E	Q	L	F	A	R	H	F	W	N	R	A	G	T	M	G	D	T	V	P	Q	S	L	I	K	G	T	...	M	R	A	S	P	G	S	289								
HPV16R	EV	PL	D	I	C	T	S	I	C	KY	PD	Y	L	K	M	V	S	E	PY	G	D	S	L	F	F	L	R	R	E	Q	M	F	V	R	H	L	F	N	R	A	G	T	V	G	E	N	V	P	D	D	L	Y	I	K	G	S	G	...	S	T	A	N	L	A	S	289						
HPV31	NV	PL	D	I	C	N	S	I	C	KY	PD	Y	L	K	M	V	A	E	PY	G	D	L	F	F	L	R	R	E	Q	M	F	V	R	H	F	F	N	R	S	G	T	V	G	E	S	V	P	T	D	L	Y	I	K	G	S	G	...	S	T	A	T	L	A	N	S	290						
HPV33	DV	PL	D	I	C	G	S	T	C	KY	PD	Y	L	K	M	T	S	E	PY	G	D	S	L	F	F	L	R	R	E	Q	M	F	V	R	H	F	F	N	R	A	G	T	L	G	E	A	V	P	D	D	L	Y	I	K	G	S	G	...	T	T	A	S	I	Q	S	289						
HPV32	IY	Y	P	T	P	S	G	S	M	V	T	S	D	A	Q	I	F	N	KPY	W	L	Q	A	Q	G	H	N	N	G	I	C	W	G	N	Q	V	F	L	T	V	V	D	T	R	S	T	N	M	T	V	C	A	T	V	T	E	D	...	T	Y	K	S	T	N	358							
HPV11R	IY	V	H	T	P	S	G	S	L	V	S	S	E	A	Q	L	F	N	KPY	W	L	Q	A	Q	G	H	N	N	G	I	C	W	G	N	H	L	F	V	T	V	D	T	R	S	T	N	M	T	L	C	A	S	V	S	K	A	...	T	Y	T	N	S	355									
HPV6bR	IY	V	N	T	P	S	G	S	L	V	S	S	E	A	Q	L	F	N	KPY	W	L	Q	A	Q	G	H	N	N	G	I	C	W	G	N	L	F	V	T	V	D	T	R	S	T	N	M	T	L	C	A	S	V	T	T	S	...	T	Y	T	N	S	354										
HPV18R	VY	S	P	S	P	S	G	S	I	V	T	S	D	S	Q	L	F	N	KPY	W	L	H	K	A	Q	G	H	N	N	G	I	C	W	H	N	Q	L	F	V	T	V	D	T	R	S	T	N	L	T	I	C	A	S	T	Q	S	P	V	...	P	G	Q	Y	D	A	T	K	360				
HPV16R	NY	F	P	T	P	S	G	S	M	V	T	S	D	A	Q	I	F	N	KPY	W	L	Q	A	Q	G	H	N	N	G	I	C	W	G	N	L	F	V	T	V	D	T	R	S	T	N	M	S	L	C	A	A	I	S	T	S	...	T	Y	K	N	T	359										
HPV31	TY	F	P	T	P	S	G	S	M	V	T	S	D	A	Q	I	F	N	KPY	W	M	Q	R	A	Q	G	H	N	N	G	I	C	W	G	N	L	F	V	T	V	D	T	R	S	T	N	M	S	V	C	A	A	I	A	N	S	D	...	T	T	F	K	S	S	N	360						
HPV33	AFF	P	T	P	S	G	S	M	V	T	S	E	S	Q	L	F	N	KPY	W	L	Q	R	A	Q	G	H	N	N	G	I	C	W	G	N	Q	V	F	V	T	V	D	T	R	S	T	N	M	L	C	T	Q	V	T	S	D	S	...	T	Y	K	N	E	N	358								
HPV32	F	K	E	Y	L	R	H	A	E	E	Y	D	I	Q	I	F	I	F	L	C	K	I	T	L	S	V	E	V	M	S	I	H	T	M	N	P	D	I	L	D	D	W	N	V	G	V	A	P	P	P	S	G	T	L	E	S	Y	R	F	V	Q	S	A	I	R	C	Q	A	K	430		
HPV11R	Y	K	E	Y	M	R	H	V	E	E	F	D	L	Q	I	F	I	F	L	C	S	I	T	L	S	A	E	V	M	A	I	H	T	M	N	P	S	V	L	E	D	W	N	F	G	L	S	P	P	P	N	G	T	L	E	D	T	Y	R	V	Q	S	A	I	T	C	Q	K	P	427		
HPV6bR	Y	K	E	Y	M	R	H	V	E	E	Y	D	L	Q	I	F	I	F	L	C	S	I	T	L	S	A	E	V	M	A	I	H	T	M	N	P	S	V	L	E	D	W	N	F	G	L	S	P	P	P	N	G	T	L	E	D	T	Y	R	V	Q	S	A	I	T	C	Q	K	P	426		
HPV18R	F	K	Q	S	R	H	V	E	E	Y	D	L	Q	I	F	I	F	L	C	T	I	T	L	A	D	V	M	S	I	H	S	M	N	S	I	L	E	D	W	N	F	G	V	P	P	P	T	T	S	L	V	D	T	Y	R	F	V	Q	S	A	I	T	C	Q	K	D	432					
HPV16R	F	K	E	Y	L	R	H	G	E	E	Y	D	L	Q	I	F	I	F	L	C	K	I	T	L	A	D	V	M	T	Y	I	H	S	M	N	S	I	L	E	D	W	N	F	G	L	P	P	P	G	T	L	E	D	T	Y	R	F	V	T	S	Q	A	I	C	Q	K	H	431				
HPV31	F	K	E	Y	L	R	H	G	E	E	F	D	L	Q	I	F	I	F	L	C	K	I	T	L	A	D	V	M	T	Y	I	H	S	M	N	P	A	I	L	E	D	W	N	F	G	L	T	P	P	S	G	S	L	E	D	T	Y	R	F	V	T	S	Q	A	I	T	C	Q	K	T	432	
HPV33	F	K	E	Y	I	R	H	V	E	E	Y	D	L	Q	I	F	I	F	L	C	K	V	L	T	L	A	E	V	M	T	Y	I	H	A	M	N	P	D	I	L	E	D	W	Q	F	L	T	P	P	P	S	A	S	L	Q	D	T	Y	R	F	V	T	S	Q	A	I	T	C	Q	K	T	430
HPV32	V	..T	A	P	E	K	K	D	P	F	S	D	Y	S	F	W	E	V	N	L	S	E	K	F	S	S	D	L	D	Q	F	P	L	G	R	K	F	L	L	Q	A	G	L	R	A	R																										

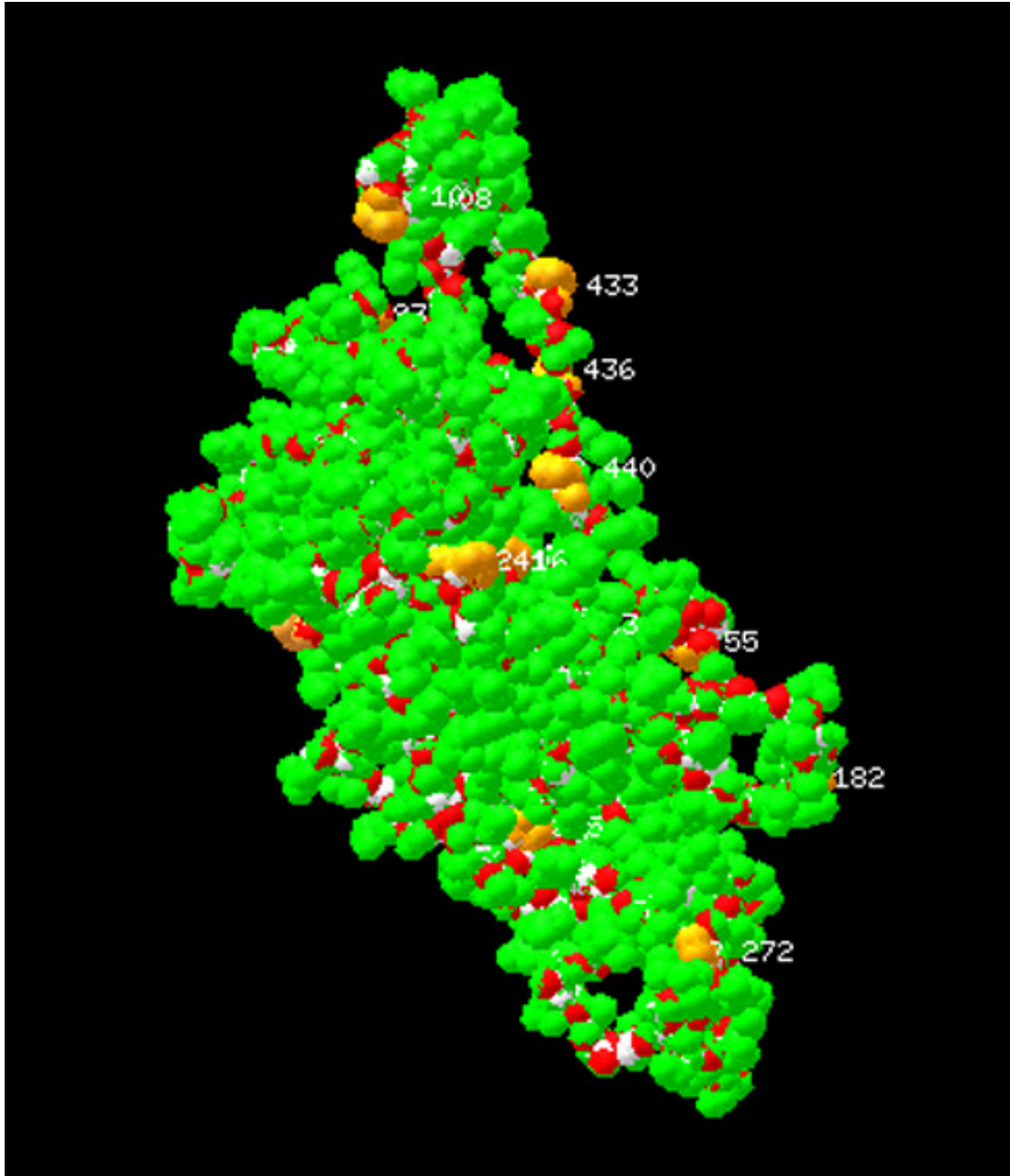


Figure 3. A model of the HPV16-L1 capsid protein responsible for generating a number of forms of squamous cell carcinomas. In the present picture green represents the surface groups and the backbone has been shown in red and white. The suitable prolines for ginger enzyme cleavage are shown in orange-yellow. The numbers represents the proline number in the HPV16-L1 protein.

Papilloma virus is of particular interest to Biohawk because it is associated with disease in humans and other species, for example equine sarcoid, bovine eye cancer, common warts, corns, some squamous cell carcinomas, and cervical cancer. Case studies have shown a very favourable response to treatment with the Biohawk creams (for example, Skin Rejuvenator). For example, the following common warts (Figure 4) were treated with a ginger cream once at 3 pm on one day and the next morning the growths had fallen off the hand.



Figure 4: Common wart treated with an active ginger cream

Bacteria

Some bacteria have a carbohydrate coating on the membrane mostly attached to the membrane proteins through the amino acids, serine (highlighted pink) and threonine (highlighted grey). This coating makes it more difficult for antibiotics to penetrate the membrane and makes it more difficult for the ginger enzyme to digest the proline-rich proteins. Biohawk's Pine Crush was developed to remove the carbohydrate coating. In other types of bacteria the proline-rich proteins project out of the membrane and although they may have serine and threonine, the carbohydrate coating does not prevent the ginger enzyme from digesting the proline-rich membrane proteins. Some examples of membrane proteins for common types of bacteria are given below.

Whooping Cough: *Bordetella pertussis*

***Bordetella pertussis*, the causative agent of whooping cough**, is an aerobic coccobacillus capsulate of the genus *Bordetella*. The acellular pertussis vaccine components, pertussis toxoid (PT), pertussis filamentous haemagglutinin (FHA) and pertactin (PRN) are extracted from phase I *Bordetella pertussis*, and are then purified and stabilised. The structures of FHA and PRN are given below.

FHA protein [*Bordetella pertussis*]

1 MNTNLYRLVF SHVRGMLV PV SEHCTVGN TF CGRTRGQARS GARATSL SVA PNALAWALML
61 ACTGLPLVTH AQGLV PQQGT QVLQGGNKVP VVNIADPNSG GVS HNKFFQQF NVANP GVVFN
121 NGLTDGV SRI GGALTKNP NL TRQASAILAE VTD TSPSRLA GTLEVYKGA DLIANPNGI
181 SVNGLSTLNA SNLTLTTGRP SVNGGRIGLD VQQGTVTIER GGVNATGLGY FDVVARLVKL
241 QGAVSSKQK PLADIADVAG ANRYDHATRR ATPIAAGARG AAAGAY AIDG TAAGAMYGKH
301 ITLVSSD SGL GVRQLGSLSSPSAITVSSQG EIALGDATVQ RGPLSLKGAG VVSAGKLAG
361 GGAVNVAGGG AVKIASASSV GNLAVQGGGK VQATLLNAGG TLLVSGRQAV QLGAASSRQA
421 LSVNAGGALK ADKLSATRRV DVDGKQAV AL GSSSNALSV RAGGALKAGKLSATGRLDVD
481 GKQAVTLG SVASD GALSVSA GGNLRANELV SSAQLEVRGQ REVALDDASS ARGMTVVAAG
541 ALAARNLQSK GAIGVQGGEA VSVANANS DA ELRVRGRGQV DLHDL SAARG ADISGEGRVN
601 IGRARSSDV KVS AHGALS I D SMTALGAIG VQAGS VSAK DMRSRGAVTV SGGGAVNLGD
661 VQSDGQVRAT SAGAMTVRDV AAAADLALQA GDALQAGFLK SAGAMTVNGR DAVRLDGAHA
721 GGQLRVSSDG QAALGSLAAK GELTVS AARA ATVAELKSLD NISVTGGERV SVQSVNSASR
781 VAISAHGALD VGKVS AKSGI GLEGWGAVGA DSLGSDGAI S VSGRDAVRVD HARSLADISL
841 GAEGGATLGA VEAAGSIDVR GGSTVAANSL HANRDVRVSG KDAVRVTAAT SGGGLHVSSG
901 RQLDLGAVQA RGALALDGGG GVALQS AKAS GTLHVQGGEH LDLGTAAVG AVDVNGTGDV
961 RVAKLVSDAG ADLQAGRSMT LGIVDTTGD LQARAQQKLEL GSVKSDGGLQ AAAGGALS LA
1021 AAEVAGALEL SGQGVTVDRASASRRARIDST GSVGIGALKA GAVEAASPRRARRALRQDFF
1081 TPGSVVVRAQ GNVTVGRGDP HQGVLAQGD IIMDAKGGTLL LRNDALTENG TVTISADSAV
1141 LEHSTIESKI S QSVLAAKGD KGKPAVSVKV AKKLFLNGTL RAVNDNNETM SGRQIDVVDG
1201 RPQITDAVTG EARKDES VVS DAALVADGGP IVVEAGELVS HAGGIGNGRN KENGASVTVR
1261 TTGNLVNKGYSAGKQGVLE VGGALTNEFL VGS DGTQRIE AQRIENRGTF QSQAPAGTAG
1321 ALVVKAAEAI VHDGVMATKG EMQIAGKGGG SPTVTAGAKA TTSANKLSVD VASWDNAGSL
1381 DIKKGAQVT VAGRYAEHGE VSIQGDYTVS ADAIALAAQV TQRGGAANLT SRHDTRFSNK
1441 IRLMGPLQVN AGGPV SNTGN LKVREGVTVT AASFDNETGA EVMAKSATLT TSGAARNAGK
1501 MQVKEAATIV AASVSNP GTF TAGKDITVTS RGGFDNEGKM ESNKDIVIKT EQFSNGRVLD
1561 AKHDLTVTAS GQADNRGSLK AGHDFTVQAQ RIDNSGTMAA GHDA TLKAPH LRNTGQVVAG
1621 HDIHIINS AK LENTGRVDAR NDIALDVADF TNTGSLYAEH DATLTLAAGT QRDLVVDQDH
1681 ILPVAEGLTR VKAKSLTTEI ETGNPGLIA EVQENIDNKQ AIVVGKDLTL SSAHGNVANE
1741 ANALLWAAGE LTVKAQNITN KRAALIEAGG NARLTA AVAL LNKLGRI RAG EDMHLDAPRI
1801 ENTAKLSGEV QRKGVQDVGG GEHGRWSGIGYVNYWLRAGNGK KAGTIAAPWYGGDLTAEQ
1861 SLIEVGKDLH LNAGARKDEH RHLLNEGVIQ AGGHGIGGD VDNRSVVRTV SAMEYFKT PL
1921 PVS LTALDNR AGLSPATWNF QSTYELLDYL LDQNRYEYIW GLYPTYTEWS VNTLKNLDLG
1981 YQAKPAPTAP PMPKAP ELDL RGHTLES AEG RKIFGEYKKL QGEYEKAKMA VQLVEAYGEA
2041 TRRVHDQLGQ RYKALGGMD AETKEVDGII QEFAADLRTV YAKQADQATI DAETDKVAQR
2101 YK SQIDAVRL QAIQ PGRVTL AKALS AALGA DWRALGHS QL MQRWKDFKAG KRGAEIAFY P
2161 KEQTVLAAGA GLTLSNGAIH NGENAAQNRG RPEGLKIGAH SATSVSGSFD ALRDVGLEKR
2221 LDIDDALAAV LVNPHIFTRI GAAQTS LADG AAGPALARQA RQAPETDGMV DARGLSADA

2281 LASSLALDAA QGLEVS GRRN AQVADAGLAG PS AVAAPAVG AADVGV EPVT GDQVDQPVVA
 2341 VGLEQPVATV RVAPPAVAL P RPLFETRIKF IDQSKFYGSR YFFEQIGYKP DRAARVAGDN
 2401 YFDTTLVREQ VRRALGGYESS RL PVRGVALV AKLMD SAGTV GKALGLKVG V APTAQLKQA
 2461 DRDFVWYVDT VIDGQKVLAP RLYLTEATRQ GITDQYAGGG ALIASGGDVT VNTDGHVSS
 2521 VNGLIQGRSV KVDAGKGVV VADSKGAGGG IEADDEVDS GRDIGIEGGK LRGKDVRLKA
 2581 DTVKVATSMR YDDKGRLAAR GDGALDAQGG QLHIEAKRLE TAGATLKGGK VKLDVDDVKL
 2641 GGVYEAGSSY ENKSSSTPLGS LFAIL SSTTE TNQSAHANHY GTRIEAGTLE GKMQLNLEIEG
 2701 GSVDAHTDL SVARDARFKA AADFAHAEHE KDVRQLSLGA KVGAGGYEAG FSLGSESGLE
 2761 AHAGRGMTAG AEVKVGYRAS HEQSSSETEKS YRNANLNFGG GSV EAGNVLD IGGADINRNR
 2821 YGGAACKGNAG TEEALRMRAK KVESTKYVSE QTSSSSGWSV EVA STASARS SLLTAATRLG
 2881 DSVAQNVEDG REIRGELMAA QVAEATQLV TADTAVALS AGISADFDSS HSRSTS QNTQ
 2941 YLGGNLSIEA TEGDATLVGA KFGGGDQVSL KAAKSVNLMA AESTFESYSE SHNFHASADA
 3001 NLGANVQGA VGLGLTAGMG TSHQITNETG KTYAGT SVDA ANV SIDAGKD LNL SGRVRG
 3061 KHVLDVEGD INATSKQDER NYNSSGGGDW ASAGVAIQNR TLVAPVGSAG FNFNTEHDNS
 3121 RL TNDGAAGV VASDGLTGHV KGDANLTGAT IADLSGKGNL KVDGAVNAQN LKDYRDKDGG
 3181 SGGLNVGSS TTLA PTVGVA FGRVAGEDYQ AEQRATIDVG QTKD PARLQV GGGVKGTLNQ
 3241 DAAQATVVQR NKHWAGGSE FSVAGKSLKK KNQV RVPVET P TPDVVDGPPSRPTT PPASPO
 3301 PIRATVEVSS PPPV SVATVE VVPRPKVETG SAASASAGGA QVVPVTPPKV EVAKVEVVPR
 3361 PKVETAQPLP PRPVVAEKVT TPAVQPQLAK VETVQPVKPE TTKPLPKPLP VAKVTKAPPP
 3421 VVETAQPLPP VKPQKAT PGP VAEVGKATVT TVQVQS APPK PAVAKQPAP APKPKPKPKP
 3481 KAERPKPKGKT TPLSGRHVVQ QQVQVLQRQA SDINNTKSLP GGKLPKPVTV KLTDENGKPK
 3541 TYTINRRDL MKLNGKVLST KTTLGLEQTF RLRVEDIGGK NYRVFYETNK

pertactin outer membrane protein [Bordetella pertussis]

1 MNMSLSRIVK AAPLRRTTLA MALGALGAAP AAHADWNNQS IVKTGERQHG IHIQSSDPGG
 61 VRTASGTTIK VSGRQAQGIL LENPAELQF RNGSVTSSGQ LDDGIRRFLL GTVTVKAGKL
 121 VADHATLANV GDTWDDDDGIA LYVAGEQAQA SIADSTLQGA GGVQIERGAN VTVQRS AIVD
 181 GGLHIGALQS LQPELPPSR VVLRDTNVTAVP ASGAPAAV SVLGASELTL DGGHITGGRA
 241 AGVAAMQGAV VHLQRATIRR GDAPAGGAV PGGAV PGGAV PGGFGPGGFGP VLDGWYGVVD
 301 SSSVELAQS IVEAPELGAA IRVGRGARVT VSGGSLSPH GNVIETGGAR RFA PQAAPLS
 361 ITLQAGAHAQ GKALLYRVL PEPVKLTLTGG ADAQGDIVAT ELPSIPGTSI GPLDVALASQ
 421 ARWTGATRAV DLSLIDNATW VMTDNSVGA LRLASDGSVD FQPPAEAGRF KVLTVNTLAG
 481 SGLFRMNVA DLGLSDKLVV MQDASGQHRL WVRN SGEPA SANTLLLVT PLGSAATFTL
 541 ANKDGKVDIG TYRYRLAANG NGQWSLVGAK APPAPKPAQ PGPQPPQPQ POPEAPAPQP
 601 PANGRELSAAA NAAVNTGGVG LASTLWYAES NALS KRLGEL RLNPDAGGAW GRGFAQRQQL
 661 DNRAGRFRDQ KVAGFELGAD HAVAVAGGRWHLGGLAGYTRGDRGFTGDGGGHTD SVHVGG
 721 YATYIADSGF YLDATLRASR LENDFKVAGS DGYAVKGYR THGVGASLEA GRRFTHADGW
 781 FLEPQAELAV FRAGGAYRA ANGLRVRDEG GSSVLGRLGL EVGKRIELAG GRQVQPYIKA
 841 SVLQEFDGAG TVHTNGIAHR TELRGTRAEL GLGMAAALGR GHSLYASYEY SKGPKLAMPW
 901 TFHAGYRYSW

Homology in sequences of above proteins

Pertactin: PKPAP- - - QPGPQPPQPPQPEAPAPQP
 PKPAP- - - QP-P- P- - - P-P-P- A - -P- P
 FHA protein: PKPAPVAKQPAPAPKPKPK PKPKAERPKP

Staphylococcus aureus

Cap5P [Staphylococcus aureus]

1 MCLNFREDNV MKKIMVIFGT RPEAIKMAPL VKEIDHNGNF EANIVITAHQ RDMLD SVLSI
 61 FDIQADHDLN IMQDQQTLAG LTANALAKLD SIINEEQPDM ILVHGDTTTT FVGS LAAFYH
 121 QIPVGHVEAG LRTHQKYS PF PEELNRVMVS NIAELNFAPT VIAAKNLLFE NKDKERIFIT
 181 GNTVIDALST TVQNDFVSTI INKHKGKVI LLTAHRRENI GEPMHQIFKA VRDLADEYKD
 241 VVFIYPMHRN PKVRAIAEKY LSGRNRIELI EPLDAIEFHN FTNQSYLVLT DSGGIQEEAP
 301 TFGKPVVLVR NHTERPEGVE AGTSRVIQTD YDNIVRVKQ LIEDEAYQR MSQANNPYGD
 361 GQASRRICEA IEYFGLRTD KPDEFVPLRH K

Protein A signal fusion protein.

1 MKKKNIYSIR KLGVGIA SVT LGTLLISGGV TPAANAAQHD EAVDNKFNKE QQNAFYEILH
 61 LPNLNEEQRN AFIQSLKDDP SSSANLLAEA KKLNDAAQAPK VDNKFNKEQQ NAFYEILHLP

121 NLNEEQRNAF IQSLKDDPSQ SANLLAEAKK LNDAAQPKVD ANSSSVPGDP LEStCRHASL
181 ALAVVLQRRD WENP GVTQLN RLAAHPPFAS WRNS EEAR TD RPSQQLRSLNGEWRFRNGW
241 R

Biofilm-associated surface protein [Staphylococcus aureus]

1 MGNKQGF LN KLNKY SIRKF TVGTASLLVGG TLLFFGIGSE AQA AELDTIT KEDVK SQDKG
61 EALDIKNIKE SEKDVTTEDDNNAEVQNSAQTVDKSEN SNTAVE STND SV KTDETKETS E
121 NKSAQDDDN IKEDSNTQEE STNTSSQSSSEVPQTKKDTNETS ETAIDEDA STKEQNNKDND
181 TAQDDDN IKE D SNTQEE STN TSSQSSSEVPQ TKKEQPKSS NSIKEPDKQQ EEVAKKEEKAI
241 TEIADKNKEL ELKNNKTDKN EESSELESNLS SSENKKTVE SFLNSQLSDS ETKKIMENAN
301 IDYDKATDEE INTEILRASL IEMANNKKKT ETLATPQRTM FRAMATPTAL RAAVNQDEEL
361 QKSLGYTDNY TFASMLFDPG KLDSDDALNS NIIPFDLHSY MSGANS GNRY KIDLKLDPII
421 AEHVTKISANPSGSKNKP VEF VRNKDENG NL TDTWEVNFIR ANDGLFGGAE ILSQYTAKNG
481 KIELDDTVGN IISNAGNLSN NKLNHQVFVR DSRENKIVRT SSSGYFLTQ ADDDLVNLEN
541 NVSTENNNAF KASSGSATYN ENVGEFGGIL IDQIMKNGI FSYSKTKANQ WAYNYQIDKD
601 LLYPIEGVEL HQDYKGLNG FDKNYDAKNK VADLTIDEVG NGTITS DNLN KLIEFNAL P
661 ETVGVRVVLK LNKSVNNILT KDAKYDSEGN LIRETTKQKE DFTFAGYLT DSKGALINNTL
721 GTSTLALQDY DKDGLLD RYE RQLSLSDAEN EDTDGDGKND GDEVVNYKTS PLVGKPAAD
781 ITTEDTVVSG SVPLKEGAAT QTAKVINAEG TTVGTATVNS DGTFSVSI PN SPEGTYTIAI
841 DSPNYDNDEV NTFEIVDN SKLPAP SINPVD DNDQQIVVNG TSGSTVTVTD SNNNVLGTVT
901 IPADDTSAAI NVDTPLEAGT VLTSTASKDG KTS DVSDQIT VTDATA PDAP TLDEVNTDAT
961 QVTGQAE P NS TVKLTFPDGT TATGTADDQG NYTIDI PSNV DLNGGEELQV TATDKDGNTS
1021 EPSSANVTD TAPDAP TVND VTS DATQVTG QAE PNSTVKLTF PDGTTATGTADDQGN YTI
1081 DIPSNVDLNG GEELQVTATD KDGNTSE P SS ANVDTTAPD APTVNDVTS DATQVTGQAE P
1141 NSTVKLTF PD GTTATGTADD QGNYTIDI PS NVDLNGGEEL QVTATDKDGNTSE P SS ANVT
1201 DTTAPDAP TV NDVTS DATQV TGQAE PNSTV KLTFPDGT TATGTADDQGN YTIDI PS NVDL
1261 NGGEELQVTA TDKDGNTSE P SS ANVDTTAPD APTVNDVTS DATQVTGQAE PNSTVKLTF
1321 PDGTTATGTA DDQGN YTI DI PS NVDLNGGE ELQVTATDKD GNTSE P SS ANVDTTAPDAP
1381 TVNDVTS DAT QVTGQAE P NS TVKLTFPDGT TATGTADDQG NYTIDI PS NVDLNGGEELQV
1441 TATDKDGNTS EPSSANVTD TAPDAP TVND VTS DATQVTG QAE PNSTVKLTF PDGTTATG
1501 TADDQGN YTI DI PS NVDLNG GEELQVTATD KDGNTSE P SS ANVDTTAPD APTVNDVTS D
1561 ATQVTGQAE P NSTVKLTF PD GTTATGTADD QGNYTIDI PS NGDLNGGEELQVTATDKDGN
1621 TSE P SS ANVT DTTASDAP TV NDVTS DATQV TGQAE PNSTV KLTFPDGT TATGTADDQGN Y
1681 TIDI PS NVDL NGGEELQVTA TDKDGNTSE P SS ANVDTTAPD APTVNDVTS DATQVTGQA
1741 EPNSTVKLTF PDGTTATGTA DDQGN YTI DI PS NVDLNGGE ELQVTATDKDGNTSE P KLTN
1801 VDTTASDAP TVNDVTS DAS QVTGQAE P NS TVKLTFPDGT TATGTADDQGN YTI DI PS NV
1861 DLNGGEELQV TATDKDGNTS EPSSANVTD TAPDAP TVND VTS DATQVTGQAE PNSTVKL
1921 TFPDGTTATG TADDQGN YTI DI PS NGDLNG GEELQVTATD KDGNTSE P SS ANVDTTAPD
1981 APTVNDVTS D ATQVTGQAE P NSTVKLTF PD GTTATGTADD QGNYTIDI PS NVDLNGGEEL
2041 QVTATDKDGN TSE P SS ANVT DTTAPDAP TV NDVTS DATQV TGQAE PNSTVKLTF PDGTTA
2101 TGTADDQGN Y TIDI PS NVDL NGGEELQVTA TDKDGNTSES T N TTIID SDDN SDNGNNSGA
2161 GDTSSDDNS GNGDN S GAGD NSDSDDN SDNGNNS GAGDN SSSDDN SDNEDN SSSNKD SIN
2221 QDSNVNSDS KHDKQNEL PE TGEKEVRNGT LFGTFLFAGLG SLLLFTKRRR KENDKK

Streptococcus agalactiae

Protein immunoglobulin-a-binding beta antigen

1 AIKQQ IFDI DNAK EVEID NLVHDAF SKM NA VAKFQKG LE N PE PD PKIPELPQA
61 PD PQAPD PHVPE SPK APE APRVPE SPK PD PHVPE SPK APE APRVPE SPK PD PHV
121 PE SPK APE APRVPE SPK PD PHVPE SPK APE APRVPE SPK PD PHVPE SPK APE APRV
181 PE SPK APE APRVPE SPK PE APKIPEPPK PDV PKLPDV P KL PDV PKLPD APKLPDGLNK
241 VGQAVF S D GN KV VVFD KP DADKLHL KEV KELA

Penicillin binding protein 1a

1 MITIKKE SVI KLLKYAFGII MGFIILAIVI GLLFAYYV S RSPKLT DQAL KSVN SSLVYD
61 GNNKLIADLG SEKRES VSAD SIPNLVNAI TSIEDKRFFK HRGVDIYRIL GAAWHNLVSS
121 N T QGG S LDQ QLIKLAYF ST NKS DQ TLKRK SQEVWLALQM ERKY TKEEIL TFYINKVYMG
181 NGNYGMRTTA KSYFGKDLKE LSIAQLALLA GI PQAPTQYD PYKNPESAQT RRNTVLQQMY
241 QDKNISKKEY DQAVAT PVT D GLKELKQKST YPKYMDNYLK QVIS EVKQKT GKDIF AGLK
301 VYTNIN DAQ KQLYDIYNSD TYIAYPNNEL QIASTIMDAT NGKVIAQLGG RHQENISFG
361 TNQSVL TDRD WGSTMKPISA YAPAISS G VY NSTGQSLNDS VYYWP GSTQ LYDWD RQYMG

421 WMSMQTAIQQSRNVPVAVRAL EAAGLDEAKS FLEKLGIIYP EMNYSNAISS NNSSSDAKYG
481 ASSSEKMAAAYSAFANGGTYYKPOYVVKIEFSDGTNDTYAASGSRAMKETIAYMMTDMLKT
541 VLTFTGTGTA AIPGVAQAGK TGTSNYTEDE LAKIEATIGI YNSAVGTMAPDENFVGYTSK
601 YTMAIWGYK NRLTPLYGSSQLDIAT EYVYRAMMSYLTGGYSADWMPPEGLYRSSSYLYING
661 TTTTGTYS SSSVYKNIYQNSG QSSQSSSSSSEKQKEDKNT ANDANSSSPQVETPNNGNAT
721 TPNNSNQVP GTHGNGNGN NNVPNGN

Pi-2a ancillary protein 2

1 MKKIRKSLGL LCCFLGLVQ LAFFSVA SVN ADTPNQLTIT QIGLQPNTE EGISYRLWTV
61 TDNLKVDLLS QMTDSELNQK YKSILTSPTD TNGQTKIALP NGSYFGRAYK ADQSVSTIVP
121 FYIELPDDKL SNQLQINPKR KVETGRRLKI KYTKEGKIKK RLSGVIFVLY DNQNQPVRFK
181 NGRFTTDDQD ITS LVTDDKG EIEVEGLLPG KYIFREVKAL TGYRISMKDA VVAVVANKTQ
241 EVEVENEKET PPPTNPKPSQ PLFPQSFLPK TGMII GGGLT ILGCILGIL FIFLRKTKNS
301 KSERNDTV

Streptococcus pneumoniae

PspA [Streptococcus pneumoniae]

1 MNKKMILTSLASVAILGAG FVASSPTFVR AEEAPVANQS KAEKDYDAAV KKS EAAKKDY
61 ETAKKAEDA QKKYDEDQKK TEAKAEKERK ASSEKIAEATK EVQQAYLAYL QASNE SQRKE
121 ADKKIKEATQ RKDEAEAAFA TIRTTIVVPEPSELAEKTKK AEEATKEAEV AKKSEEA
181 EVEVEKNKIL EQDAENEKKI DVLQNKVADL EKGIA PYQNE VAELNKEIAR LQSDLKDAEE
241 NNVEDYIKEG LEQAITNKA ELATTQQNID KTQKDLEDAE LELEKVLATL DPEGKTQDEL
301 DKEAAEAELN EKVEALQNQV AELEEELSKL EDNLKDAETN NVEDYIKEGL EEAIATKKA
361 LEKTQKELDA ALNELGPDGD EEETPAPAPQPEKPAEEPEN PAPAPEKPSADQQA EEDYA
421 RRS EEEYNRLTQQPPKAEKPAPAPQPEQPAPAPKIGWKQENGMWYFYNTDGS MATGWLQ
481 NNGSWYYLNSGAMATGWLQYNGSWYYLNANGAMATGWLQYNGSWYYLNANGAMATGWLQ
541 YNGSWYYLNANGMATGWLQYNGSWYYLNANGMATGWAKVHGSWYYLNANGMATGWVK
601 DGETWYYLEA SSMKANQWF QVSDKWYYVN GLGSLSVNTT VDG YKVNANG EWV

cbpA [Streptococcus pneumoniae GA19998] gram pos signal peptide

1 MFASKSERKV HYSIRKFSIG VASVVVALF LGGVVHAEV RRGNNLTVTS SSGDEVE SHYQ
61 SILEKVRKSL EKDRHTQNV D LIKKLQDIK TYLYNLKEKPEAELT SKTKK ELDAAFEKFK
121 KEPELTKKLA EAEKKAKDQK EEDHRNYPTN TYKTIELEIA EAEVGVAKAE LELVQAQVQI
181 PQDTEKINAA KAKVEAAKSN VKKLEKIKSD IEKTYLYKLD NSTKETPKSR VRRNSPQVGD
241 SRELKETIDK AKETLSTYMV TRLTKLDP SV FWFADLLMDA KKVVEEYKTK LEDASDKKSV
301 EDLRKEAEGK IESLIVTHQN REKENQPAPQ PGGQAGGSMV VPPVTQT PPS TSQSPGQKAT
361 EAEKKKLQDL IRQFQEALNK LDDETKTVPD GAKLTGEAGK AYNETRTRYAK EVVDKSKLL
421 SQTAVTMDL AMQLTKLND MSKLKEAKAK LVPEVKPQPE NPEPKPQPEG EKPSV PDINQ
481 EKEKAKLAIA TYMSKILDDI KKHHLKKEKH HQIVALIKDL DKLKKQALSE IDNVNTKVEI
541 ENTVHKVFAD MDTVVTKFQK GLIQNTPQVPEAPKSPPEVPK VSDTPKAPDT PQVPEAPKSP
601 EVPKVPEAPK APDTPQVPEA PKSPEVPKVDTPKAPDTPQ VPEAPKAPDT PQIPEAPE
661 TPAPAPEAPK TGWKQENGMWYFYNTDGS MATGWLEYNGSWYYLNANGAMATGWLEYNGSW
721 YYLNTNGAMETGWLEYNGSWYYLNTNGAMETGWLEYNGSWYYLNTNGAME TGWLEYNGSW
781 YYLNTNGAMETGWLEYNGSWYYLNTNGAMETGWLEYNGSWYYLNANGSMA TGWLKDGDTW
841 YYLEASGAMK ESQWFKVSDK WYYVNGS GAL AVNTTVGGYR VNANGKWVN

Pseudomonas aeruginosa

ExoU [Pseudomonas aeruginosa PA103].

1 MHIQSLGATA SSLNQEPVET PSQAAHK SAS LRQEPSGQGL GVALKSTPGI LSGKLPESVS
61 DVRFSSPQGG GESRRLTDSA GPRQITLRQF ENGVTELQLS RPPLTSLVLS GGGAKGAAYP
121 GAMLALEEKG MLDGIRSMSG SSAGGITAAL LASGMSPAAF KTLSDKMDLI SLLDSSNKKL
181 KLFQHISS EI GASLKKGLGN KIGGFS ELLL NVLPRIDSRA EPLERLLRDE TRKAVLGQIA
241 THPEVARQPT VAAIASRLQS GSGVTFGDL RLSAYIPQIK TLNITGTAMF EGRPQLVVFN
301 ASHTPDLEVA QAAHISGSFP GVFKVSLSD QPYQAGVEWT EFQDGGVMIN VVPEMIDKN
361 FDSGPLRRND NLILEFEGEA GEVAPDRGTR GGALKGWVVG VPALQAREML QLEGLEELRE
421 QTVVVPLKSE RGD FSGMLGG TLNFTMPDEI KAHLQERLQE RVGEHLEKRL QASERHTFAS
481 LDEALLALDD SMLTSSVAQQN PEITDGAVAF RQKARDAFTE LTVAVISANG LAGRLKLDEA

541 MRSALQRLDA LADTPERLAW LAELNHADN VDHQQLLDAM RGQTVQSPVL AAALAEQRR
601 KVAVIAENIR KEVIFPSLYR PGQPD SNVAL LRRAEELRH ATSPAENQA LNDIVDNYSA
661 RGFLRF GKPL SSTTVEMAKA WRNKEFT

exotoxin A, partial [*Pseudomonas aeruginosa*].

1 ALLERNYPTG AEFLGDGGDV SFSTRGTQNW TVERLLQAHR QLEERGYV FV GYHGTFLEAA
61 QSIVFGG VRA RSQDLDAIWR GFYIAGDPAL AYG YAQDQEP DARGRIRIGA LLRVYV PRSS
121 L PGFYRTGLT LAAP EAAGEV ERLIGHPLPL RLDAITG PEE EGG RLETILG WPLAERTVVI
181 PSAIPTDPRN VGGDLDPSSI PDKEQAISAL PDYASQPGKPPREDLK

toxA gene product [*Pseudomonas aeruginosa* PAO1]

1 MHLT PHWIPL VASLGLLAGG SFASAAEEAF DLWNECAKAC VLDLKDGVRS SRMSVDP AIA
61 DTNGQGVLHY SMVLEGGNDA LKLAIDNALS ITS DGLTIRL EGGV EPNKP V RYSYTRQARG
121 SWSLNWLVP I GHEKPSNIKV FIHELNAGNQ LSHMSPIYTI EMGDPELLAKL ARDATFFVRA
181 HESNEMQPTL AISHAGVSVV MAQAQPRREK RWS EWASGKV LCLLDPLDGV YNYLAQQRCN
241 LDDTWEGKIY RVLAGNPAKH DLDIKPTVIS HRLHF PEGGS LAALTAHQAC HLPLETFFRH
301 RQPRGW EQLE QCGYPVQRLV ALYLAARLSW NQVDQVIRNA LSPGSGGDL GEAIREQPEQ
361 ARLALT LAAA ESERFVRQGT GNDEAGAASA DVVSLTCPVA AGE CAGPADS GDALLERNYP
421 TGAEFLGDGG DISFSTRGTQ NWTVERLLQA HRQLEERGYV FVGYHGTFLE AAQSIVFGGV
481 RARSQDLDAI WRGFYIAGDP ALAYG YAQDQ EPDARGRIRN GALLRVYVPR SSLPGFYRTG
541 LTLAAPEAAG EVERLIGHPL PLRLDAITGPEEEGG RLETI LGWPLAERTV VIPSAIPTDP
601 RNVGGDL DPS SIPDKEQAIS ALPDYASQPG KPPREDLK

Escherichia coli

Surface protein [*E. coli*]

1 MTT PNP LAKT KGAGTTFW MY TGKGD AFANP LS DTDWLRLA MVKDL QP GEM TADAEDDTYL
61 DDEDADWKTT TGGQK SVGDT SATLAW RPD SGQKKLVQLF DS GEVCAFRI KYPNGTVDVF
121 RGWLS LGKT IAS KDVMTRT VKISGVGRPY LAEEGXETVG VTGLTVAPAS ASVKAGATTT
181 LTFTV KPDGA SDKAIS VHSS DPQTASVTL S GLVATVKGVK QG SVSIVGMT SDGEFVAVAA
241 VTVSAP

Plasmodium falciparum - malaria

PfMSA2

1 SIRR SMAESK SPTGTGASGS AGSGDGASGS AGSGDGASGS AGSGDGAVAS ARNGANP GAD
61 AEGSSSTPAT TTTTTTTTTT TTTND AEAST STSS ENPNHN NAET NPKGNG EVQEPNQANK
121 ETQNN SNVQQ DSQTK SNVPP TQDADTKSPT AQPEQAENSA PTAEQTESPE LQSA PEN

P proline **P** where ginger enzymes digest protein

T threonine that can bind to carbohydrate

S serine that can bind to carbohydrate