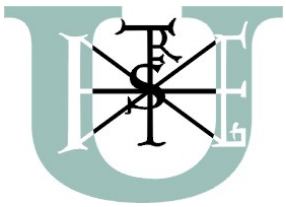

SZENT ISTVÁN UNIVERSITY

KEY COMPONENTS OF THE *16-3* PHAGE DNA PACKAGING

PhD thesis

Ganyu Anita

Gödöllő
2005

PhD school

The name of the school: **The School of Biological Sciences**

Field of science: **Biology**

Head of school: **Prof. Dr. Tuba Zoltán DSc**
Professor, Head of Department
Department of Botany and Plant Physiology
Faculty of Agricultural and Enviromental Sciences
Szent István University

Supervisors: **Dr. Papp Péter PhD**
Principal Investigator
Institute of Genetics
Agricultural Biotechnology Center

Prof. Dr. Orosz László DSc
Academican, Head of Department
Department of Genetics
Faculty of Science
Eötvös Loránd University

Approved by the head of school:

…….

Dr. Tuba Zoltán

Approved by the supervisors:

…… **……**

Dr. Papp Péter **Dr. Orosz László**

BACKGROUND, AIMS

After attachment to a specific receptor in the bacterial cell wall, the phage genome enters the cell. In case of many phages this is followed by the circularization of the genome by ligation of short complementer end sequences (*cos*, cohesive end site). During lytic cycle of the bacteriophages rolling circle replication generates concatameric substrates for DNA packaging. In parallel with DNA replication new phage proheads are produced.

The terminase enzyme is part of a large nucleoprotein complex that packages viral DNA into the capsid. The terminase enzime usually consists of a small and a large subunit and is encoded in the vicinity of the *cos*. During DNA packaging terminase recognizes phage DNA at the cohesive end sites, recruits a prohead and *cos* cleavage occurs. Then the terminase proceeds along the concatamer to the next cohesive end site while the DNA is translocated into the prohead. *cos* cleavage then occurs at the terminal end of the packaged chromosome and the filled head is released from the packaging complex.

The phage DNA packaging is ATP dependent. Previous sequence alignment studies with the large subunits of various viral terminase molecules showed that they contain putative ATPase domains (Walker A and B, Adenin binding, Motif III).

Our aims were to identify the key components of *Rhizobium meliloti* 41 (*Synorhizobium meliloti* 41) phage *16-3* in DNA packaging process. Our first aim was to identify and characterize the *cos* region of phage *16-3*. Since the coding sequence of a DNA binding protein and its target site are often found in close proximity in procaryotic system, we expected the terminase genes to be located in the vicinity of *cos*. The terminase recognizes the *cos* sites on the concatemer chain and cleaves unit-lenght phage cromosomes. If one or more ORF found by sequence analysis encoded the subunits of a terminase enzyme, their inactivation would result in uncut *cos* sites and failure in packaging of the phage DNA.

MATERIALS AND METHODS

Living materials and microbiology methods

Escherichia coli strain DH5α was used in cloning experiments and served as the host for donor plasmids used for triparental mating. *Rhizobium meliloti* 41 was used for the *16-3* phage experiments and as a recipient host for triparental mating. The *16-3* phage techniques (growth conditions, triparental mating, prophage induction) were carried out following procedures available in the literature.

Molecular biology methods

Isolation and manipulation of nucleic acids (nucleic acid preparation, restriction endonuclease cleavage, electrophoretic separation, cloning into vectors, polymerase chain reaction, Southern hybridization, DNA sequencing), protein techniques (protein overexpression, electrophoretic separation) and band-shift experiment were carried out following commonly used techniques in molecular biology as well as by following guidelines provided by the suppliers.

Bioinformatic analysis

Putative ORFs were searched against the available databases (nonredundant protein database) by using BLAST (Basic Local Alignment Search Tool) and various databases to establish amino acid sequence homologies.

RESULTS

Identification of the cohesive ends of phage 16-3

To identify the cohesive ends, we isolated the *16-3* phage DNA fragments which contain the cohesive ends. Primers were designed close to the expected ends and we determined the nucleotide sequence of the fragments. We also identified the sequence of the ligated cohesive ends. If the phage genome has 5' protruding ends, their sequences should show up in both end-fragment templates. If the phage has 3' overhangs, their sequences will not appear in either end-fragment template. A 10-bp sequence region existing in the original phage DNA fragment was absent from the sequences of both end-fragments, hence we concluded that the chromosome of phage *16-3* ended in 10-base-long, 3'-protruding, single-stranded, complementary sequences ($^{5'\text{-...CCGGCGTCGG-3'}}_{3'\text{-GGCCGCAGCC...-5'}}$). The *cos* sequence has high G/C content and shows dyad symmetry.

Identification of genes around the cos region essential for phage production

Since the coding sequence of a DNA binding protein and its target site are often found in close proximity in prokaryotic systems, we expected the terminase genes to be located in the vicinity of *cos*. ORF171 and ORF249 on the right end and ORF154 and ORF525 on the left end of the phage genome were identified as candidate genes and tested for terminase function.

Kanamycin cassette mutagenesis

To investigate the roles of the different open reading frames (ORFs), kanamycin (Km) cassette insertional mutagenesis was applied. Plasmids were constructed, which contained the kanamycin cassette in ORF249, ORF171, ORF154, and ORF525. The four plasmids were introduced into the lysogenic strain *R. meliloti* 41(*16-3cti3*) by triparental mating. Since these plasmids were not able to replicate in *R. meliloti*, Km resistant colonies indicated homologous recombination between the plasmid and the prophage either with one crossing over (resulted in the integration of

the entire plasmid into the prophage) or with two crossing overs (resulting in the insertion of the Km cassette into the targeted ORF of the prophage). Selection for Km-resistant but streptomycin-sensitive (antibiotic resistance of the vector part of the plasmids) colonies resulted in mutant lysogenic *R. meliloti* strains *Rm*41*(16-3cti3Km-249)*, *Rm*41*(16-3cti3Km-171)*, *Rm*41*(16-3cti3Km-154)*, and *Rm*41*(16-3cti3Km-525)*. These prophages contained the Km gene in ORF249, ORF171, ORF154, and ORF525, respectively.

Phage yield detection after mutant prophage induction

After prophage induction we analysed if mutations of the four ORF localised in the *cos* region influenced phage replication. Phage titration of the cell lysates on the *R. meliloti* 41 lawn indicated that, in the case of *16-3cti3*Km-171, normal phage burst occurred, however, no phage productions were detected after induction of prophages *16-3cti3*Km-249, *16-3cti3*Km-154, and *16-3cti3*Km-525. The results indicated that cistrons, represented by ORF249, ORF154, and ORF525, are vital for phage growth.

Complementation test

Phage productions were fully regained at the wild-type level when plasmids containing the wild type ORFs were introduced into lysogenic strains *R. meliloti* 41(*16-3cti3*Km-249), *R. meliloti* 41(*16-3cti3*Km-154), and *R. meliloti* 41(*16-3cti3*Km-525), respectively. Since the function of these three cistrons could be reestablished by complementation, we conclude that the insertion of the Km resistance gene did not disrupt any *cis*-acting control element of the cognate genes, and that in each case, loss of the function was due to the inactivation of the gene product itself.

Functional tests for identification of the terminase genes of phage 16-3

If ORF249, ORF154, and ORF525 encoded the subunits of a terminase enzyme, their inactivation would result in uncut *cos* sites and failure in packaging of the phage DNA. To test this hypothesis we analysed *cos* cleavage following heat

induction of the mutant prophages. Southern hybridization of samples obtained from prophage induction in *R. meliloti* 41(*16-3cti3*Km-249), *R. meliloti* 41(*16-3cti3*Km-154), and *R. meliloti* 41(*16-3cti3*Km-525) showed only covalently closed *cos* sites, indicative of the malfunction of terminase activity. In contrast to this, Southern analyses of samples derived from induction of prophages, containing the kanamycin cassette inserted into ORF171 indicated wild-type terminase activity.

Bioinformatic analyses of the ORFs

The ORF249 showed significant homology to homing endonucleases (44% identity, 61-63% similarity). Homology of ORF171 to other proteins was based on its C-terminal residues (110-171 aminoacids). The closest relatives were holin proteins of different phages (41-54 % identity, 55-62% similarity).

ORF154 showed high homology only to the small terminase subunit of phage *HK022* and *HK97* (34% identity, 53% similarity) and no reasonable homology was found to other proteins in the database.

In case of ORF525, the highest homology was found to the terminase large subunit of phages *HK97* and *HK022* (52% identity, 66% similarity).

Functional tests and bioinformatical analysis with the four ORFs proved that the ORF154 is the small subunit and the ORF525 is the large subunit of the *16-3* terminase. We renamed ORF154 and ORF525 as *terS* and *terL*, *respectively*.

Gel mobility shift assay with the TerS

We analyzed the DNA binding of the overexpressed TerS protein. Gel mobility shift assay indicated that TerS protein can bind to a DNA fragment carrying the *cos* site and the flanking regions.

ORFs with alternative start codons in ORF171

Although ORF171 did not seem to be essential for *cos* cleavage and phage production, there are also alternative translational start sites (ATG) resulting in

ORF165 or ORF76. Insertion of a Km cassette at codon position 56 (counting according to ORF171) ruled out the role of ORF171 and ORF165 but not ORF76.

We have tested a mutant prophage in which the Km cassette replaced almost the entire ORF76 (1 to 63 codons deleted). The mutation resulted in no *cos* cutting and no phage production. Since phage production could not be complemented by ORF76 supplemented in *trans*, it is likely that the changes disrupted a *cis* element involved in *cos* cutting and/or in phage production.

Novel results

1) We identified and characterised the *cos* sites of the *16-3* phage which are 10-base-long, 3'-protruding, G/C rich sequences
2) We identified four ORFs (154, 171, 249 and 525) in the vicinity of the *cos* region. It was proved that ORF249, ORF154 és ORF525 are vital for lytic growth, while the ORF171 was not
3) We proved that ORF154 and ORF525 encode the subunits of the terminase enzyme of phage *16-3*
4) In databanks ORF525 showed striking similarity to the terminase large subunits of *Escherichia coli* phage *HK97* and *HK022* (52% identity, 66% similarity), while ORF154 showed 34% identity and 53% similarity to the small terminase subunits of the same *HK* phages. Based on the results of the functional experiments and bioinformatic analysis we renamed ORF154 and ORF525 as *terS* and *terL, respectively*.
5) Gel mobility shift assay indicated that TerS protein can bind to a DNA fragment carrying the *cos* site and the flanking region.

CONCLUSIONS AND SUGGESTIONS

Our experiments aimed at the identification of the key components of the *16-3* bacteriophage DNA packaging process. We have identified the cohesive ends and the terminase genes (*terS* and *terL*) which are essential elements in the formation of the new phage population.

We have determined that the chromosome of temperate phage *16-3* of *Rhizobium meliloti* 41 has 10-base long, 3' protruding cohesive ends with complementary G/C rich (90%) sequences. Four open reading frames (two from each side of *cos*) have been tested by experimental and bioinformatic analyses to identify genes encoding the subunits for a terminase enzyme, required for cutting *cos* and producing unit length phage DNA. Three of the candidate genes (ORF249, ORF154 and ORF525) turned out to be vital for phage production. Knock-out of any of these three genes resulted in the loss of *cos* cleavage, while products of these three genes supplemented

from plasmids restored the missing functions.

ORF249 indicates the longest open-reading frame, although use of alternative start codons (to ATG) would result in shorter products (ORF215 and ORF197) in the same frame. Insertion of the Km cassette at codon position 60 (counting according to ORF249) disrupts even the shortest possible ORF. To mark the gene we prefer to use ORF249 since homology with homing endonuclease sequences starts with its first codon. Determination of the role of the protein encoded by ORF249 requires further studies. If the protein acts as an endonuclease, it could have a direct role in DNA cleavage at the *co*s site, alone or complexed with other terminase subunits as a part of the packaging machinery. It has been shown that lambda phages with an amber mutation in any head gene or in FI, the gene encoding the accessory packaging protein gpFI, results in the loss of *cos* cutting and phage production. The product of ORF249 could therefore be an accessory packaging protein.

Although ORF171 did not seem to be essential for *cos* cleavage and phage production, there are also alternative translational start sites resulting in ORF165 or

ORF76. Insertion of a Km cassette at codon position 56 (counting according to ORF171) ruled out the role of ORF171 and ORF165 but not ORF76. Homologies found with other known sequences (holin proteins of different phages) fall into the region of ORF76. We have tested a mutant prophage in which the Km cassette replaced almost the entire ORF76 (1-63 codons deleted). The mutation resulted in no *cos* cutting and no phage production (data not shown). Since phage production could not be complemented by ORF76 supplemented in *trans*, it is likely that the changes disrupted a *cis* element involved in *cos* cutting and/or in phage production.

In databanks we found genes and proteins of various phages which showed extensive homologies with ORF154 and ORF525, and from the alignment we concluded that these ORFs encode the subunits of the terminase enzyme of phage *16-3*. The highest homologies were found with the terminase of phage *HK97* and *HK022*. Indeed these phages also have 10 bp long, 3' protruding cohesive ends. However, the sequence of the cohesive ends of phage *16-3* and *HK* phages are not similar. We renamed ORF154 and ORF525 as genes *terS* and *terL*, respectively.

PUBLICATIONS

Articles:

Ganyu, A., Csiszovszki, Z., Ponyi, T., Kern, A., Buzas, Z., Orosz, L. and Papp, P. P. (2005). Identification of cohesive ends and genes encoding the terminase of phage *16-3. Journal of Bacteriology* 187, 2526–2531.

Ferenczi, S., **Ganyu, A.**, Blaha, B., Semsey, S., Nagy, T., Csiszovszki, Z., Orosz, L. and Papp, P. P. (2004). Integrative Plasmid Vector for Constructing Single-Copy Reporter Systems to Study Gene Regulation in *Rhizobium meliloti* and Related Species. *Plasmid* 52, 57-62.

Posters at international conferences

Ganyu, A., L. Orosz and P. P. Papp (2003). Identification of the cohesive ends and the genes encoding the terminase of phage *16-3*. *American Society for Microbiology 103rd General Meeting*, Washington DC, p. 393

Ganyu, A., Z. Csiszovszki, L. Orosz and P. P. Papp (2003). Identification of the cohesive ends and the genes encoding the terminase of phage *16-3*. *The 2003 Molecular Genetics of Bacteria and Phages Meeting*. University of Wisconsin-Madison, Madison, Wisconsin, p. 282

Oral and poster presentations at national conferences:

Ganyu A., Orosz L., Papp P. (2003). Kulcsfontosságú elemek azonosítása a *16-3* fág DNS pakolásában. Poszter. *V. Magyar Genetikai Kongresszus*, Siófok

Ganyu A., Orosz L., Papp P. (2002). A *16-3* fág ragadós végeinek és termináz enzimének azonosítása (2002). Előadás. *MBK Napok*, Gödöllő

Ganyu A., Orosz L., Papp P. (2002). A *16-3* fág DNS pakolásában résztvevő főbb elemek: A *cos* régió és a termináz enzim azonosítása. Poszter. *Magyar Biokémiai Egyesület Molekuláris Biológiai szakosztálya 7. Munkaértekezlete*, Keszthely

Ganyu A., Orosz L., Papp P. (2000). A *16-3* fág ragadós végeinek meghatározása és a termináz enzim azonosítását célzó kísérletek. Poszter. *Magyar Biokémiai Egyesület Molekuláris Biológiai szakosztálya 5. Munkaértekezlete*, Sopron

Ganyu A., Orosz L., Papp P. (1999). A *16-3* fág cos régiójának szekvencia-szintű vizsgálata és a termináz enzim azonosítását célzó kísérletek. Poszter. *IV. Magyar Genetikai Kongresszus*, Siófok