Sub classification and targeted characterization of prophage-encoded two-component cell lysis cassette

K V Srividhya and S Krishnaswamy*

Centre of Excellence in Bioinformatics, School of Biotechnology, Madurai Kamaraj University, Madurai 625 021, India

**Corresponding author (Fax, 91-425-2459105; Email, krishna@mrna.tn.nic.in)*

Bacteriophage induced lysis of host bacterial cell is mediated by a two component cell lysis cassette comprised of holin and lysozyme. Prophages are integrated forms of bacteriophages in bacterial genomes providing a repertoire for bacterial evolution. Analysis using the prophage database (*http://bicmku.in:8082*) constructed by us showed 47 prophages were associated with putative two component cell lysis genes. These proteins cluster into four different subgroups. In this process, a putative holin (*essd)* and endolysin (*ybcS)*, encoded by the defective lambdoid prophage DLP12 was found to be similar to two component cell lysis genes in functional bacteriophages like p21 and P1. The holin *essd* was found to have a characteristic dual start motif with two transmembrane regions and C-terminal charged residues as in class II holins. Expression of a fusion construct of *essd* in *Escherichia coli* showed slow growth. However, under appropriate conditions, this protein could be over expressed and purified for structure function studies. The second component of the cell lysis cassette, *ybcS*, was found to have an N-terminal SAR (Signal Arrest Release) transmembrane domain. The construct of *ybcS* has been over expressed in *E. coli* and the purified protein was functional, exhibiting lytic activity against *E. coli* and *Salmonella typhi* cell wall substrate. Such targeted sequence-structure-function characterization of proteins encoded by cryptic prophages will help understand the contribution of prophage proteins to bacterial evolution.

[Srividhya K V and Krishnaswamy S 2007 Sub classification and targeted characterization of prophage encoded two-component cell lysis cassette; *J. Biosci.* **32** 979–990]

1. Introduction

Bacterial genomes are composed of a core minimal species genomic backbone decorated with variety of genetic elements including prophages, conjugative transposons and insertion sequence elements. Prophages are distinct from other genomic segments as they are mostly defective or cryptic that has been proposed to be acquired by horizontal gene transfer events (Casjens 2003). The impact of prophages on bacterial evolution has been reviewed extensively (Brussow *et al* 2004). Virulence factors in many pathogenic bacteria are observed to be located on prophage loci, indicating the possible role played by prophages in conferring pathogenicity to host bacterium (Waldor 1998; Davis and Waldor 2000; Boyd and Brussow 2002). The prophage database (Srividhya *et al* 2006) (*http://bicmku.in:8082*) details the annotation of prophages, cryptic phages and phage remnants from literature data. Prophages maintain most of the gene clusters associated with viable bacteriophages including integration-excision genes, head-tail terminase cluster, regulator-repressor and two component cell lysis genes.

1.1 *Phage lysis strategies*

Bacteriophages get away from bacterial cells at the end of their vegetative cycle. This crucial event involves the dissolution of the continuous strong peptidoglycan network in bacteria (Young 1992). This phenomenon associated with all eubacterial phages is defined as "two component cell lysis system" (Young *et al* 2000; Young 2002). The two

Keywords. DLP12: endolysin; holin; lambdoid; lysis cassette; over expression; prophages; purification; two-component

components are holin and lysozyme (also termed endolysin or lysins). Holins are small bacteriophage-encoded cytoplasmic membrane proteins that accumulate during the period of late-protein synthesis after infection and produce non-specific lesion on the membrane aiding the transport of murien degrading enzyme, endolysin resulting in host lysis. Additionally in many phages two overlapping genes encode auxillary lysis proteins, Rz and Rz1, which are reported to attack oligopeptide links in the outer membrane (Young *et al* 2000). Bacteriophage encoding larger genomes encode two component cell lysis cassette involving holin – lysozyme/ endolysin (Young 2005). Contrastingly phages with small genomes employ only one protein in accomplishing peptidoglycan cleavage. Single gene lysis systems are reported in phage *Φ*X174, MS2 and Q*β* (Atkins *et al* 1979; Winter and Gold 1983; Young *et al* 2000). As reported in *Φ*X174, single gene E targets cell wall biosynthesis pathway (Bernhardt *et al* 2000).

1.2 *Holin forms and characteristics*

Holins, belonging to pores and channel family of transport proteins have dual start motif, two or more membrane spanning alpha helical domains separated by *β* turn and highly charged residues at the C-terminus (Young and Blasi 1995, Blasi and Young, 1996). These mediate facilitated diffusion in addition to endolysin export (*http://www.tcdb.org/tcdb/*). They are grouped into three classes on the basis of sequence similarity and the number of potential transmembrane domains (Young and Blasi 1995; Young *et al* 2000; Young 2005).

Lambda holin protein designated as S^{λ} perhaps is the extensively studied among the holin class of proteins both in terms of genetic analysis and biochemical approaches (Chang *et al* 1995; Raab *et al* 1988; Smith *et al* 1998a,b; Grundling *et al* 2000a,b, 2001; Deaton *et al* 2004). Although the lysis systems of lambda, p21 and P22 were considered to be functional homologs due to the advent of their dual start motif conservation (Bonovich and Young 1991), recent reports suggests the possibility of different regulation mode operative in p21 holin involving dynamic topology changes of the TMD1 (transmembrane domain 1) (Park *et al* 2006). Holins are also seen to be associated with toxicity, as evidenced from *Clostridium difficile*, where they are reported to mediate the transport of toxic proteins TcdA and TcdB (Tan *et al* 2001). Understanding of holin function requires more sequence-structure-function relation data.

1.3 *Murein degrading endolysins: modular features*

Phage encoded endolysins are diverse with four different kind of muralytic activities directed against the three different covalent linkages that maintain integrity of the cell wall (Young 1992; Loessner 2005). Structure and lytic activity of *Bacillus anthracis* prophage BaO2 endolysin (plyL) has been reported which contains an N-terminal amidase and a C-terminal domain (Low *et al* 2005). Most of them lack a signal peptide sequence and depend on cognate holins for release into periplasm for cleaving peptidoglycan. Interestingly few endolysins (P1, F0g44) carry signal transmembrane domain, Signal Arrest Release (SAR domain) for their export to membrane for lysis (Sao-Jose *et al* 2000; Xu *et al* 2004). The nature of regulatory mechanism of such enzymes via disulphide bond isomerization has been established structurally in P1 endolysin. Essentially the SAR domain mediates association with the membrane by acting as a signal sequence and also favours cysteine residue isomerization event upon membrane association (Xu *et al* 2005).

2. Materials and methods

2.1 *Sequence analysis*

Prophage encoded holins and endolysins were fetched form prophage database**.** Bacteriophage holin and endolysin sequences were retrieved using the EXPASY server (*www.expasy.org*). Primary sequence analysis were done using EMBOSS (Rice *et al* 2000). PSI-BLAST (Altschul and Koonin 1998) and HMM (*http://hmmer.janelia.org/*) searches were considered only when significant at *E*-values of 0.01 or less. Multiple sequence analysis of protein sequences was carried out using CLUSTALW (Thompson *et al* 1997). NJ plot (Perrière and Gouy 1996) and PHYLODRAW (Choi *et al* 2000) were employed for constructing cluster trees. Only nodes with greater than 50% bootstrap values over 1000 trials were considered significant. Transmembrane spanning segments were predicted using TMHMM prediction server at the centre of Biological Sequence Analysis (CBS) website *http://www.cbs.dtu.dk*.

2.2 *Bacterial strains and plasmids*

The constructs of *essd* and *ybcS* were obtained as clones in pCA24N vector. This vector has the ORF cloned at the *SfI* restriction site with N terminal His Tag with and GFPuv4 at the C terminus transformed into *E. coli* AG1 strain [recA1 endA1 gyrA96 thi-1 hsdR17 (rK_mKþ) supE44 relA1]. The ORF is placed under the control of IPTG inducible promoter, pT5/lac. The clones were obtained form ASKA collection constructed and maintained by Dr Hirotada Mori, Nara institute of Science and technology, Japan (Kitagawa *et al* 2005). Further the deletion constructs harbouring inframe *ybcS* deletion in *E. coli* BW25113 (rrnB3

DElacZ4787 DEphoBR580 hsdR514 DE (ara-BAD) 567 DE (rha-BAD) 568 galU95 DEendA9::FRT DEuidA3:: pir(wt) recA1 rph-1) strain was also obtained from the same source (Baba *et al* 2006).

2.3 *Over expression using glucose-IPTG Switch (catabolite repression)*

In the case of HolinGFP, *E. coli* AG1 strain harbouring pCA24N-essdGFP was grown overnight in LB medium supplemented with 1% Glucose containing 25 μg/ml Chloramphenicol. Five percent inoculum was used in fresh LB medium with 1% Glucose and grown under shaking at 37° C to an OD_{600} of 1. Over expression was induced by pelleting cells and suspending them in LB with isopropyl-β-D thio galactopyranoside (IPTG) to a final concentration of 1mM and incubated further for 5 hours at 25°C in a shaking incubator. The cell pellet obtained after induction showed green fluorescence. In the case of *ybcS* (pCA24N-*ybcS*) over expression was carried out as like HolinGFP except that IPTG to a final concentration of 0.4mM was used.

2.4 *HolinGFP purification using two-step lysis buffers*

Cell pellets were suspended in cold lysis buffer I (250 mM sucrose, 20 mM Tris pH 7.5, 1 mM EDTA, 0.3 mg/ml lysozyme) for 20 min in ice followed by cold lysis buffer II (20 mM Tris pH 7.5, 1 mM EDTA, 300 mM NaCl, 0.1% Triton X-100) incubation for further 20 min. The resulting cell suspension was lysed by sonication and centrifuged at 12000rpm (REMI CPR30 rotor No2) for 15 min at 4°C to remove insoluble material. The crude protein was purified using talon affinity column. Unbound protein were washed with buffer A (Tris pH 7.5, 300mM NaCl, 10%Glycerol, 1mM PMSF) and eluted with 50 mM and 150 mM imidazole elution with 0.2% Triton X 100 in buffer A. The green fluorescing fractions were pooled for further experiments.

2.5 *ybcS* purification by metal chelate affinity chromatography

Cells were harvested by centrifugation after 5 h induction. Pellets were resuspended in lysis buffer (sodium phosphate buffer pH 7.0 with 250 mM sucrose, 5 mM imidazole, 500 mM NaCl and 0.3 mg/ml lysozyme). Cells were incubated on ice for 30 minutes and lysed by sonication and centrifuged at 12000 rpm (REMI CPR30 rotor No2) for 15 min at 4°C to remove insoluble material. Endolysin containing supernatants were sterilized by filtration (0.22 um-pore size, Millipore). The His-tagged Endolysin was purified from crude cell extracts under native conditions using cobalt based talon affinity column (BD biosciences). Unbound proteins were washed from the column using buffer B (sodium phosphate buffer pH 7.0, 500 mM NaCl, 1 mM PMSF, 10% glycerol). Further pure fractions of *ybcS* could be eluted with 100 mM and 150mM imidazole in buffer B. The proteins appeared as homogenous band when analysed on SDS-PAGE.

2.6 *Detection of lytic activity by SDS polyacrylamide gel electrophoresis*

The zymogram protocol reported for LLH lysin was followed (Vasala *et al* 1995). Enzymatic activity was detected *in situ* in 12% polyacrylamide gel containing either *E. coli* BW25113 having an inframe deletion of *ybcS* or *S. typhi* cell wall as substrates. Bacterial cells from overnight culture were harvested by centrifugation [7000 rpm (REMI CPR30 rotor No 2), 5 min 4°C], suspended in 4% SDS and boiled for 10 min. This was followed by centrifugation at room temperature and the pellet was washed once with distilled water, collected by centrifugation and suspended in 1 ml of distilled water. The cell suspension was mixed with other components of the 12% separating gel mix (10 ml). Samples were mixed with 5X sample buffer. Samples were not heated prior to electrophoresis. After electrophoresis the gel was incubated for 8–10 h in 100 ml of 25 mM Tris pH 7.5 and 1% Triton X-100 to permit protein renaturation. Transparent bands were visible in gel where lytic activity was present. The gel was photographed against a dark background.

2.7 *Fluorescence spectroscopy*

HolinGFP was checked for conformational variation of the GFP counterpart by using fluorescent emission spectroscopy using F-2500 FL spectrophotometer (HITACHI). Around 8 μg of protein was diluted in buffer A. The emission spectrum was scanned from 250 nm to 800 nm after excitation at 494 nm. Cuvette with 1 cm path length was used with parameters: scan speed 1500 nm/minute with response of 0.08 s and emission and excitation slit width of 10 nm.

3. Results and discussion

3.1 *Prophage encoded two-component cell lysis cassette*

A total of 73 endolysins and 41 holins are reported to be encoded by bacteriophages. Molecular details on the lytic pathway of prophages are unclear; as such prophage-encoded systems have not been examined in detail. Among the prophage encoded systems studied, PBSX gene of *Bacillus subtilis* (Krogh *et al* 1998) and *Borrelia burgorferi* plasmid cp31 plasmid associated BlyA and BlyB encoding holin

and endolysin respectively were the first evidence of prophage encoded lysis systems (Damman *et al* 2000). Additionally, the *B. anthracis* prophage endolysin, PlyL, harbouring the N-terminal catalytic domain and a C-terminal cell wall-binding domain has been structurally characterized (Low *et al* 2005). Thus structure, function and enzymatic features on prophage encoded lysis genes remain to be elucidated.

The analysis of the prophage database (*http://bicmku.in:8082*) (Srividhya *et al* 2006) showed 56 putative holins and 63 putative endolysins that are prophage encoded (table 1). Forty-seven prophages are seen to have both partners of the lysis cassette. Amongst the 47, a non-redundant set of 35 prophage encoded two-component cell lysis genes holin-endolysin were taken for further analysis (table 2). In addition the auxillary proteins of lysis cassette namely Rz and Rz1 (Young *et al* 2000) are also seen to been associated with some prophages. Ten prophages harbor Rz endopeptidase (table 3). Four prophages (DLP12, Sp4, Sp5, and Sp6) have both the Rz endopeptidase and Rz1 precursor component (table 3). The sequence based profile analysis of these prophage-encoded components of the lysis cassette has been used to cluster them and examine their relationship with phage encoded lysis systems. This helped to identify and characterize a cryptic prophage encoded lysis system from DLP12 that has similarity to functional bacteriophage p21 and P1 lysis components.

Table 1. Prophages encoding holin-endolysin

Host organism	Prophage name
Escherichia coli K-12	DLP12, Qin
Escherichia coli EDL933	Cp933w, cp933r, cp933n, cp933u , cp933p cp933oc
	Cp933k ,cp933oa, cp933xa ,cp933m , cp933v
Escherichia coli CFT073	Cp073-2 , cp073-1, cp073-5 , cp073-4
Escherichia coli O157:H7 Sakai	Sp17, sp18, sp5 sp9 sp10 sp4 sp6 sp15 sp12 sp11 Sp14 sp3 sp8
Salmonella LT2	Fels2 gifsy2
Salmonella enterica CT-18	Sti4b
Shigella flexneri2a301	Flex3
Yersinia pestis CO92	Yp5
Ralstonia solanacearum GMI1000	Rs1
Pseudomonas putida KT2440	pp01
Shewanella oneidensis MR-1	Lambdaso Muso1
Xylella fastidiosa 9a5c	Xfp5 Xfp4 Xfp3
Xylella fastidiosa temecula	Xpd4 xpd6 Xpd7 Xfp8

3.2 *Prophage encoded holins: sub classification, over expression and purification*

Holin sequences from bacteriophage perspective in gram positive and gram negative genomes have been extensively reviewed (Young and Blasi 1995; Wang *et al* 2000). The putative holins from the prophages (table 2) do not have any structural homologs and no holin is characterised structurally till date. Clustering was seen in prophage holins based on the dual start motif and transmembrane domains. The four groups with reference to holin (figure 1A) include DLP12 group (dual start with lysine residue motif conserved), CP073 and Sp group (dual start with tyrosine residue motif conserved), Phispec group (no dual start motif and conserved TMDs, has C- terminal charged residues) and List group (no dual start motif and no TMDs, no C-terminal charged residues). On comparison with the bacteriophage sequences, although few bacteriophage sequences clustered out separately, most of the prophages were seen to cluster with few other bacteriophages. Interestingly the cryptic DLP12 prophage and functional bacteriophage p21 lysis cassette genes were seen to cluster together. In HMM and PSI-BLAST based searches of prophage holins, holins from bacteriophages P21, A18, A50 were retrieved along with DLP12 and Qin prophage putative holins. Using the prophage and bacteriophage holin set, 19 bacteriophages sequences along with Qin and DLP12 prophage holins could be identified by the HMM approach (figure 2A).

The putative holin encoded by DLP12 prophage (*essd*) has 2 potential transmembrane as seen with class II holins. Using TMHMM, two potential membrane loci were predicted to be from residue 7 - 26 and residue 36 - 58 (figure 3A (I)). Interestingly the same regions are seen to align with p21 holin transmembrane domains that seem to be consistent with class II holin membrane topology model. The dual start motif and C-terminal charged residues are also conserved (figure 3A (II)). The DLP12 holin shows 96% identity to p21 holin over the full length of 71 residues. The p21 holin is reported to be functional encoding two products with 71 and 68 amino acids due to the presence of dual start motif as seen in the case of lambda but takes up different regulation strategy (Young and Blasi 1995; Barenbiom *et al* 1999; Park *et al* 2006). The sequence similarity of the putative holin from the cryptic DLP12 prophage with the functional holin from the bacteriophage p21 suggests the possibility of exchanges or transfer between bacteriophages and microbes via prophage states. The phenotype contributed by prophage holin to their respective host cells still remains to be explained. Thus functional and structural characterisation of selected prophage targets such as the DLP12 holin will help providing greater insights into such a process.

3.2.1 *Over expression, purification and characterization of DLP12 holin:* The study of phage lysis genes has been

Table 2. Prophage holins and endolysins

	Endolysins				Holins		
Prophage	location	length	Acc no	Pdb homolog	Acc no	length	location
Fels2	2859833..2860342	169	q8zmt7	1xju (P1)	q8zmt8	142	2859028..2859456
Dlp12 **	576836..577333	165	p78285	-	P77242	71	576621..576836
cp073-4 *	1421893..1422390	165	q8fi81	-	P0A9R3	71	1421678..1421893
cp933xa*	1710857..1711333	158	q8x705	1d9u (lambda)	Q8X706	101	1710565..1710870
Sp5**	1272867..1273400	177	Q8X526	-	Q9KXD4	71	1272647..1272862
Sp3	902021..902518	165	Q8X877	-	Q8X3E5	71	901806..902021
cp933m *	1270754..1271104	116	q8x4i0	-	Q8X573	68	1270148..1270354
Sp9	1777316..1777849	177	Q7AEQ9	-	Q7ABG7	71	1776701..1776916
Sp6**	1558372..1558905	177	Q7ADW1	-	Q8X4K3	71	1557757..1557972
Lambdaso	3104244..3104756	170	q8ecz8	-	Q8ED02	109	3102501..3102830
Flex3*	722382-722915	177	q83s56	-	q83s57	71	721468..721683
Sp4**	1184947..1185480	177	Q8X539	-	Q8X538	68	1186168..1186374
Sp14	2692072..2692605	177	Q8X4U1	-	Q8X4J9	68	2693005..2693211
cp073-2*	1345465..1345998	177	q8fii6	-	Q8fii9	106	1344452..1344772
cp073-5	3049754..3050287	177	q8fev1	-	q8feu9	71	3051165..3051380
cp933u	2769889..2770422	177	q7dbg5	-	Q8X8Y2	104	2770822..2771136
pp01	4378171..4378710	179	q88g71	1d9u (lambda)	Q88G43	106	4403714..4404034
Bh2	1037887..1038903	338	Q9KE90	-	Q9KE89	78	1038924..1039160
bl1	481637..482515	292	Q8G788	-	Q8G787	87	482512..482775
cp933n *	1644354..1644887	177	Q8XEE4	-	Q8X894	104	1643640..1643954
pi3	1416048..1416827	259	Q9CFT6	-	Q9CFT5	99	1416827..1417126
List-I4	1715216..1716142	308	Q92B50	1yb0 (prophage lambdaba02)	Q92B48	86	1716378..1716638
List-I1	113381..114250	289	Q92FH8	-	Q92FH9	93	113100..113381
List-M2	2364754..2365599	281	O52604	-	Q8Y500	93	2365599..2365880
ob1	239037..239888	283	Q8ETN8	-	Q8ETN9	77	238736..238969
pp02	3446684..3447235	183	Q88ID3	-	Q88IF9	115	3424921..3425268
phin315	2008308..2008571	87	Q99QR8	-	Q7A4M9	84	2010535..2010789
Phimu50B	2086939..2087670	243	Q931M5	-	Q99SU5	84	2087682..2087936
phisa2mw	1531688..1533142	484	Q8NWL7	-	Q8NWL6	100	1533153..1533455
phisa3mw	2049938..2050693	251	Q8NVR1	-	Q7A0F5	84	2050705..2050959
lambdasa2	1834175..1835581	468	Q8DXL3	-	Q8DXL2	109	1835585..1835914
phispecA	330218..331435	405	Q8P2E4	-	Q7CNI6	75	329872..330099
315.5	1316849..1317613	252	Q8K6K0	-	Q8K6J9	110	1317615..1317947
315.3	1139031..1140245	404	Q8K6W9	-	Q8K6W8	151	1140357..1140812
qin	1637548..1638081	177	P76159	-	P77237	96	1638394..1638684

*Prophages harbouring endopeptidase; **prophages having endopeptidase precursor.

complex since the proteins encoded by these genes are designated to kill bacteria and are usually expressed at low levels. Consequently little information is known about prophage encoded lysis systems. Overexpression of DLP12 holin could not be accomplished, due to slow growth or death of the cells, by the classical IPTG induction methods. Instead over expression of holinGFP was successful by glucose IPTG switch (catabolite repression) resulting in a 36-kDa-protein product corresponding to the theoretical molecular mass of holin-GFPuv4 fusion protein (figure 4A).

Table 3. Prophages encoding holin-endolysin with endopeptidase and endopeptidase precursor

	Endopeptidase (Rz endopeptidase)			Endopeptidase precursor (lipoprotein Rz1 precursor)		
Prophage	location	length	Acc no	location	length	Acc no
DLP12	577330..577791	153	P75719	577550..577732	60	P58041
cp073-4	1422375..1422848	157	Q8FI80			
cp933xa	1711474..1711791	105	Q8X704			
Sp5	1274394..1274858	154	Q9KXD2	1274614..1274799	61	q9kxd1
cp933m	1272282..1272749	155	Q9EYC9			
Sp6	1559899..1560366	155	Q8X2A4	1560119..1560325	68	q8x3h6
Flex3	722897..723316	139	Q83S55			
Sp4	1184144..1184611	155	Q8X5F1	1184206..1184391	61	Q8X5F1
cp073-2	1345980..1346462	160	Q8FII5			
cp933n	1646079..1646552	157	Q8X3X6			

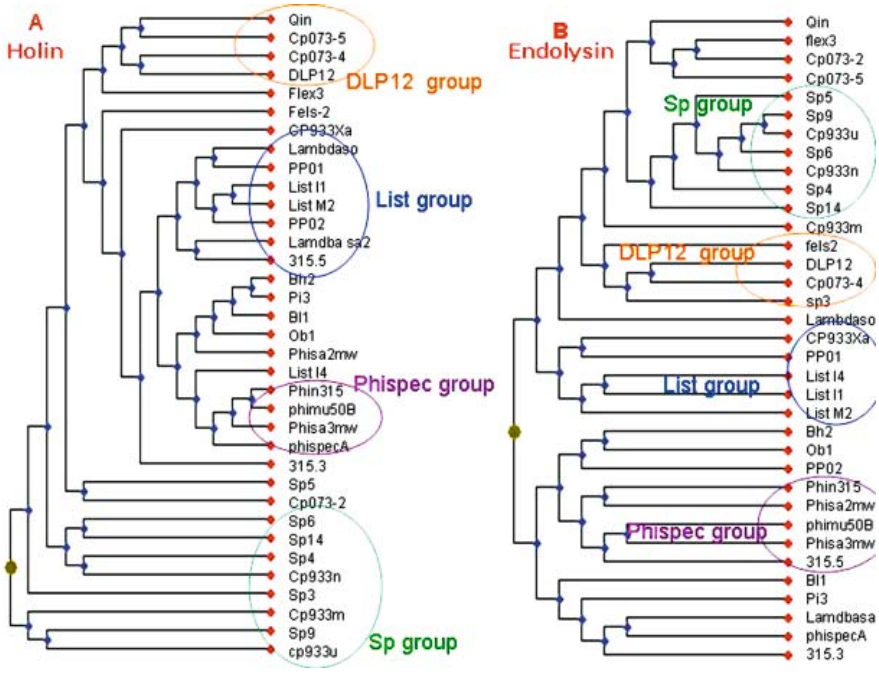

Figure 1. Clustering of prophage encoded holins and their cognate endolysins**.** For holin:DLP12 group (orange): Dual start with lysine residue motif conserved; for CP073 and Sp group (green): dual start with Tyrosine residue motif conserved; for Phispec group (magenta) : no dual start motif and no conserved TMDs; for List group (blue): no dual start motif and no TMDs. In the case of endolysins, the SAR domain conserved group, DLP12 group clusters separately from the rest.

The conformational analysis of GFPuv4 in the holin GFP fusion protein was assessed by fluorescence spectroscopy. The emission spectrum (figure 4B) at 494 nm excitation with a 512nm peak characteristic of GFPuv4 as reported by Ito *et al* (1999) confirmed the native folded state of recombinant fusion protein. The putative holin encoded by the cryptic prophage DLP12 is possibly functional and structurally stable. Being a close relative of S^{21}, which does encode two functionally opposed polypeptides, mutagenesis experiment needs to be carried out to understand the role of if any of the dual start motif of holin encoded by DLP12 prophage.

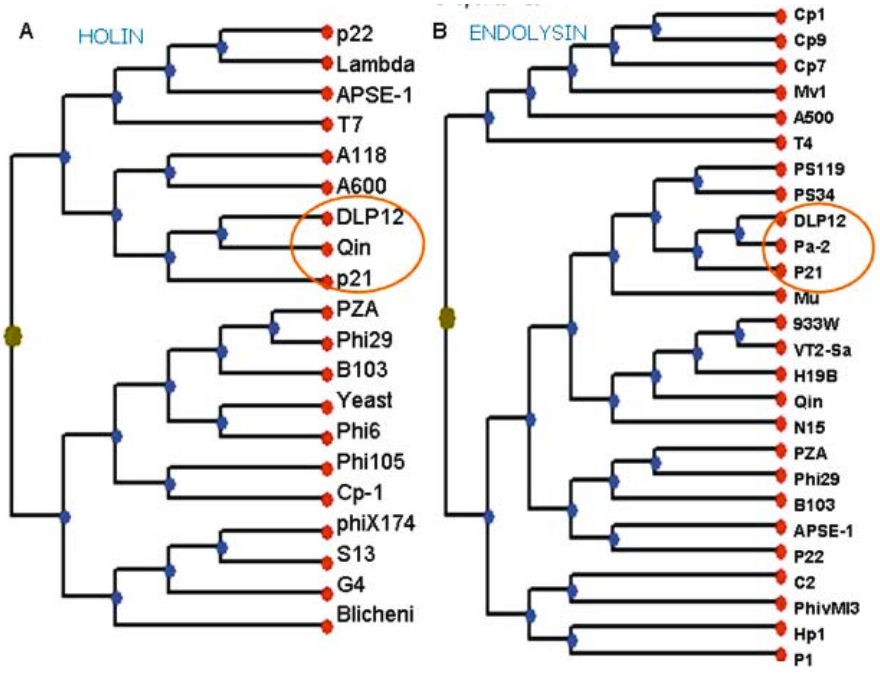

Figure 2. NJ plot of HMM hits of bacteriophage –prophage lysis cassette.

3.3 *Prophage endolysins: subclass categorization, purification and characterization*

Endolysin sequences were seen to cluster similar to holin groups (figure 1B). The DLP12 cluster with sp3 and CP073-4 prophages had the SAR domain sequences conserved. Prophages from *E. coli sakai* strain were seen to have a different N terminal domain instead of SAR and clustered separately. *Streptococcus* prophages Phispec group and *Listeria* prophages with no SAR domain clustered separately into two different groups (figure 1B). Alignment of prophage- bacteriophage endolysin sequences show clear clustering of DLP12 endolysin with p21 phage. By applying HMM search methods bacteriophage encoded endolysin from p33, H19, N15 and P21 along with DLP12, Qin prophage endolysin could be selected. With endolysin a total of 24 bacteriophages along with DLP12 prophage sequence could be retrieved by HMM based search based on prophage-bacteriophage alignment. Among the retrieved genomes, cognate cell lysis partners were retrieved for 8 bacteriophages along with DLP12 prophage (table 4). Figure 2B shows the clustering out of DLP12 with p21 phage endolysin.

From the analysis most of the bacteriophage and prophage encoded two-component cell lysis genes are seen to be coevolved. On the basis of similarity searches and HMM based approach, it is likely that DLP12 cell lysis cassette is a close relative of p21 bacteriophage cell lysis genes (figure 2B), having 98% identity over 165 residues. Interestingly DLP12 Endolysin has the SAR domain conserved with p21, PA-2 endolysin, except for Cys13 that is seen only in P1 (figure 3B (I)). Subsequently over expression and purification of *ybcS* was carried out to examine the sequence-structure-function relation.

3.3.1 *Over expression, purification and characterization of* ybcS*:* Interestingly over expression of *ybcS* cloned under an inducible promoter in log phase cells in the absence of cognate holin gene resulted in overt lysis by 35 minutes after induction. The culture turbidity significantly decreased (figure 5A) after induction at 37° C. Inspection of *ybcS* sequence revealed the conserved SAR signal sequence as that of P1 (figure 3B (I)). Further in DLP12 endolysin the SAR domain was seen to be conserved with that of P1, p21 and PA-2 endolysins. Additionally N-terminal SAR domain residues are predicted to be membrane associated as predicted by TMHMM (figure 3B (II)).

This SAR domain is proposed to be potentially capable of serving as signal peptide for traversing the membrane. The export of P1 endolysin is reported to be mediated by SAR (Signal Arrest Release) by employing the host

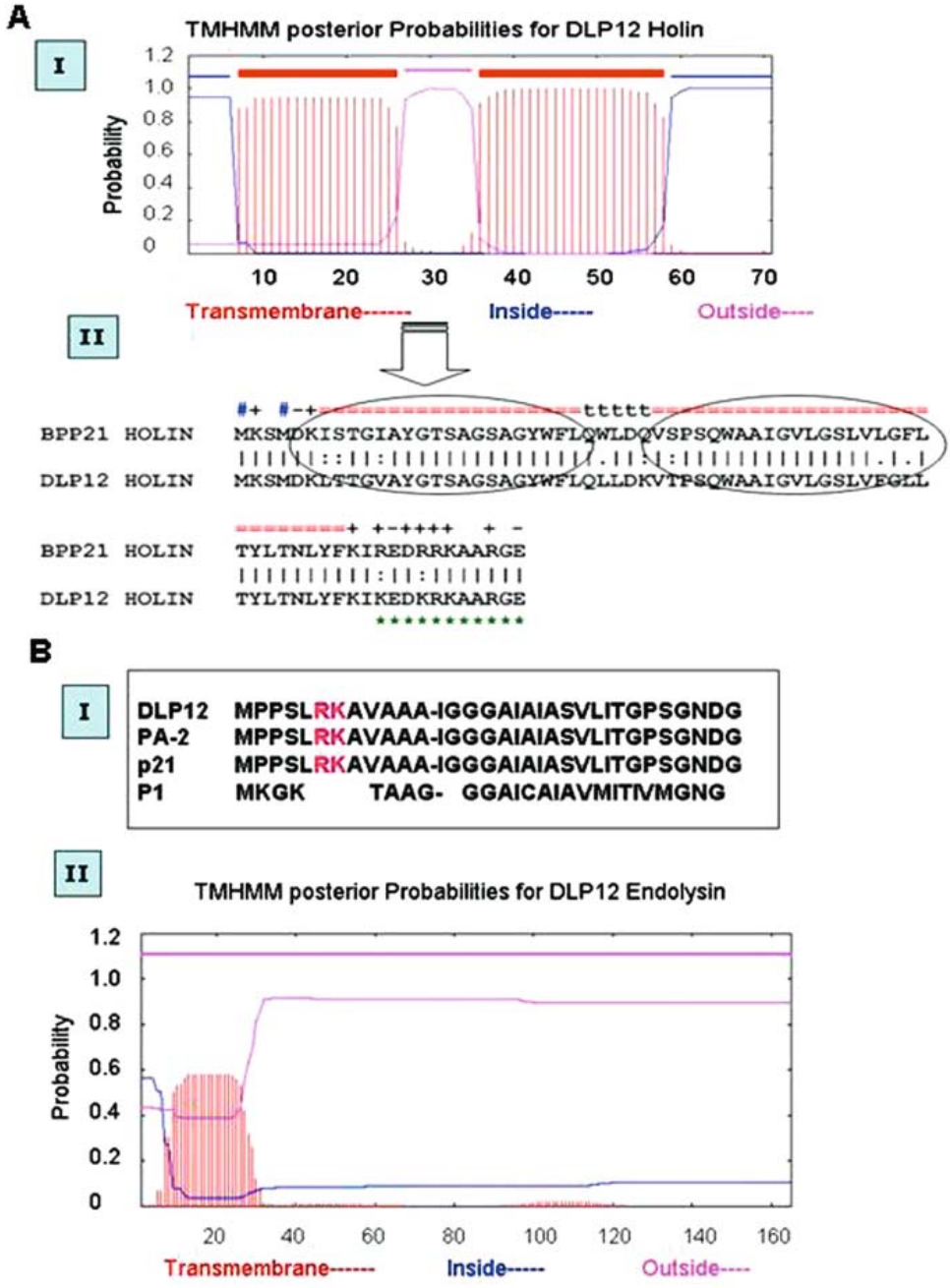

Figure 3. **(A)** (I) Predicted probability for transmembrane domains (transmembrane domains in red) as determined by TmHMM. (II) Alignment of DLP12 prophage encoded holin and predicted transmembrane segments. For comparision, the primary sequence of p21 holin is also depicted. Predicted transmembrane segments shown in bars (===). Potential beta turn regions indicated by't'. Charged residues indicated by '+'and '–'. Beginning of dual start motif indicated by #. Highly charged C terminus indicated by *. **(B)** (I) Alignment of N-terminal SAR domain sequence of DLP12 endolysin with bacteriophage PA-2, p21 and P1 endolysins. (II) Predicted probability for SAR transmembrane domain at N – terminus (transmembrane domains in red) as determined by TmHMM in *ybcS* sequence**.**

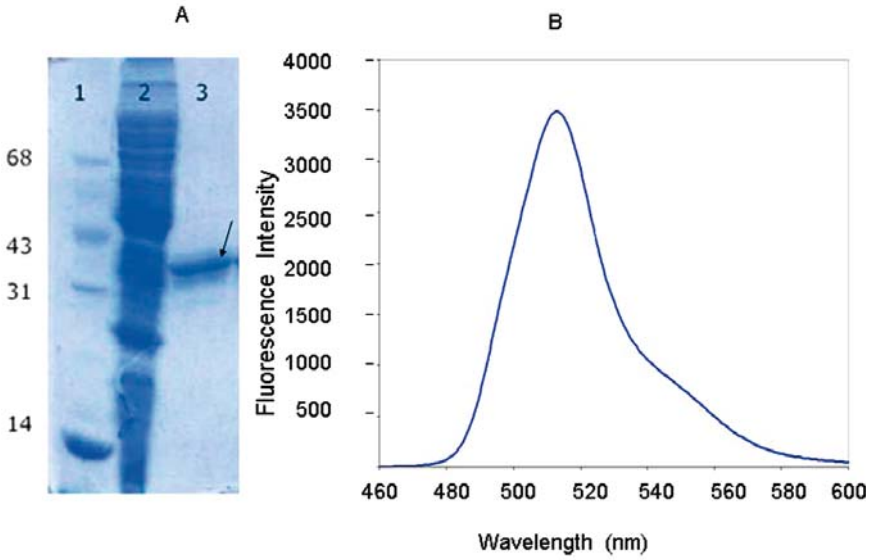

Figure 4. Purification and fluorescence emission spectra of holinGFP. **(A)** Lane 1, marker; lane 2, crude supernatant; lane 3, 50 mM imidazole eluted. **(B)** Fluorescence measurement of HolinGFP. Characteristic emission peak at 512 nm upon excitation at 494 nm as reported for GFPuv4 (Ito *et al* 1999).

Table 4. Cognate holin-endolysin in bacteriophages and prophages

Holin	Bacteriophage/prophage	Endolysin
LYS_BPB03	Bacteriophage B103	VG14_BPB03
LYS_BPPZA	Bacteriophage PZA	VG14_BPPZA
LYS_BPPH2	Bacteriophage phi-29	VG14_BPPH2
LYSD_ECOLI	prophage DLP12	ESSD_ECOLI
LYS_BPP21	Bacteriophage 21	VLYS_BPP21
LYS_BPAPS	Bacteriophage APSE-1.	VLYS_BPAPS
LYS_BPP22	Bacteriophage P22	VLY1_BPP22
LYS_BPCP1	Bacteriophage Cp-1.	VLYS_BPCP1
AEPE_BPA50	Bacteriophage A500	HOLI_BPA50

secretory function (Xu *et al* 2004). In the case of *ybcS*, with Glucose-IPTG switch method and at low temperature induction, successful over expression of *ybcS* could be achieved and detected using electrophoresis (figure 5B). The over expressed endolysin from DLP12 was shown to be active using a zymogram containing *E. coli BW25113* (with inframe *ybcS* gene deletion), a lucent band at 20 kDa was detected (figure 5C). The results suggest that the over expressed protein is active against *E. coli* and *S. typhi* cell wall as substrate.

In P1 phage, endolysins are thus capable for exporting themselves and their cognate holins are thought to play role in membrane depolarisation release of endolysins. (Xu *et al* 2005). A classical one-component system is also encoded by F0g44 (Sao-Jose *et al* 2000). In well-studied cases of lambda, holin is absolutely essential for scheduling lysis.

The P1 endolysin and DLP12 do not share significant sequence similarity except over the SAR domain at N terminal region. Therefore, as proposed by Xu *et al* (2004) as like the P1 endolysin, DLP12 might possibly get localised in cytoplasm. However unlike P1, DLP12 endolysin lacking the Cys residue might follow some other catalytic mechanism rather than disulphide bond isomerization. We have shown conclusively that lysis is mediated by *ybc*S without its cognate holin when over expressed *in vitro* in *E. coli* .To our knowledge, this is the first report of its kind for a prophage encoded endolysin with SAR domain. This supplements the results reported with other SAR endolysin P1 and F0g44, which harbour SAR secretary signal peptide to traverse the cell membrane. A possible explanation could be that N terminal SAR contributes to its export from periplasm as in the case of P1, but takes up a different mechanism other than disulphide bond isomerization.

3.4 *DLP12 prophage: cell lysis cassette*

DLP12 a cryptic prophage of 21302 bp is immediately downstream of argU tRNA gene located in *E. coli* K-12 (locus 564025-585326) and has a GC content of 0.43. It is lambda–like (Lindsey *et al* 1989) and encodes the putative two-component lysis cassette Holin-endolysin.

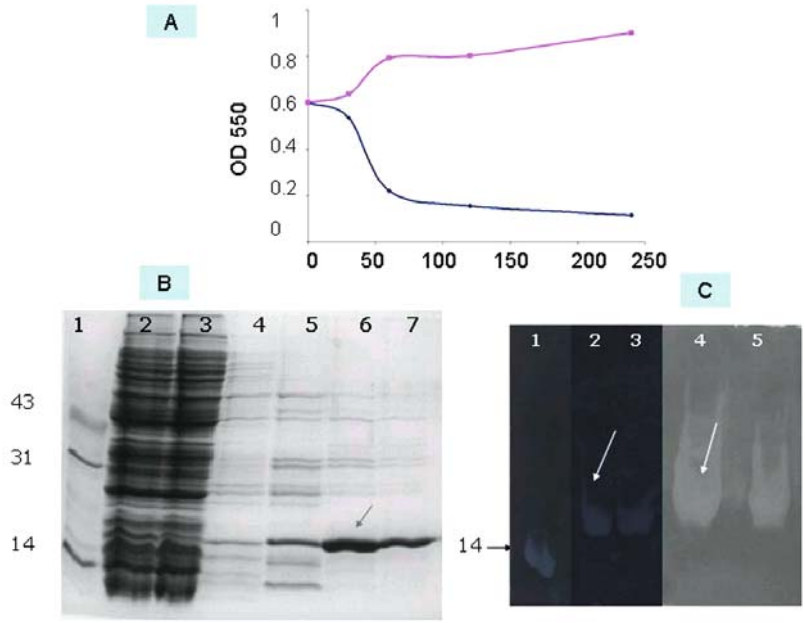

Figure 5. **(A)** Functionality of DLP12 endolysin. Cell lysis observed within 30 minutes of induction with 0.1mM IPTG (direct induction in LB media). **(B, C)** Purification of DLP12 endolysin and demonstration of lytic activity in a zymogram with *S. typhi Ty21a* and *E. coli* cell wall as substrate. **(B)** Lane 1, marker; lane 2, crude supernatant; lane 3, flow through; lane 4, 10 mM imidazole eluted; lane 5, 50 mM imidazole eluted; lanes 6, 7, 100 mM imidazole eluted. **(C)** lane 1, chicken egg white lysozyme; lanes 2, 3, *ybcS* against *E. coli* cell wall substrate; lanes 4, 5, *ybcS* against *S. typhi* cell wall substrate.

DLP12 prophage lysis module contains a putative lysin gene designated as *ybcs* and a putative holin designated as *essd* placed upstream of the endolysin similar to the genetic arrangement seen in many bacteriophages. DLP12 prophage is also seen to have the auxiliary proteins Rz endopeptidase and Rz1. Out of the total prophages, only 4 prophages are seen to have holin, endolysin, endopeptidase (Rz) and endopeptidase precursor (Rz1) (table 3). Further in HMM searches of prophage Rz endopeptidase and Rz1 sequences, the Rz-Rz1 proteins of bacteriophage lambda, p21 and p22 could be retrieved. The investigation of the DLP12 prophage cell lysis cassette would throw more light on the role of endopeptidase and the precursor in the context of the two-component cell lysis cassette.

4. Conclusions

In conclusion cell lysis by phages is a highly regulated and evolved sophisticated process. This machinery, in general, involves holins and endolysins. Surprisingly prophages that are cryptic or defective are also found to harbor the components of the lysis cassette. We have shown that the prophage and bacteriophage lysis components are related. Moreover, the *in silico* analysis of the cryptic DLP12 prophage which showed similarity in its cell lysis cassette to the functional bacteriophages like p21 has led to the choice of targets for *in vitro* and *in vivo* characterization. The holin and endolysin from DLP12 have been purified and partially characterized. We also suggest that from our analysis that the endolysin from DLP12 is likely to be different in the mechanism of function from the P1 endolysin whose structure is known. So far no holin has been structurally characterized. Thus the understanding of the DLP12 lysis cassette in terms of the sequence-structure-function relationship is likely to enrich the understanding of one of the most dramatic events of controlled lysis in the virus-host cycle, while also enlightening the role of prophages in bacterial evolution.

Acknowledgements

We acknowledge the use of Bioinformatics Centre facility funded by the Department of Biotechnology (DBT) New Delhi, and DBT for INDO-ISRAEL project support. We thank Dr Hirotoda Mori, Nara Institute for Science and Technology, Japan for providing *essd* and *ybcS* clones. We

thank Dr A Arockiasamy and Mr M R Sankar Narayanan for timely help in full text reference collection of all crucial articles during the study.

References

Altschul S F and Koonin E V 1998 Iterated profile searches with PSI-BLAST - a tool for discovery in protein databases; *TIBS* **23** 444–447

Atkins J F, Steitz J A, Anderson C W and Model P 1979 Binding of mammalian ribosomes to MS2 phage RNA reveals an overlapping gene encoding a lysis function; *Cell* **182** 247–256

Baba T, Ara T, Hasegawa M, Yuki Takai Y, Yoshiko OkumuraY, Baba M, Datsenko K A, Tomita M, Wanner B L and Mori H 2006 Construction of *Escherichia coli K-12* in-frame, single-gene knockout mutants the Keio collection; *Mol. Syst. Biol.* doi101038/

Barenboim M, Chang C Y, Dib H F and Young R 1999 Characterization of the dual start motif of a class II holin gene; *Mol Microbiol*. **32** 715–727

Bernhardt T G, Roof W D and Young R 2000 Genetic evidence that the bacteriophage phi X174 lysis protein inhibits cell wall synthesis; *Proc. Natl. Acad. Sci. USA* **97** 4297–4302

Blasi U and Young R 1996 Two beginnings for a single purpose the dual-start holins in the regulation of phage lysis; *Mol Microbiol*. **21** 675–682

Bonovich M T and Young R 1991 Dual start motif in two Lambdoid genes unrelated to l S; *J Bacteriol.* **173** 2897–2905

Boyd E F and Brussow H 2002 Common themes among bacteriophage-encoded virulence factors and diversity among the Bacteriophages involved; *Trends Microbiol*. **10** 521–529

Brussow H, Canchaya C and Hardt W D 2004 Phages and the evolution of bacterial pathogens from genomic rearrangements to lysogenic conversion; *Microbiol. Mol. Biol. Rev*. **68** 560–602

Casjens S 2003 Prophages and bacterial genomics what have we learned so far?; *Mol Microbiol*. **49** 277–300

Chang C Y, Nam K and Young R 1995 S gene expression and the timing of lysis by bacteriophage lambda; *J. Bacteriol.* **177** 3283–3294

Choi J H, Jung H Y, Kim HS and Cho HG 2000 PhyloDraw: a phylogenetic tree drawing system; *Bioinformatics* **16** 1056–1058

Damman C J, Eggers C H, Samuels D S and Oliver D B 2000 Characterization of *Borrelia burgdorfer*i BlyA and BlyB proteins a prophage-encoded holin-like system; *J. Bacteriol.* **182** 6791–6797

Davis B M and Waldor M K 2000 Filamentous phages linked to virulence of *Vibrio cholerae; Curr. Opin. Microbiol.* **36** 35–42

Deaton J Savva CG Sun J Holzenburg A Berry J and Young R 2004. Solubilization and delivery by GroEL of megadalton complexes of the lambda holing; *Protein Sci*. **13** 1778-1786

Grundling A, Blasi U and Young R 2000a Genetic and biochemical analysis of dimer and oligomer interactions of the lambda S holin; *J. Bacteriol*. **182** 6082–6090

Grundling A, Smith D L, Blasi U and Young R 2000b Dimerization between the holin and holin inhibitor of phage lambda; *J. Bacteriol*. **182** 6075–6081

Grundling A, Manson M D and Young R 2001 Holins kill without warning; *Proc. Natl. Acad. Sci. USA* **98** 9348–9352

Ito Y, Suzuk M and Husimi Y 1999 A Novel Mutant of Green Fluorescent Protein with Enhanced Sensitivity for Microanalysis at 488 nm Excitation; *Biochem. Biophys. Res. Commun*. **264** 556–560

Kitagawa M, Takeshi A, Arifuzzaman M, Ioka-Nakamichi T, Inamoto E, Toyonaga H and Mori H 2005 Complete set of ORF clones of *Escherichia coli* ASKA library A Complete Set of *E coli K-12* ORF Archive Unique Resources for Biological Research; *DNA Res.* **12** 291–299

Krogh S, Jorgensen S T and Devine K M 1998 Lysis genes of the *Bacillus subtilis* defective prophage PBSX; *J. Bacteriol*. **180** 2110–2117

Lindsey D F, Mullin D A and Walker J R 1989 Characterization of the cryptic lambdoid prophage DLP12 of *Escherichia coli* and overlap of the DLP12 integrase gene with the tRNA gene argU; *J. Bacteriol*. **171** 6197–205

Loessner M J 2005 Bacteriophage endolysins-current state of research and applications; *Curr. Opin. Microbiol*. **8** 480–487

Low L Y, Yang C, Perego M, Osterman A and Liddington R C 2005 Structure and lytic activity of a *Bacillus anthracis* prophage endolysin; *J. Biol. Chem.* **280** 35433–35439

Park T, Struck D K, Deaton J F and Young Ry 2006 Topological dynamics of holins in programmed bacterial lysis; *Proc. Natl. Acad. Sci. USA* **103** 19713–19718

Perrière G and Gouy M 1996 WWW-Query: An on-line retrieval system for biological sequence banks. *Biochimie* **78** 364–369

Raab R, Neal G, Sohaskey C, Smith J and Young R 1988 Dominance in lambda S mutations and evidence for translational control; *J. Mol. Biol.* **199** 95–105

Rice P, Longden I and Bleasby A 2000 EMBOSS: The European Molecular Biology Open Software Suite; *Trends Genet.* **16** 276–277

Sao-Jose C, Parreira R, Vieira G and Santos M A 2000 The N-terminal region of the *Oenococcus oeni* bacteriophage fOg44 lysin behaves as a bona fide signal peptide in *Escherichia coli* and as a cis-inhibitory element, preventing lytic activity on oenococcal cells; *J. Bacteriol.* **182** 5823–5831

Smith D L, Struck D K Scholtz J M and Young R 1998a Purification and biochemical characterization of the lambda holin; *J. Bacteriol.* **180** 2531–2540

Smith D L, Chang C Y and Young R 1998b The lambda holin accumulates beyond the lethal triggering concentration under hyperexpression conditions; *Gene Expr*. **7** 39–52

Srividhya K V, Geeta V Rao, Raghavenderan L, Preeti Mehta, Jaime Pirulsky, Sankarnarayanan Manicka, Joel L. Sussman and S Krishnaswamy 2006, Database and Comparative Identification of prophages. *LNCIS* **344** 863–868

Tan K S, We B Y and Song K P 2001 Evidence for holin function of tcdE gene in the pathogenicity of *Clostridium difficile*; *J. Med. Microbiol*. **50** 613–619

Thompson J D, Gibson T J, Plewniak F, Jeanmougin F and Higgins D G 1997 The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 244 876–4882

Vasala A, Valkkila M, Caldentey J and Alatossava T 1995 Genetic and biochemical characterization of the *Lactobacillus subsp*

lactis bacteriophage LL-H lysin; *Appl. Environ. Microbiol*. **61** 4004–4011

Waldor M K 1998 Bacteriophage biology and bacterial virulence; *Trends Microbiol.* **6** 295–297

Wang I N Smith D L and Young R 2000 Holins the protein clocks of bacteriophage infections; *Annu. Rev. Microbiol*. **54** 799–825

Winter R B and Gold L 1983 Overproduction of bacteriophage Q3 maturation A, protein leads to cell lysis; *Cell* **33** 877–885

Xu M, Struck D K, Deaton J, Wang N and Young R 2004 A signal-arrest-release sequence mediates export and control of the phage P1 endolysin; *Proc. Natl. Acad. Sci USA* **101** 6415–6420

Xu M, Arulandu A, Struck D K, Swanson S, Sacchettini J C and Young R 2005 Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme; *Science* **307** 113–117

Young R 1992 Bacteriophage lysis mechanism and regulation; *Microbiol. Rev*. **56** 430–481

Young R 2002 Bacteriophage holins deadly diversity; *J. Mol. Microbiol. Biotechnol*. **4** 21–36

Young R, 2005 Phage lysis; in *From phages–Their role in bacterial pathogenesis and biotechnology* (eds) M K Waldor, D L Friedman and S L Adhya (Washington: ASM press) pp 92–127

Young R and Blasi U 1995 Holins form and function in bacteriophage lysis; *FEMS Microbiol. Rev.* **17** 191–205

Young R, Wang I and Roof W D 2000 Phages will out strategies of host cell lysis; *Trends Microbiol*. **8** 120–128

*e*Publication: 21 June 2007