SZENT ISTVÁN UNIVERSITY

REGULATORY REGIONS OF *16-3* PHAGE: REPRESSORS AND OPERATORS

PhD thesis

Csiszovszki Zsolt

Gödöllő
2003

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

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

……
Dr. Tuba Zoltán
Approved by the head of school

……
Dr. Orosz László
Approved by the supervisor

BACKGROUND, AIMS

After infection the temperate phages are able to integrate into chromosome of the host bacterial strain and replicate. A bacterium carrying a prophage is immune to superinfection by a homoimmune phage. Superinfection immunity is due to the binding of the prophage-encoded repressor protein(s) to specific operator sites of the phage chromosomes. The repressor prevents vegetative phage development by acting on both the resident prophage and the newly injected DNA of a homoimmune phage. Although, the repressor systems may differ in specificity, topology and position of the relevant genes on the genetic map, two major types are emerging among temperate phages investigated by genetic analyses: phages with one immunity control region (canonized by coliphage λ), phages with dual control regions (the archetype is *Salmonella* phage *P22*). Both types have been described for phages of either Gram positive or negative bacteria. The highly complex *P22 immI* region encodes an antagonist (Ant) of the C2 immunity repressor and proteins that control *ant* expression (*mnt*, *arc*). Even more complex immunity system is represented by phage *P1* where three immunity regions function in a complex network.

The superinfection immunity system of *Rhizobium meliloti* temperate phage *16-3* was mapped to three distinct regions: the *immC*, the *immX* and the *avirT* regions. The genetic anatomy and function of the *immC* region are well documented. The *immC* region codes for a ”typical” repressor, by cistron *c* with operators at its flanks. Mutations in the *c* cistron lead to clear plaques. In its sequence specific DNA binding, the *16-3* C repressor utilizes a helix-turn-helix operator recognition motif with significant homology (and partial cross functionality) to the helix-turn-helix motif of the CI repressor of coliphage *434*. The cognate operators are also of *434* type operators.

Little is known of the other two elements *immX* and *avirT*. In the host cell the *immC* and *immX* regions together provide very high level immunity to superinfection of the homologous phages. Phage titers drop 10^7 to 10^9 magnitudes. In itself *immX* or *immC* still inhibit phage growth drastically, 10^5 to 10^6-fold. Like *immC*, *immX* also

expresses a repressor activity, but no mutant phenotype has been identified for it. The *avirT* locus has been localized close to the early genes. Its function is not known. One mutant allele $avirT_{1\text{-}9}$, which has been studied in some detail, gives to the superinfecting phage full escape from *immX* immunity and in lesser degree escape from *immC* immunity.

According to these I aimed to examined the immunity complexity of *16-3* phage: the *immC* region and primarily the *immX* region. We focused on the specification of the structure of the *immX* region, so the regulatory *trans* (coding repressors and other molecules) and *cis* elements (active DNA surfaces for example operators, promoters). It seemed to be important to describe the interaction between *immX* and *immC*.

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, superinfection immunity test, marker rescue analysis, phage crosses, constructing mutant phages) 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, hybridization techniques, sequencing), total protein analysis were carried out following techniques commonly used in molecular biology as well as by following guidelines provided by the suppliers.

RESULTS

Mapping and cloning of the repressor function of the immX gene

Earlier studies localized the *immX* gene in the *Eco*RI L and H fragments of the *16-3* chromosome. Setting out from this premise, the position of the gene was narrowed down to a 442 bp region, which included the *Eco*RI(41) site. The 442 bp region provided complete ImmX function from *trans* in the carrier cells, as shown by a 5-7 order of magnitude reduction in the efficiency of plating (e.o.p.) for the superinfecting *16-3* strains. Sequencing of this locus revealed two overlapping ORF-s, 116 and 127 putative codons. The two ORFs were in the same frame with opposite directions.

Gene for ImmX function is assigned to two overlapping cistrons

Products encoded by 442 bp region (in *trans*) provided ImmX function while a 4-bp insertion in ORF116/127 eliminated the above function. In order to determine the active product(s), *R. meliloti* 41 strains carrying this region were tested for protein expression. Two proteins expressed from this region and were identified by SDS-PAGE and confirmed by N-terminal sequencing. The proteins corresponded to ORF116 and ORF127. The ORF127 and the corresponding protein followed the general rule (i.e., ATG for start), while the 116 amino acids long protein and the cognate ORF (ORF116) started at CTG (Leu) codon frequently utilized in rhizobia. No proteins were detected from the 4 bp inserted allele.

In order to provide further evidence of the action of the two proteins, various mutations were introduced at specific sites in the two ORFs. The ORF127 was destroyed by a nonsense (opal) mutation at codon 45 (arginine→stop), while this change led to a synonymous codon in ORF116 (serine, TCG→TCA). The ORF116 was destroyed by an amber mutation at codon 10 (glutamic acid→stop), while this mutation led to a synonymous codon in ORF127 (leucine, CTC→CTA). Both ORF127 and 116 were destroyed by a 4-bp insertion leading to a frameshift mutation. Since the ImmX function was inactivated by these changes we concluded that two

cistrons (and two proteins) were involved in ImmX activity: X_U and X_L (corresponding to ORF116 and ORF127, respectively). The results from the protein expressions are fully in agreement with this mutation analyses confirming the existence of two products. It is worthy to note that no homologous sequences for X_U and X_L have appeared so far in databanks EMBL, SWISSPROT, and PDB.

Km^r transducing phage mutants in which gene immX is knocked out

The Km^r cassettes have been inserted at the non-overlapping regions of X_U and X_L in order to disrupt the action of X_U or X_L. These allelic variants were then transferred to the *16-3* chromosome, resulting in Km^r transducing phage mutants (*16-3cti3Km^R-U463* and *16-3cti3Km^R-L469*). Similarly, a third mutant allele of *immX* was also constructed, but the Km^r cassette was inserted in the overlapping region of X_U and X_L, destroying both cistrons (*16-3cti3Km^{R}6-1*).

These mutations were not lethal since the phages formed normal turbid plaques on *R. meliloti* 41 with somewhat different appearance compared to *16-3cti3*. Upon lysogenisation, the three phages were able to integrate into the bacterial chromosome via the *16-3* int/att integrative recombination pathway. Although, the *immC* region was intact in the *immX* mutants, the lysogenic *R. meliloti* 41 derivatives which carried them as prophage have shown sensitivity to superinfection of homoimmune phages. Phage like *16-3cti3* plated with high e.o.p. in *R. meliloti* 41 lysogenic for *immX* mutant phages. Furthermore, these *immX* mutants were not immunity insensitive, i.e., they did not grow on either in *R. meliloti* 41(*16-3cti3*) lysogenic strain, or on *R. meliloti* 41(*immX*).

Genetic complementation between mutants X_U and X_L

Our conclusion that ImmX function was due to two cistrons, X_U and X_L,was probed by genetic complementation tests with the Km^r transducing ("*immX* knock out") phage mutants carrying *Km^R-U463*, *Km^R-L469*, *Km^{R}6-1* insertion and $X_{U/L}$ wild type alleles. Two different double lysogenic *R. meliloti* 41 strains were constructed: *R. meliloti* 41(*16-3cti4Sp4*, *16-3cti3Km^{R}6-1*) for *cis* heterozygotic arrangement of the

X_U and X_L alleles (X_{U+L+}/X_{U-L-}) and *R. meliloti* 41(*16-3cti4Sp4Km*^{R}*-U463, 16-3cti3Km*^{R}*-L469*) for *trans* heterozygotic arrangement (X_{U-L+}/X_{U+L-}). The two double lysogens were tested for superinfection immunity, i.e., for the ImmX function. The *cis* and *trans* heterozygotes were immune to superinfection of *16-3cti3* and *16-3cti4Sp4*, but sensitive to the tester ImmX insensitive *16-3v17-1* strain. Hence, these tests confirmed that the two cistrons (X_U and X_L) were in the genetic background of the ImmX function, since X_U and X_L mutants complemented each other in *trans*, i.e., the ImmX function were completely restored in the *trans* double lysogen.

Mapping of mutations leading to ImmX insensitivity

After proving the involvement of the two cistrons in ImmX function, target elements for the ImmX repressors were searched for. We were especially interested in sites which might be nearby the *immX* genes. These should be other than *avirT*$_{1-9}$ which mapped on the opposite half of the *16-3* chromosome. Since the *immX* genes were tagged physically with the Km^r cassette, consequently the Km^r was to be useful as a genetic marker for them in phage crosses, we hoped that mutations linked to *immX* could be identified at sequence level. We tested the *vir* mutants for growth on strains carrying the cloned *immX* genes or the *immC* gene. Tester bacterial strains were *R. meliloti* 41(*immX*+, in low copy plasmid), *R. meliloti* 41(*immX*+++, in high copy plasmid) for *immX* and *R. meliloti* 41(*immC*) for *immC*. It was observed that *16-3 vir* mutant stocks lysed these bacterial lawns, while wild type *16-3* strains and clear mutant *16-3cti4Sp4* carrying mutation in the *c* cistron did not (in agreement with earlier studies). The transparency and size of plaques and the e.o.p-s varied widely for the *vir* mutants in the presence of the ImmX function.

In order to assign this ImmX sensitivity/insensitivity phenotype to a major region of the *16-3* chromosome, 3-point mappings have been done by crosses between *vir* stocks *16-3v17*, *16-3v17-1* (ImmX insensitive) and *16-3cti3Km*^{R}*6-1* (ImmX sensitive). First, recombinants were scored by Southern hybridization assays for intact *immC* region (DNA marker for *16-3cti3Km*^{R}*6-1* parent) and for the intact

immX gene (i.e., absence of the Km^r cassette, DNA marker for the *vir* parents). Out of 108 randomly chosen progeny in the *16-3v17-1* cross, 14 were intact for both *immC* and for *immX*. The recombination frequency for the *16-3v17* cross was 4/72. The recombinants were then tested for the ImmX sensitivity/insensitivity phenotype (i.e., this is being the nonselective marker). All the recombinants (14/14; 4/4) grew on ImmX expressing strain *R. meliloti* 41(*immX*+) and formed turbid (*cti3*) plaques on *R. meliloti* 41 (i.e., they were intact for *immC* and ImmX insensitive). This outcome was compatible with a linkage of ImmX sensitivity/insensitivity and the left flanking region of the Km^r marker in *16-3cti3Km^{R}6-1*. Furthermore, the manifestation of the ImmX insensitivity phenotype of these recombinants indicates independence from the function of the *immC* region. Mutant pairs, one with intact, the other with deleted *immC* region, plated with equal efficiency in ImmX expressing *R. meliloti* 41 strains.

Results of the phage crosses were confirmed by marker rescue analyses on the *16-3v17* and *16-3v17-1* phage chromosomes: double crossovers between pDH1 (which carried immX region) and chromosomes of the above *vir* stocks resulted in few progeny (for *16-3v17* 6/2500 and for *16-3v17-1* 4/2500) unable to grow on *R. meliloti* 41(*immX*+) (i.e., they were ImmX sensitive).

Combinations of mutations for ImmX insensitivity in the vicinity of the immX genes

Refined mapping of the ImmX insensitivity mutations of *16-3v17* and *16-3v17-1* was carried out by a series of marker rescue experiments where cloned fragments from *16-3v17-1* were probed against the *16-3cti3* phage chromosome. Segments from *16-3v17-1* providing ImmX insensitivity were readily exchanged in the *16-3cti3* chromosome (by double crossing over). The mutations carried by these recombinants were ascribed to a 1302 bp section of the phage chromosome. Sequencing of this section revealed three mutant sites (X_{V1}, X_{V2}, X_{V3}) outside but in the close vicinity of the $X_{U/L}$ cistrons: X_{V1}, X_{V2}, X_{V3} in *16-3v17-1*; X_{V1}, X_{V3} in *16-3v17*. The same chromosomal segment was isolated by PCR from three *vir* stocks of independent origin. Combinations of X_{V1} and X_{V3} sites were identified in the mutants. Strong expression of the ImmX insensitive phenotype was linked to the double

mutation X_{V1}, X_{V3}. Mutation in X_{V2} enhanced the expressivity of X_{V1} and X_{V1}, X_{V3} double mutants.

Novel results

1. $X_{U/L}$ and X_V elements were identified. These elements are unique in the databanks.
2. X_U and X_L are cistrons, both of them need for the function of *immX* repressor. These *trans* elements complement each other.
3. X_{V1}, X_{V2} and X_{V3} are assessed to be *cis* elements, and hierarchical order of their phenotypes are determined.
4. *immX* (*Eco*RI L and H) region is bioinformatically examined.
5. *immC* is unnecessary for the function of *immX*.
6. We determined the epistatical relationship between *immX* and *immC*, so *immC* the downstream element (see next section).
7. The Km^r cassette mutagenesis procedure is developed, which opens the possibility to analize complementations and recombinations among phage mutants. This method can be generalized and applied to different genes in theory.

CONCLUSIONS AND SUGGESTIONS

This thesis shed insight into a complex regulatory region of *Rhizobium* phage *16-3*. The *immX* region contributes to the function of immunity to homoimmune phage superinfection. We dissected the *immX* region in two parts, $X_{U/L}$ and X_V. The $X_{U/L}$ region contains two overlapping cistrons, X_U and X_L, which code for proteins pX_U and pX_L, respectively. The X_V region functioned as a cognate target for the two proteins. To date, no homologous DNA or peptide sequences to the *immX* region were found in the databanks (EMBL, SwissProt, PDB).

Mutations identified in the X_V region overcame the $X_{U/L}$ repression. Generally, these mutations could define cognate *cis* element(s). The sequence around the X_{V3} site is comparable to known repressor binding sites. A CG to TA base change at the mutant X_{V3} site weakens the palindromic symmetry of the sequence box 5'-ATGGCCGGGCAT-3' to 5'-ATGGCCGGGTAT-3'. Three copies of this box are separated by 11 and 9 base pairs on the genome. Consequently, proteins bound to the boxes should lie on the same face along the B-DNA also providing opportunity for cooperative interactions. The sequences for X_{V3} (as well as for X_{V1}, see below) are not related to the operator sites O_R and O_L bind the *16-3* C repressor (i.e., 5'-ACAA-4/6 bp-TTGT-3').

The sequence around the X_{V1} site (i.e., 5'-CGACCGATCGCTGTCGTTTTATT-3') can be evaluated in two ways. The first 16 bp segment contains a strong palindromic symmetry like X_{V3} and may also be a protein binding site. However the whole sequence, if transcribed in mRNA (bottom strand) could be folded into a Rho independent transcription termination structure. The mutations found at the 13th (G to T) and 14th (T to C or A) position would weaken both symmetry of a putative binding site as well the stability of a putative stem loop.

Phages with the sole mutation at the X_{V2} site has not been isolated. The overlapping DNA sequence did not provide hint for the function of X_{V2}. Phenotypically, the X_{V2} mutation enhanced the expression of the ImmX insensitivity when added to either a single mutation in X_{V1} or double mutation in $X_{V1}X_{V3}$. The phenotypic change caused by mutant X_{V2} was mild according to the e.o.p. assays, although its effect was very strong for plaque morphology. Mutants with X_{V2} mutation ($X_{V2}X_{V1}$ vs. X_{V1}, and $X_{V3}X_{V2}X_{V1}$ vs. $X_{V3}X_{V1}$) formed significantly bigger and more transparent plaques in the lawn of ImmX expressing bacteria.

Mutations in X_U or X_L destroyed the ImmX controlled immunity function. The immunity function was restored in *R. meliloti* 41*(16-3cti3Km^R-L469*, *16-3cti4Sp4Km^R-U463*) double lysogens, i.e., the X_U and X_L mutant prophages complemented each other in *trans*. Our genetic analyses were compatible with the

fact that both X_U and X_L cistrons were directly involved in the ImmX repression. Whether the cistrons act independently in a simple additive way or cooperatively by formation of complex structures remains open for further studies.

A number of temperate phages possess two repressor regions. One of the regions, often named *immC* (after the clear plaque phenotype rendered to it) shows structural similarity and sequence homology to the *λcI* region. The second immunity regions, here in *16-3* is *immX* (in *P22* is *immI*), seem to be more diverse. The second immunity region of *16-3*, *immX*, is unique. It does not show structural similarity or sequence homology to immunity regions of other phages with dual control regions.

In this study we focused only on the *immX* region of *16-3* and investigated (investing less time and energy) its functional connections with the *immC* region. Plaques of *immX* mutants are turbid (an indication for ImmC activity), while those of *immC* mutants and *immCimmX* mutants are clear. This result is compatible with the ImmC function being epistatic to the ImmX function, that is, according to the rules of epistases may indicate that ImmC acts downstream to ImmX in the pathway toward the development of the lysogenic state. However, our results do not show how the roles for ImmX and ImmC are shared in establishing and maintaining lysogeny. The heat inducibility of *16-3cti3* lysogens (i.e., carrying *immC ts* mutant prophages) and the lack of immunity to superinfection of the lysogens carrying $X_{U/L}$ mutant prophages (i.e., Km^r insertion mutants) show that the functions of *immC* and *immX* are involved in at least maintaining lysogeny.

The *avirT* locus may be a ground where both the C and X repressors interact since mutant *16-3c*$^+$*avirT*$_{1\text{-}9}$ escaped from *immX* and (in lesser extent) from *immC* actions. To date there is no handle for the function of *avirT*. Mutant *16-3c*$^+$*avirT*$_{1\text{-}9}$ formed turbid plaques like the *16-3cti3X*$_V$-s, however unlike these the former made very unstable *R. meliloti* 41 lysogens. The turbid plaques of *16-3c*$^+$*avirT*$_{1\text{-}9}$ cleared up after 2-3 days, its lysogenic broth cultures lysed. From the genetic analyses of the lysogeny of the *16-3* that can be deduced that the ability of *16-3* to overcome efficiently the immunity of the lysogenic cells needs at least the combination of two mutant elements: mutant *immC* combined with mutant X_V or with mutant *avirT*. Our

next efforts are aiming the detailed functional dissection of these two regulatory regions.

PUBLICATIONS

Articles:

Csiszovszki Zs., Zs. Buzás, Sz. Semsey, T. Ponyi, P. P. Papp, and L. Orosz. (2003) *immX* immunity region of *Rhizobium* phage *16-3*: two overlapping cistrons of repressor function. *Journal of Bacteriology*. **185**:4382-4392.

Papp P. P., T. Nagy, Sz. Ferenczi, P. Élő, **Zs. Csiszovszki**, Zs. Buzás, A. Patthy and L. Orosz (2002) Binding sites of different geometries for the *16-3* phage repressor. *Proc. Natl. Acad. Sci.* U.S.A. **99**:8790-8795.

Posters at international conferences

Csiszovszki Zs., Zs. Buzás, Sz. Semsey, P. P. Papp, L. Orosz. (2003) Identification of the *cis* and *trans* regulatory elements in the *immX* immunity region of *Rhizobium meliloti* temperate phage *16-3*. Poster. American Society for Microbiology 103rd General Meeting. Washington, USA, május 18-22. Abstract, 390. o.

Ganyu A., **Z. Csiszovszki**, L. Orosz and P. P. Papp. (2003) Identification of the cohesive ends and genes encoding the terminase of phage *16-3*. Poster. Molecular Genetics of Bacteria and Phages, Madison, WI, USA, augusztus 5-10. Abstract, 282. o.

Presentations at national conferences:

Csiszovszki Zs., Papp P. P., Orosz L. (2003) A *16-3* fág *immX* szabályozó régiója. Előadás. *V. Magyar Genetikai Kongresszus*. Siófok, április 13-15. Összefoglaló, 30. o.

Csiszovszki Zs., Papp P. P., Orosz L. (2002) A *16-3* fág *immX* szabályozó régiója. Előadás. *MBK Napok*. Gödöllő, október 2-3.

Nagy T., Élő P., **Csiszovszki Zs.**, Papp P. P., Orosz L. (1997) A *Rhizobium meliloti 16-3* fágjának represszora szerkezetileg különböző operátorokon szabályoz. Előadás. *A Magyar Biokémiai Egyesület Molekuláris Biológiai Szakosztálya 2. Munkaértekezlete*. Lillafüred, május 13-16. Összefoglaló, 134. o.

Posters at national conferences

Csiszovszki Zs., Papp P. P., Orosz L. (2002) A *16-3* fág immunitás régiójának vizsgálata. Poszter. *A Magyar Biokémiai Egyesület Molekuláris Biológiai Szakosztálya 7. Munkaértekezlete.* Keszthely, május 14-17. Összefoglaló, 122. o.

Csiszovszki Zs., Orosz L., Papp P. P. (2000) A *16-3* fág immunitás régiójának vizsgálata. Poszter. *A Magyar Biokémiai Egyesület Molekuláris Biológiai Szakosztálya 5. Munkaértekezlete*. Sopron, május 8-11. Összefoglaló, 136. o.

Csiszovszki Zs., Orosz L., Papp P. P. (1999) A *16-3* fág immunitás régiójának vizsgálata. Poszter. *IV. Magyar Genetikai Kongresszus*. Siófok, április 11-14. Összefoglaló, 124. o.