Structural Characterization of the *virB* Operon of the Hairy-root-inducing Plasmid pRiA4

Yajie Liang, Takashi Aoyama, and Atsuhiro Oka

The virulence (*vir*) loci of the hairy-root-inducing plasmid pRiA4 are essential for pathogenicity of its host *Agrobacterium rhizogenes* A4. We have now determined the nucleotide sequence of *virB* region (9,926 bp). It was found that the *virB* operon is composed of 11 genes (*virB1* to *virB11*), whose products mostly appear to be associated with the cell membrane. A novel structural characteristic is frequent overlapping between the translation termination and initiation codons of the adjacent genes This is indicative of fine tuning of relative translation frequencies for each VirB protein, supporting the view that VirB multisubunit complexes provide facilities for T-DNA transfer at the bacterial cell membrane.

Keywords: *Agrobacterium*/ Membrane protein/ Plant-microbe interaction/ Virulence regulon

The agropine-type hairy-root-inducing plasmid pRiA4 confers tumorigenic symptoms at wound sites on a wide variety of dicotyledonous plants upon infection by its host bacterium, *Agrobacterium rhizogenes* A4. Tumorigenesis is caused by the transfer of a defined DNA segment (T-DNA) on the plasmid from a bacterium into the plant nuclear genome and subsequent constitutive synthesis of plant phytohormones directed by the T-DNA. T-DNA transfer requires the 25 bp imperfect direct repeats at both extremities of T-DNA as *cis* factors, and the virulence (*vir*) loci outside of T-DNA as *trans* factors. The *vir* genes constitute the six transcriptional units, *virA*, *virB*, *virC*, *virD*, *virE*, and *virG*. Their expression is tightly regulated as a regulon by both the *virA* and *virG* gene products, being inducible by plant phenolic compounds such as acetosyringone. Nucleotide sequences of all of the *vir* loci, except the largest *virB* operon, have been determined by our laboratory, and compared with those of tumor-inducing plasmids harbored by *A. tumefaciens* (for reviews, see ref. 1 and 2). Here, we report the nucleotide sequence of the *virB* operon together with its flanking regions, thus completing sequencing of the entire *vir* loci of pRiA4.

The *vir* loci of pRiA4 have been identified with transposon (Tn*3*-HoHo1) insertion mutations that lead to defects in pathogenicity, and the *vir* gene organization examined from the similarity with that of the tumor-inducing plasmid pTiA6 by hybridization experiments. As a result, it has been found that pRiA4 *virB* is located in a region corresponding to the *Hin*dIII-10, -8 and -32b frag-

MOLECULAR BIOLOGY AND INFORMATION — Molecular Biology —

Scope of research

Attempts have been made to elucidate structure-function relationships of genetic materials and various gene products. The major subjects are mechanisms involved in signal transduction and regulation of gene expression responsive to environmental stimuli, differentiation and development of plant organs, and plant-microbe interaction. As of December 1997, study is being concentrated on the roles of homeo domain proteins, MADS box proteins, and DDK response regulators of higher plants in developmental and signal transduction processes.

Prof
OKA, Atsuhiro
(D Sc)

Assoc Prof
AOYAMA, Takashi
(D Sc)

Instr
GOTO, Koji
(D Sc)

Students:
TSUKUDA, Mayumi (DC)
HOMMA, Takashi (DC)
OHGISHI, Maki (DC)
SAKAI, Hiroe (DC)
LIANG Yajie (MC)
UEDA, Yumi (MC)
YANO, Hiroyuki (MC)
OHASHI, Yohei (MC)

ments. These three fragments were purified from a cosmid clone (pBANK1329) covering the DNA region from the *Hin*dIII-23b to *Hin*dIII-25 fragments (3), and then separately inserted into pUC18. With the resulting recombinant plasmids, progressive deletion derivatives were constructed by either ejecting appropriate restriction fragments or recloning smaller fragments. Each of the inserted DNA segments was sequenced from either side for both strands by an ABI Prism 377 DNA sequencer. In addition, adjacent restriction fragments actually being contiguous was confirmed by sequencing overlapping fragments with appropriate custom primers. By arranging overlapping sequences of both strands, the nucleotide sequence from the *Eco*RI site downstream of *virA* to the *Dra*I site upstream of *virG* was deduced (9,926 bp; see http://molbio.kuicr.kyoto-u.ac.jp/virb).

The occurrence of initiation and termination codons for protein synthesis in all possible reading frames suggests that the *virB* locus carries information to code for 11 polypeptides of 26.4 kDa, 12.3 kDa, 11.7 kDa, 87.7 kDa, 23.1 kDa, 31.8 kDa, 5.9 kDa, 26.2 kDa, 32.2 kDa, 40.3 kDa, and 38.2 kDa. Each of the presumed coding frames was named *virB1* to *virB11*, respectively, all of which were accompanied by a potential ribosome-binding sequence (4). The initiation codon for all *virB* genes except *virB1* was the conventional AUG codon. However, *virB1* appeared to use a rare initiation codon AUC. This was not a result of base substitution occurred during cloning experiments because direct sequencing of the corresponding region of pRiA4 DNA gave an identical result. The transcription start site is located 65 bp upstream of the *virB1* translation initiation site (5). There are helically phased vir box sequences essential for VirG binding and the promoter -10 region, though no typical promoter -35 region is present (5). These structural features strongly support all of *virB1* to *virB11* being actually structural genes and constituting an operon.

A novel structural characteristic of the *virB* operon is frequent overlappings of the translation termination codon of the preceding gene with the translation initiation codon of the following gene. Those from *virB1* to *virB4* overlap for 1 bp, and those from *virB8* to *virB10* do for 2 bp. Furthermore, in case of *virB7* and *virB8*, the coding region itself (11 bp) is overlapped with each other. Such kinds of overlapping for compacting have never been found in other *vir* operons.

Computer analyses of amino-acid sequences of each VirB suggest the presence of N-terminal signal sequences in VirB1, VirB5, VirB7, and VirB9, and of the transmembrane regions in VirB2, VirB3, VirB6, VirB8, VirB9, and VirB10. Since the smallest VirB7 protein contains a putative target sequence for signal peptidase II, the enzyme involved in lipoprotein recognition and processing, it appears to associate with lipid. Besides, VirB4 and VirB11 are presumed to be the ATP-binding proteins.

It is well known that *vir* regions of pRi and pTi are exchangeable without loss or alteration of the pathogenic specificity, and DNA fragments carrying *vir* of a given plasmid hybridize with those of other plasmids (3). Therefore, the pRiA4 *virB* operon should structurally be close to those of pTi's, as already shown for the other *vir* genes of pRiA4 (3). Indeed, both the gene organization in the pRiA4 *virB* operon and the amino-acid sequence of each VirB protein were exceedingly similar to those of pTi's at the nucleotide sequence level. The identities of amino-acid residues between pRiA4 and pTiC58 VirB's and between pRiA4 and pTiA6 VirB's occurred at 90%-95% and at 70%-92% residues, respectively. Also the overlapping feature is completely identical with that of pTiC58 and close to that of pTiA6. Conservation of the overlapping feature seems to reflect the control of translation initiation frequency relative to that of the preceding gene, providing the exact molar ratio of synthesized VirB proteins. This view supports the idea that VirB proteins make multisubunit complexes at the bacterial cell membrane, presumably giving facilities for the passage of T-DNA from a bacterium to a plant cell.

The similarity scores of VirB between pRiA4 and pTiC58 (compatible combination) are significantly higher than those between pTiA6 and pTiC58 (incompatible combination), whereas the similarity scores between pTiA6 and either pRiA4 or pTiC58 are comparable. These are consistent with the view that the *vir* regions of the three plasmids have evolved from a common ancestral *vir* gene set from which pTiA6 *vir* was initially separated. Since plasmids belonging to a single incompatibility group are usually more cognate to each other than those involved in different incompatibility groups, functional domains of pRi and pTi, at least those determining virulence on plants (*vir* and T-DNA) and autonomous replication and incompatibility (*ori*/*inc*), appear to have been shuffled during evolution of these plasmids.

References

1. Hooykaas PJJ, *Plant Mol. Biol.*, **13**, 327-336 (1989).
2. Oka A, *Tanpakushitsu-Kakusan-Kouso*, **40**, 1010-1021 (1995).
3. Hirayama T, Muranaka T, Ohkawa H, and Oka A, *Mol. Gen. Genet.*, **213**, 229-237 (1988).
4. Shine J and Dalgarno L, *Proc. Natl. Acad. Sci. USA*, **71**, 1342-1346 (1974).
5. Aoyama T, Takanami M, and Oka A, *Nucl. Acids Res.*, **17**, 8711-8725 (1989).