INTRODUCTION

Molecular mechanisms of recognition processes between biological structures in symbiotic or pathogen interactions were always in the center of scientific interest. The *S. meliloti* – *Medicago* symbiosis is an important model for endosymbiotic nitrogen fixation. Phage *16-3* is a temperate double stranded DNA phage of *S. meliloti* strain 41. It is the far best studied rhizobiophage that serves as a tool in rhizobium genetics, in isolation of some symbiotic mutants and in construction of special vectors. Genetic determinants and molecular mechanisms of many aspects of the *16-3* life cycle have been examined in detail, such as phage integration and excision, regulation of the lytic/lysogenic switch, superinfection immunity system and phage DNA packaging. Moreover, the complete 60 kb phage genome sequence has been determined recently (AC: DQ500118). However, little is known about the genes and structural elements involved in the interaction between the phage and its host, and only one study have been reported on the *16-3* virion proteins.

The initial interaction between a tailed phage and its bacterial host cell is mediated by the distal part of the phage tail, which specifically binds to the phage receptor located on the host surface. Earlier results demonstrated that phage *16-3* adsorption is connected to the strain specific capsular polysaccharide of *S. meliloti* 41, the K_R5 antigen. So far, three bacterial gene clusters involved in K_R5 antigen production have been described, such as *rkp-1*, *rkp-2* and *rkp-3* regions. The *rkp* mutants are also defective in the invasion of the host plants in symbiosis. In addition, they could not adsorb phage *16-3*, suggesting that the K_R5 antigen is required for both functions.

AIMS

In order to elucidate the molecular mechanism of phage *16-3* and *S. meliloti* 41 recognition, the first aim of this study was to isolate bacterial mutants carrying an altered phage receptor and host range phage mutants that were able to overcome the adsorption block. With the help of these bacterium and phage mutants we were intend to identify genetic determinants involved in *16-3* phage infection. This was an appropriate method to investigate recognition processes between the phage and bacterium. Furhermore, our pourpose was to analyse a gene region of the phage chromosome supposed to responsible for tail formation.

MATERIALS AND METHODS

Bacterial strains, bacteriophages, plasmids and growth conditions

Escherichia coli strains XLI-Blue, DH5α and WA321 were used for cloning procedures. *S. meliloti* strain 41 (Rm41) was used for bacteriophage propagation and genetic experiments. The *S. meliloti* derivatives were propagated on complete TA and minimal GTS media at 28°C. The *E. coli* strains were propagated on LB media at 37°C. The appropriate antibiotic was applied to control the strains or to select for transconjugants or transformants. Phage strains *16-3*ΔNC and *16-3cti3* were used as a background for the isolation of host range and insertional mutants, respectively. Plasmids pBluescriptII SK(+), pBBR1MCS-5 and pPAG160 were used for cloning. In triparental matings helper plasmid pRK2013 was used for transfer of pBBR1MCS-5 and pLAFR1 cosmid clones, and pCU101 for the transfer of pPAG160 derivatives.

Isolation of receptor mutant bacteria and host range phage mutants and identification of the mutant genes

Among more than hundred spontaneous phage resistant mutants we screened for those bacteria that presented an altered receptor and therefore allow the isolation of host range phage mutants. When a wild type phage population was titered on the phage resistant bacteria, on a few strains host range phage mutants appeared (with a frequency of 10^{-5} to 10^{-6}), suggesting that the phage receptor of these mutant bacteria had been modified but were still present on the surface. Host range phages appeared were supposed to carry a mutation in genes necessary for host recognition.

We have localized the bacterial mutations in complementation experiments. Cosmid clones carrying either wild-type or mutant *rkp* regions (with different Tn*5* insertions) were introduced by conjugation into the receptor mutants and the transconjugant strains were tested for phage-resistance/sensitivity by the wild-type *16-3* phage.

Temperature sensitive host range mutations (*h109* and *h843*) were localized by marker rescue analysis. Mutants produced minute, hardly visible plaques while wild type phages formed normal plaques at 37 ^{o}C. In these experiments mutant phages were propagated by single-step growth on *S. meliloti* 41 transconjugants that harbored different fragments of the *16-3* chromosome. Phage progeny were incubated at restrictive temperature (37 ^{o}C) on Rm41 lawn in order to detect wild type recombinants (and revertants). The frequency of reversion was estimated by propagating the mutant phages on Rm41 harboring empty vectors. When wild type phages appeared at least two orders of magnitude more frequently than in the

control population (recombinants versus revertants) the cloned fragment was considered to carry the wild type sequence of the mutated region.

Alleles encoded for either an altered receptor in *S. meliloti* or an altered host range phenotype in *16-3* phage, were amplified by PCR reactions and their DNA sequence were determined.

DNA procedures, sequence determination and bioinformatics

Basic DNA manipulation procedures including DNA isolation, restriction enzyme digestion, agarose gel electrophoresis, ligation and transformation of *E. coli* were performed according to standard techniques or as recommended by the suppliers. DNA sequence determination was done with the BigDye terminator kit on an Applied Biosystems 373A sequencer (Perkin Elmer, Wellesley, MA USA).

For bioinformatic analysis of the predicted proteins the following websites were used: BLAST server (blast.ncbi.nlm.nih.gov), HHpred server (toolkit.tuebingen.mpg.de/hhpred), BetaWrap prediction (groups.csail.mit.edu/cb/betawrap/).

Temperature shift experiments

Cultures and lysates were warmed up to the restrictive temperature (37 ^{o}C) and approximately 10^5 PFU of *h109* phage particles were added to $2x10^8$ CFU Rm41 bacteria. Cultures were shaken at restrictive temperature and were cooled down to the permissive temperature (25 ^{o}C) at different time points. Phages were propagated altogether for 180 minutes to complete the life cycle. The number of phage progeny was compared to control lysate which was grown for the 180 minutes at 25 ^{o}C.

Construction of insertional mutants of phage *16-3*

DNA fragments of the *16-3* phage genome were cloned into pBluescriptII SK(+) vector and were used for insertional mutagenesis. Mutations were constructed *in vitro* by MuA transposase using the kanamycin resistant entranceposon F779 (TGSII Kit, FINNZYMES, Espoo, Finland). The approximately positions of each insertions were determined by PCR. The correct positions of the selected insertions were determined by DNA sequencing.

DNA fragments carrying the insertions were recloned into plasmid pPAG160 harboring a spectinomycin resistance marker. These plasmid derivatives were introduced by triparental mating into a lysogenic Rm41 strain harboring a temperature inducible prophage (*16-3cti3*). Since pPAG160 is unable to replicate in *S. meliloti*, kanamycin resistant but spectinomycin sensitive colonies should have acquired the mutation by homologous recombination. The

absence of the vector sequences were tested by PCR using primer pairs specific for the pPAG160 sequences. Integration of the mutations were verified by PCR using primers specific for the entranceposon and for the target site. Excision of prophage mutants were induced at 37 ^{o}C for 30 min, and phage maturation was completed by incubation of the cultures for additional 180 minutes at 28 ^{o}C.

Purification of phage particles and protein techniques

Purification of bacteriophages by cesium-chloride step gradient centrifugation was performed by conventional method. Purified phage samples were dialyzed against SM buffer.

Protein separation was carried out in NuPage gel electrophoresis system according to the Invitrogen protocols (Invitrogen Co Carlsbad, CA). Purified, dialyzed phage samples were loaded into 4 to 12% NuPage Novex Bis-Tris polyacrylamide gels. Electrophoresis was performed at 200V constant voltage. Protein bands were visualized by SimpleBlue^{TM} staining. Western blots were performed with phage proteins separated by electrophoresis and subsequently electroblotted onto polyvinylidene difluoride membranes using an Xcell II Blot module and NuPage transfer buffer. N-terminal amino acid sequencing was performed with an Applied Biosystems (model 471) protein sequencer using the Edman degradation chemistry.

Electron microscopy methods

Formvar-coated copper grids (300 or 400 mesh) were floated on a drop of CsCl-purified, dialyzed phage samples for 5 minutes, excess fluid was carefully removed with the edge of a filter paper, and the preparations were negatively stained with 3% (wt/vol) phosphotungstic acid (PTA) solution, pH 6.6 for 40 seconds. Excess PTA was removed and the remainder was left to air dry. The preparations were examined on a JEOL 1200 EXII transmission electron microscope.

Deoxycholate-polyacrylamide gel electrophoresis (DOC-PAGE)

Isolation of the surface polysaccharides (KPS, LPS) were carried out by a modified hot phenol-water extraction method. The samples were dialyzed against distilled water and were dried by lyophilization, finally were solved in distilled water. The phenol-water extracted materials from bacteria were separated in 18% DOC-PAGE (BioRad Miniprotean II) and were visualized using Alcian blue silver staining. Photos were taken by a UVP BioDoc-It System.

RESULTS AND DISCUSSION

IDENTIFICATION OF PHAGE AND BACTERIUM GENES INVOLVED IN THE PHAGE INFECTION

To investigate phage-host recognition, receptor mutant bacteria (GH4046, PP4073, GH4180) and host range phage mutants (*h5, h105, h842, h182, h109, h843*) were isolated and investigated (for details of isolation see in Materials and Methods; the summary of mutants see in Table 1.).

TABLE 1. Characterised phage and bacterium strains

Phage strain	Phage genotype or mutated gene (amino acid change)	Genotype of bacteria and phage propagation efficiency				
		Rm41 (wild type)	**GH4046** *rkpM* L252F	**PP4073** *rkpY* L552P	**GH4180** *rkpZ*	**AT313** *rkpZ*::Tn5
16-3	Wild type	++	-	-	+	+
h5	*hI* G588D [(a)]	++	+ *	+	++	++
h105	*hI* G588D [(a)]	++	+	+	++ *	++
h842	*hI* G588V [(a)]	++	++	++ *	++	++
h182 [ts]	*hII* N666K [(b)]	++	-	+	++	++ *
h109 [ts]	*hII* D783N [(b)]	++	-	+	++ *	++
h843	*hII* D783G [(b)]	++	-	+ *	++	++

ts) temperature sensitive mutant;
a) *hI* (*ORF022*) is 2112 bp long, mutations affected the GGC codon from 1762 to 1764 bp. Mutants have GAC or GTC triplet;
b) *hII* (*ORF023*) is 2412 bp long, mutations affected either the AAC codon from 1996 to 1998 bp, and mutant *h182* has an AAA triplet or the codon GAC from 2347 to 2349 bp. Mutants *h109* and *h843* have AAC and GGC codons, respectively;
*) the host range mutant was isolated on bacteria of the indicated strain; ++ large, sharp-contoured plaques, with spots of confluent lysis; + small plaques with pale spots; – no plaques formation occured.

DOC-PAGE experiments showed that the phage receptor mutant bacteria affected in the K_R5 antigen production: while PP4073 and GH4180 mutants possess an altered capsule, strain GH4046 produce no detectable K_R5 antigen, suggesting that K_R5 antigen is not or not obligate part of the phage receptor.

Receptor mutations of *S. meliloti 41* affects the *rkpM*, *rkpY* and *rkpZ* genes

In order to determine which *rkp* genes harbors mutations in the receptor mutant bacteria, genetic complementation experiments were carried out by introducing the known *rkp* genes on different cosmid clones. According to these results, mutations were localized in genes *rkpM* (GH4046), r*kpZ* (GH4180) and *rkpY* (PP4073).

Since Tn5 transposon mutants of the *rkpM* and *rkpY* genes were unable to bind phage *16-3* and were not suitable for the isolation of host range phage mutants, we concluded that $rkpM_{4046}$ and $rkpY_{4073}$ are special alleles that blocked the infection of the wild type phage but

allowed the isolation of host range mutants. Because of the special features of the above mutations, DNA sequences of the mutant alleles were determined. In both mutants, one missense mutation was found resulting a Leu_{252} to Phe_{252} substitution in *rkpM*$_{4046}$ allele and a Leu_{552} to Pro_{552} substitution in *rkpY*$_{4073}$ allele, respectively. We suggest that the RkpM and the RkpY proteins take part in the formation of the phage *16-3* receptor.

A third host gene, *rkpZ* also influences the *16-3* phage infection, but contrary to RkpM and RkpY the role of RkpZ in phage adsorption/infection may be indirect, since Tn5 insertional mutant (strain AT313) as well as spontaneous mutant (strain GH4180) was both suitable for the isolation of host range phage mutants. It is likely that RkpZ protein does not take part in *16-3* receptor formation but influences phage adsorption through its effect on capsular polysaccharide production (*rkpZ* mutants show chain length modification of K_R5 antigen).

hI* and *hII* genes represent host range loci of phage *16-3

Earlier a host range mutation was localized in the late gene region of phage *16-3* by marker rescue experiments. We have established the DNA sequence of this region and an 2112 bp long open reading frame (ORF) was identified representing the putative *h* (host) gene. In order to prove that this ORF is involved in phage receptor recognition, the nucleotide sequence of host range alleles were also established. Three host range mutants (*h5, h105 and h842*) carried missense mutations on the 3' region of the putative *h* gene. In all cases the wild type GGC codon encoding the Gly_{588} was altered resulting either a GAC (Asp) or a GTC (Val) triplets (Table 1.).

Since some host range mutations (*h182, h109, h843*) could not be localized within the *h* gene we supposed that beside the *h* gene, there is at least one additional gene influencing host recognition. To delimit the location of temperature sensitive *h109* mutation on the physical map of the phage genome marker rescue experiments were carried out. The shortest DNA fragment that resulted in wild type recombinants was the 1,5 kb long *Eco*RI (L) fragment. This region covers the majority of the 2412 bp long coding region, *ORF023*, that is located downstream from the *h* gene. After PCR amplification nucleotide sequence of the *h109* allele was determined. One missense mutation (GAC to AAC) resulting substitution Asp_{783} to Asn_{783} was detected in the coding frame of the putative protein. Missense mutations *h182* (AAC to AAA, resulting Asn_{666} to Lys_{666} substitution) and *h843* (GAC to GGC, resulting Asp_{783} to Gly_{783} substitution) were also identified within *ORF023* (Table 1). This new host range locus was designated as *hII*. The original *h* locus is hereafter referred to as *hI*.

hII gene encodes a tail protein

hII locus was identified by host range mutations as a functional gene. The 803 amino acid long putative HII protein showed no significant homology to any other proteins with known function using the BLAST server. However, by the HHpred tool a well defined part of the HII protein (from amino acid residues 300 to the end) shows strong homology to sugar interacting proteins. All of these proteins display a parallel beta-helical structure. This fold seems to be the feature of the HII as well, indicating that it may interact with the host surface capsular polysaccharide.

Predicted molecular mass of HII is about 85 kDa. A protein with similar molecular mass was identified earlier as a tail component of phage *16-3* in purified samples. To determine whether the detected protein corresponds to HII, phage proteins were separated by SDS-PAGE and the N-terminal sequence of the appropriate band was determined. The resulted sequence (AITAAEAFRDY) showed the same amino acid sequence that was encoded by the 5' part of the predicted *hII* gene.

To get more information about the role of protein HII, temperature shift experiments were carried out using host range phage *h109* harboring a temperature-sensitive allele of *hII*. We were curious to know when the restrictive temperature (37 ^{o}C) affects the propagation of the mutant. We found that this mutant was not sensitive to the restrictive temperature in the first 120 minutes of the phage life cycle including the adsorption of the phage particles to the host cells. When maturated *h109* phage particles were incubated at 37 ^{o}C, no effect of the elevated temperature was detected suggesting, that the assembled tail structure of the mutant phage is not temperature sensitive any more. These results indicate a role for HII protein in a late event in virus morphogenesis (probably in assembly).

IDENTIFICATION OF FUNCTIONAL GENES IN THE PHAGE TAIL REGION

Complete sequence of the *16-3* phage genome was deposited to the nucleotide databases recently (AC: DQ500118). Upstream of genes *hI* and *hII* additional five genes were predicted to encode proteins for the tail structure (from *ORF017* to *ORF021*). In order to identify tail structural and tail assembly protein genes this region was selected for directed insertional mutagenesis. Phage *16-3* mutants were obtained by introducing insertions with a kanamycin resistant marker into each of the selected ORFs by homologous recombination. For this purpose a lysogenic Rm41 strain harboring a thermo-inducible prophage (*16-3cti3*) was used. Into this background derivatives of pPAG160 vector carrying insertions in the putative genes were introduced.

In order to determine whether insertions influence phage assembly and infection efficiency, lysogen cultures were induced by heat shock. The lysates were titered on strain Rm41 and were examined by electron microscopy. We could detect only a very few plaque forming units ($<10^3$ pfu/ml versus 10^{10} pfu/ml wild phage yield) from lysates of prophages carrying most of the insertions. We proved that these infectable phage particles were revertants. Thus the above mutations blocked phage maturation.

Particles were collected by CsCl step gradient and the purified phages were examined by electron microscopy. Analysis of the wild type phage showed that the *16-3* virion has an isometric head with 55 nm diameter and a 97 nm long flexible, noncontractile tail ending in a baseplate with six club-shaped spikes. These features are characteristic for *Siphoviridae* family. Insertional mutations in ORFs *017, 018a, 020, 021, hI* and *hII* resulted in head structures only. Therefore we concluded that all of these ORFs represent functional genes essential for the tail formation. In the case of insertion in *ORF018* no virions were detected in lysates, therefore gp018 must be essential for phage head formation.

All of the above mutations could be complemented by the DNA fragment that carries only the given ORF, suggesting that they represent functional genes essential for phage viability.

A MODEL OF THE *16-3* PHAGE INFECTION

All of the *16-3* host range mutations isolated resulted in amino acid residue substitutions in protein HI and HII. Host range mutant phages affected in genes *hI* and *hII* show different host specificities. Strain GH4046, an *rkpM* bacterial mutant, was a suitable host for *hI* mutants but not for *hII* mutants. In contrast, both *hI* and *hII* mutant phages could be isolated on *rkpZ* (strain GH4180) and *rkpY* (strain PP4073) receptor mutants. We suppose that this difference is derived from the remarkable different capsular polysaccharide surfaces of the bacterial mutants. Both *rkpZ* and *rkpY* mutants possess an altered capsule, while *rkpM* mutants produce no detectable K_R5 antigen. The *rkpM* mutant adsorbed poorly the wild-type phages (below 50%) while *rkpZ* and *rkpY* mutants adsorbed 95% of the phages in 5 minutes. Based on these data, it is likely that K_R5 antigen is involved in an initial polysaccharide-controlled phage binding, which is followed by a secondary binding step that is protein (RkpM and RkpY) dependent. Presumably, *hII* host range mutants penetrate more efficiently into the altered *S. meliloti* 41 capsule than into the wild-type capsule (and reach the proteinaceous receptor), while host range mutations in *hI* result in stronger interactions between proteins involved in recognition.

SUMMARY OF THE RESULTS

In this work we have:

- isolated spontaneous bacterial mutants carrying an altered phage receptor;
- isolated spontaneous host range phage mutants which adapt to the altered bacterial receptors;
- proven that the *16-3* phage receptor has a proteinaceous part in *S. meliloti* 41 and that the bacterial RkpM and RkpY proteins are components of the phage receptor;
- shown that *rkpZ* gene influences phage infection indirectly;
- identified HI and HII proteins of phage *16-3* as structures necessary for host recognition;
- provided direct evidence that protein HII is present in the phage particle;
- built a two step model for *16-3* phage infection where an initial binding is controlled by the HII phage protein and the K_R5 antigen of the cell surface. This is followed by a second protein-protein interaction where host proteins RkpM and RkpY and the HI protein of the phage take part;
- described morphology of the *16-3* phage;
- identified seven new genes essential for phage assembly.

PUBLICATIONS are in the hungarian version.

The dissertation is public on: http://www.ttk.pte.hu/biologia/phd/phdfok.htm

UNIVERSITY OF PÉCS

Biological Doctoral School

Identification of molecular elements involved in the *16-3* bacteriophage and *Sinorhizobium meliloti* 41 recognition

Ph.D. thesis

Veronika Deák

Supervisor:
Dr. Péter Putnoky
DSc.

PÉCS, 2010.