

Characterization of the *virA* gene of the agropine-type plasmid pRiA4 of *Agrobacterium rhizogenes*

Hideki Endoh, Takashi Hirayama, Takashi Aoyama and Atsuhiko Oka

Laboratory of Molecular Biology, Institute for Chemical Research, Kyoto University, Uji-Shi, Kyoto-Fu 611, Japan

Received 25 June 1990; revised version received 31 July 1990

We sequenced a 4.2-kb DNA region encompassing the *virA* locus of the hairy-root-inducing plasmid pRiA4, and compared its sequence with the published *virA* region sequences of four tumor-inducing plasmids. An open reading frame capable of coding for 829 amino acids was identified for *virA*. Deletion mutants of *virA* constructed by fusing to *lacZ*, but not the wild-type gene itself, were efficiently expressed in *Escherichia coli* when they were put downstream from the *lac* promoter. These fused gene products became soluble or insoluble depending on the length of their *lacZ* moieties.

Agrobacterium; Agropine; Plasmid evolution; Signal transduction; Transcriptional regulation; Virulence gene

1. INTRODUCTION

The plasmids pRi and pTi confer tumorigenic symptoms at wound sites on a wide variety of dicotyledonous plants upon infection by their host bacteria, *Agrobacterium rhizogenes* and *A. tumefaciens*, respectively. This tumorigenesis is caused by the transfer of a defined DNA segment (T-DNA) from pRi or pTi into the plant nuclear genome and subsequent production of plant phytohormones directed by the T-DNA. The T-DNA also carries genes coding for enzymes that synthesize the unique amino acid derivatives called opines, by which pRi and pTi are generally classified. The *vir* loci, which are located outside the T-DNA, contain many genes essential for the T-DNA transfer (for reviews see [1,2]). These genes constitute six transcriptional units with similar linear organizations for pRi and pTi (*virA*, *virB*, *virG*, *virC*, *virD* and *virE*), and their expressions are tightly regulated, being inducible by plant phenolic compounds such as acetosyringone [3–6]. VirG has been thought to be the positive regulator for expression of a set of *vir* genes

Correspondence address: A. Oka, Laboratory of Molecular Biology, Institute for Chemical Research, Kyoto University, Uji-Shi, Kyoto-Fu 611, Japan

Abbreviations: bp, base-pair(s); IPTG, isopropyl- β -D-thiogalactopyranoside; kb, kilobase or 1000 bp; nt, nucleotide position(s) corresponding to the numbers in Fig. 2; ORF, open reading frame; pRi, hairy-root-inducing plasmid; pTi, tumor-inducing plasmid; SDS, sodium dodecyl sulfate; *vir*, virulence gene(s); VirA, protein coded for by *virA*; VirG, protein coded for by *virG*

EMBL accession number: The accession number for the sequence reported in this paper is X51418

[7–10], and VirA might have a role both in sensing plant phenolics and in the activation of VirG [10–12]. The four published *virA* sequences of pTi (one is a nopaline-type pTiC58 and three are octopine-type pTi15955, pTiA6, and pTiAg162) seem to support this view [13–15]. We here sequenced the *virA* locus of an agropine-type pRi, pRiA4, and compared its nucleotide sequence with those of the four pTi's.

2. MATERIALS AND METHODS

Standard procedures for DNA experiments and most of the materials used were previously described [5]. pUAO28 carries the HindIII-10 fragment of pRiA4 inserted in pUC18.

Fused genes of which the head and tail portions were derived from *lacZ* of *E. coli* and *virA*, respectively, were constructed as follows. The 5'-portion of *virA* was deleted to the *Sma*I site at nt 1118, the *Sul*I site at nt 1199, the *Nco*I site at nt 1470, or the *Sal*I site at nt 1599, and the resulting 5'-truncated *virA* genes were named *virA*'-1 to *virA*'-4, respectively (see Fig. 1). Each *virA*' was connected in-phase with the *lacZ* 5'-portion carrying the first 5 codons (*lacZ*') on pUC18 through a synthetic oligonucleotide linker, giving four fused genes (*lacZ*::*virA*'-1 to *lacZ*::*virA*'-4). Similarly, each *virA*' was placed after the *lacZ* 1021st codon (the last codon but two) on pUR291 [16], and four fused genes capable of producing active β -galactosidase (*lacZ*::*virA*'-1 to *lacZ*::*virA*'-4) were obtained. These fused genes were expressed in *E. coli* by IPTG-induction. Soluble and insoluble fractions of proteins were separated by centrifugation (100000 \times g, 30 min) after disruption of induced cells with sonication. Proteins were analyzed by electrophoresis on 4–20% or 10–20% gradient gels of SDS-polyacrylamide.

3. RESULTS AND DISCUSSION

3.1. Nucleotide sequence of the *virA* gene of pRiA4

We previously showed that the *virA* gene lies on the HindIII-10 fragment [5]. By cleavage with EcoRI, this fragment gave four subfragments of 1.9 kb, 1.6 kb,

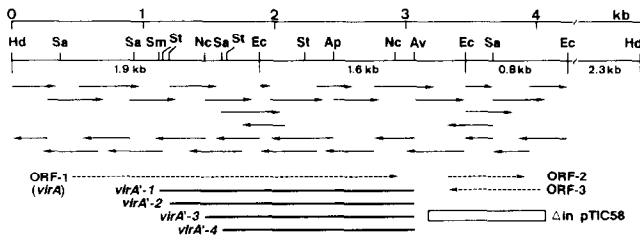


Fig. 1. Cleavage map and sequencing strategy of the *Hind*III-10 fragment. The scale is graduated in kb from the left *Hind*III site toward the right. Relevant cleavage sites are shown below the scale (Hd, *Hind*III; Ec, *Eco*RI; Sa, *Sal*I; Sm, *Sma*I; St, *Stu*I; Nc, *Nco*I; Ap, *Apa*I; Av, *Ava*I). Sequenced regions are indicated by arrows in the 5'-to-3' direction. The three broken arrows represent ORFs. Filled bars and an open bar indicate DNA regions carried by four *virA'* genes and the 900-bp region missing on pTiC58, respectively.

0.8 kb, and 2.3 kb, the former two of which were hybridized to the probe carrying the pTiA6 *virA* gene (data not shown). After these two subfragments and the next 0.8-kb subfragment were separately cloned on M13mp18 and M13mp19, many progressive deletion mutants were isolated by ejecting appropriate restriction fragments, and were then sequenced. The three *Eco*RI-subfragments being contiguous was confirmed by sequencing proper fragments across their junctions. These sequencing strategies are shown in Fig. 1. By arranging overlapping sequences of both strands, the entire nucleotide sequence of the 4.2-kb region was determined as shown in Fig. 2. Occurrence of translation start and stop codons in this region showed three ORF's longer than 400 bp. ORF-1 (nt 463-2949) and

```

                30                                60
AAGCTTTCGGGATGACCTAAAATGAAAATGCCCGCTGAAGAATATATAATGGATATCGG
    90                                120
CATGAAACAGTCGATTCGGTTTCAAATATCGCCGATGCTCTGGATCTTACGAAGACGACA
    150                                180
CGTTTGAGCGGTACGAACACACACTTAAATCTCCGCGCGCCCGCCACGGCAGTCTCG
    210                                240
GCGCTCGTAGCATAGTAGCTTGTCTCAGCGGATGATGATCAGCCATTTTGTCTATGCTGG
    270                                300
CGATTCATTTCTTTTCGATACAAACCGCGATATCTGAAAGCCACCACGCGCAGTCAAA
    330                                360
CAACACTCGGGCAGTGGAAACATCGTTGGAAATGCCAGGGCTGTGTTTCAATTTGAAACAAA
    390                                420
CTGAGTCGACGCTCTGTGATTTCAAACCCATTTACAAAGCCCTACCGTGC GGCCTAAGCGCC
    450                                480
ACGGGAGTGGGACTGAGGCTTGGCAGGAGACATAAGTGGGATGAATGGAAAGTATTCA
    510                                540
CCGACTCGGCAAGATTTCAAGACAGCGCCCAAGTCTTGGTCTGTCTGGCTCTGATAGT
    570                                600
GCTGCAATGATTTTGGCTTGGTAGCGATTTGGCTGTGGCAGGACAATGAAACCAATCGG
    630                                660
GCAATCTGACCCAGTTGGCGGCGATTAACATCGACAGTCCCTCGCTGCAGCGGGATGTA
    690                                720
CTCCGCGGGATGCGGGTATGTTGGGGAATACCGGCCATTTATCTCCAGTTGGGGCT
    750                                780
TTGCGGAAGAACCTGAAAATCTGAAGCGATTAATTAACAATCCCATCTTGTGAGCGGC
    810                                840
AACGATTTCTCAACTGCTAGACAGCTAAAGATGCTGTGGATACGACCGACGCGGCC
    870                                900
GTCCGATCTCTTGGAAACGAAACCTGCTCTGCAAGATTCGCTTGGCAGCTTCACTCGC
    930                                960
GCGCTGAGTATTTCCCAAGATGCTGCGAGGATCAGACGGTCGAAAAATTCGATCGAA
    990                                1020
TTGGGCAGCGTATGCTGCAAGTTTACCGCTCAGCCAAAGCCAGCACTTTCGTTAGACATC
    1050                                1080
AGTCGGAACCTGACATGCTCCAAAAGCTAGCGGTGGGGATGAAGCTCCTATCCGATATA
    1110                                1140
CTTACACCGGAGGTCACGTTATCTTGTCTGTTTGGCCCGGGTCAACGATGCGGTAAC
    1170                                1200
ATGATTCAGGCTCCGACACCGCTGAATATGCGGAAAGTTGCAACGTAAGTGTGAG
    1230                                1260
GCCATAGCTTACAAAGCGTAGAGAGACGCGCGCGGATCTTCTTGGATTCGCTGTC
    1290                                1320
GTGGCCCTTTGCATCTACATCACTCTGCTTATAGTTGCGCGGAAACCGATGTTG
    1350                                1380
TTAGCGCGGCTTTGGATTATGAAGAGCTAATCAAAGAGATTGGGGTTTGGTTTGAAGGC
    1410                                1440
GGTGGGCCACAGCTTCGTCGCGCAGGCTGCGCTTGAATATTACAGCTTTCTTTTAC
    1470                                1500
GCCGATTCGTGTGCGTTAGCATTAGTGGACCATGGAACAGTGGGCTGTCGAGAGTTTC
    1530                                1560
GCTGCGAAGCTGCTGAGCCAGCTGCGGAGGACGAGGCGCTACCGAAATGGTCTCTTTT
    1590                                1620
GCCAAGCGGACGAACTGCGTCAGTATTCCCGCATCATGTCGACGCGAAAGTGGCGTGC
    1650                                1680
TTTTCTCCAGAAATCCAGGCTTTTCGATACTATTGGCACAAATCTACCGATCACTG
    1710                                1740
ACAGCGATTTGTTACATAGTACCAGTACTGCTGCGACCTGTCGCGGCAAAAT
    1770                                1800
CAGCTTCTTGGACTCGCCACCGCTGCTGCGCACTATATCGATGTCGCGCAAGCAG
    1830                                1860
TCGAAATGCGCAATTCGAGAGCGGATAGAGACGCGGAACGCTTACGGCAGTTSGT
    1890                                1920
ACGCTTGTGGTGAATAGCGCATGAAATCAACAACATTTGGGAGCAATCTCCGCTAC
    1950                                1980
GCTGAAATGGCCAAAACCTCGCTGCGCGGCTCATCGGTACCCCAAGATATGTTGACCAA
    2010                                2040
ATTATTTCGTCGGGTGACAGACCGGCTCATATTCGATCAGATCTTAGCTTAAGCCGA
    2070                                2100
AAACTAGAACCGGTGACAAAGCGTTTATGCTCCGAGCTCGTATGAGATGCTGCTCC
    2130                                2160
TTATTTGCTGTGCTTTGCAACGCAACATCGAGCTGAATTTCAAGTTTATGATGAAGAAG
    2190                                2220
AGCGTAGTCGAAGGAGCCCGCTTGAAGTTACGAGACTACTGATGAACCTTTGCAAGAAC
    2250                                2280
GCTTCCAGGCGCTTACCAGCGGATGCGCAAAATGACATCTTATTAGCCGAACCTTTTAA
    2310                                2340
TCTCGACAGAAAATCTGGCAGATGCGCTATGCGCAGCGCGGACTATGTTCTTCTTCT
    2370                                2400
GTCAGCGACAACGGTGAAGGCATTTGCCGAGGCTGTGCTCCCCACATTTTTTGAGCCCTTC
    2430                                2460
TTTACACACGCTCTCGCAGCGGGGTTACGGGTCTAGGCTTGTGCTGCCGTGCACGGTCA
    2490                                2520
GTCAGCGCGCTTGCAGGATACATTCAGCTAACTTCAGCTATGGGGCGAGGGACGCGCTTC
    2550                                2580
GACATTTATCTACCTCTTCTCGAAGAACCCTGACGCCGACGCGGTTTTTTGGCCGC
    2610                                2640
TGCAAAACACCGCTGGCAACGGAGAAATTTGGCATTTGGTTAGCCAGATCTCTGCTTG
    2670                                2700
CGAGAGCGTACGAAGATAAGATCGCCGCTCTGGGCTATGAGCCGTTGGGCTTTAAGACA
    2730                                2760
TGTGCAGAACTTTGCAATTTGGATGTCAAAAGCGCAACAGCGATCTGGTCTGTTGAC
    2790                                2820
CAATGCTCTTTCCGCAACTCAGAGTCTTCTGCTTGCATGACGCCTTCAAGACAGCC
    2850                                2880
TCCATCATCTTGGAGAAAGTATCTTAATAATTCGCTTTCCAGGATGACATGACGCTCC
    2910                                2940
GCGCTTTTTCGCCAAACCTATATCTCTCAGGACATGGCCATACGCAATTCGACCAAG
    2970                                3000
ATCAAGCGCTAGATTTGAGGATTAATTCGGGAAACTTAAAGCTGTTGCGCGCGCTTAAGA
    3030                                3060
CAGATGGTCAGGCGAGATAGCTGTGGTTTCTTCCAGTCTGCTTTGGCAGCGCGTTTGG
    3090                                3120
CGGTCGTTTCCGAGACCCAAAGTTGGGATTAATCTGCTGGCCCTTTCCGATACGGCTTT
    3150                                3180
CGGTCAACAAAGCAATATCTTCACTAAGCTCAACAGCAGGATAGTCTGATGGCGGACA
    3210                                3240
AACAGCGCTCTATCGAGAATCTGCAATGCGGTAATCACTAAATGCTTTTGTCTGTTTTC
    3270                                3300
GAGCGCTGAGCGGCTGTCGACTAAATTAAGCGAGGATTTGGACAACCCATATGGGGA
    3330                                3360
CCCCGCTTTGCTTCCCGCGGGTCTGATCAATGTTGCTTCTGCTTTTTCGCAAGCGCC
    3390                                3420
GCTTTTCCGAGCGGTTCCCGCCGAGGAGTTTACGATACGCTCTTGGACAACATCCGCTC
    3450                                3480
CTGATCTCGGCGAATCAACACGGCCGCTCAAGCGCATGAATTCGCGATGCTGAGGA
    3510                                3540
TCCGCGGTTTGGCCATTTCTTACAGGAAGCTCCGTAAGCGGATTTTCGGCCCAACTCC
    3570                                3600
TCCGCGCGCCCTGAGCTCACCTCGAGGAAGCGGATACTCCAGGAACGGATTGCGCA
    3630                                3660
TCTTGTATGGCTTCGATGTGATGGCGTCTGCCAGTTATCACCGAAGTCATAGTATAG
    3690                                3720
GTGAAGTGCACACACCGGATGATGAGTGTGCAATGACAGCTTTCGAGCGGCATAA
    3750                                3780
GTCTTGGCTGCTAGGCGCTTCTCAGGATCAGTACATCATCTACCACCGGCTCA
    3810                                3840
AACTCGAAGAGATGATAATCTCAAACGATCTGTCGCTGAATGACGCTGCGGAGACCC
    3870                                3900
TTGAGGCTCAAGTCAAGTGGCACTTGCATCTTCCGAGATCGCTGGTCTTCCATCCCGT
    3930                                3960
AGTTGAAGCAGCAGCGAGCTAAGCGATCGGGCGGGTCAATGGCCCAAAACCAAGTTG
    3990                                4020
CGGCTGATCTTCCAACTCGGGGAGCAATCTCTGATATGTTGGTGGCATGGCC
    4050                                4080
AGAGCATCGGAGCAGCTGCTTCCGAGCTGATCTGTTGAGCTTTCCGCTCAGGACTA
    4110                                4140
IS66 -CAGCCACACTCGCGATCGATGCTGAGCTTGGAGCTTTCGGCTGATGA-
    4170                                4200
CCCTCGCTACGCAATGACACTATCGACCCCTGATCGGAACCTGCGAAGAACACAGTT
    4230                                4260
CATTGAAACATCTCCGCTCAATTTCCGCCATATCGACCCGATATCGACAGTGAATCCG
    4290                                4320
AGTTCAACATCGTTGCCGATAACGTTAAGCGTACGAATTC
    
```

Fig. 2. Nucleotide sequence of the pRiA4 *virA* region (EMBL accession number X51418). The sequence of *virA*-coding strand is shown in the 5'-to-3' direction. The nucleotides are numbered from the left *Hind*III site toward the right, corresponding to the scale in Fig. 1. The IS66 sequence homologous to pRiA4 is below the corresponding pRiA4 sequence.

ORF-2 (nt 3330–3920) were oriented from left to right, and ORF-3 (nt 4040–3330) had an opposite orientation. These ORF's were all preceded by potential ribosome-binding site sequences [17] marked 'SD' in Fig. 2. ORF-1 could code for 829 amino acids with a molecular weight of 91 171, which was concluded to be pRiA4 VirA on the basis of the strong amino-acid sequence similarity to pTi VirA (Fig. 3) [13–15]. ORF-2 and ORF-3 appeared to be nonessential for pathogenicity because they were not conserved in pTi's (see below). As previously reported, there are two promoters for *virA*: the upstream one is constitutively functional and the downstream one is acetosyringone-inducible [6]. Their -35 and -10 region sequences are indicated by '-35' and '-10', respectively, in Fig. 2. The downstream promoter region includes two phased 6-bp *vir* box sequences preceded by an inverted *vir* box (indicated by 'Vir Box' in Fig. 2), which work as recognition signals for VirG [7–9].

An attempt was made to overproduce VirA in *E. coli*. The *virA*-coding sequence together with its upstream region was placed downstream from the *lac* promoter, but no synthesis of VirA upon IPTG-induction was detected by gel electrophoresis (data not shown). As a next effort, we tried overproduction of truncated VirA carrying the cytoplasmic domain(s) (see

below). Various lengths of the 5'-portion of *virA* were substituted in-phase with either the almost entire *lacZ* sequence or the short *lacZ'* sequence carrying only the first 5 codons, downstream of the *lac* promoter (see section 2). All of these chimeric genes yielded fusion proteins of the expected sizes (Fig. 4a, b). However, the latter group of proteins were completely localized in the insoluble fraction, while the former group of proteins with β -galactosidase activity partly remained in the soluble fraction (Fig. 4c). These results suggest that the NH₂-terminal portion of VirA acts repressively for overproduction, and that the tertiary structure of VirA fusion proteins varies depending on the length of LacZ linked to their NH₂-termini. In support of this view, simple deletion derivatives of *virA*, lacking the *lacZ* moiety, were expressed as well as the above fused genes, but the gene products were extremely unstable and quickly degraded (data not shown).

3.2. Sequence comparison between the *virA* loci of pRiA4 and four pTi's

The structures of the *virA* gene and its flanking regions of pRiA4 were compared with those of the four pTi's. Within the *virA*-coding sequence, 90%, 78%, 75%, and 62% of those of the nucleotide residues of pRiA4 were identical to those of pTiC58, pTi15955, pTiA6,

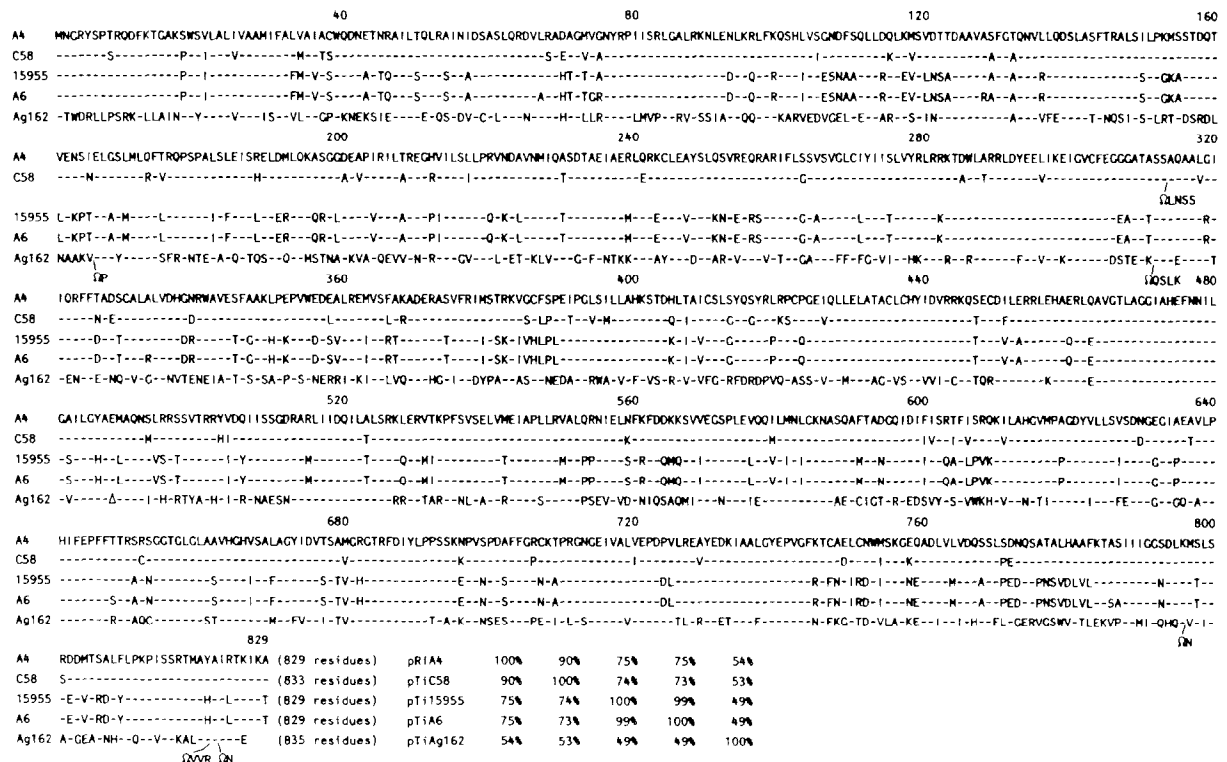


Fig. 3. Alignment of the VirA amino-acid sequences of pRiA4, pTiC58, pTi15955, pTiA6, and pTiAg162 [13–15]. Residues which differ between pRiA4 and pTi's are shown, and those which are identical are represented by hyphens. Greek delta and ohm indicate spacing and looping-out, respectively, for best matching. Each vertical row at the end of figure shows similarity (% identical amino-acid residues) between a standard VirA (indicated as 100%) and the other four VirA's.

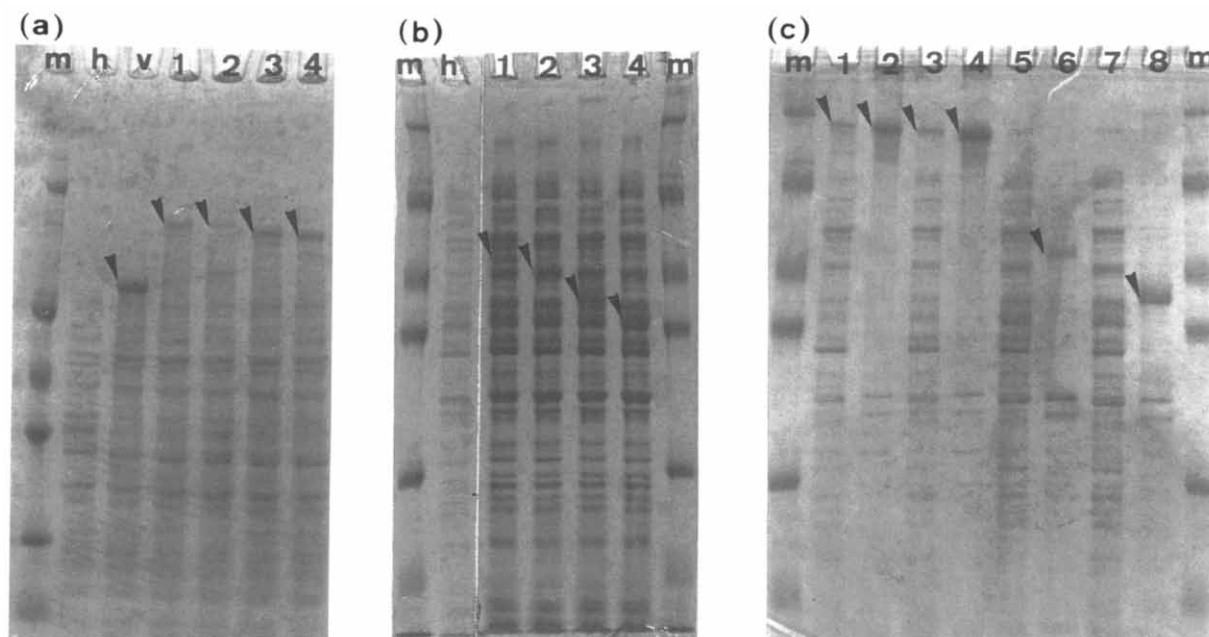


Fig. 4. Electrophoresis in 4–20% (a) or 10–20% (b, c) gradient gels. (a, b) Total proteins isolated from IPTG-treated *E. coli* cells carrying no *virA/lacZ* (h, host), *lacZ* (v, vector), *lacZ::virA'-1* (a1), *lacZ::virA'-2* (a2), *lacZ::virA'-3* (a3), *lacZ::virA'-4* (a4), *lacZ::virA'-1* (b1), *lacZ::virA'-2* (b2), *lacZ::virA'-3* (b3), or *lacZ::virA'-4* (b4). Overproduced protein bands indicated by arrowheads were not detected without IPTG-induction (data not shown). (c) Proteins in soluble (odd number) and insoluble (even number) fractions derived from *E. coli* cells carrying *lacZ::virA'-1* (1, 2), *lacZ::virA'-3* (3, 4), *lacZ::virA'-1* (5, 6), and *lacZ::virA'-3* (7, 8). The molecular sizes of marker proteins (Amersham) at lane (m) are 200, 97.4, 69, 46, 30, 21.5 kDa, respectively.

and pTiAg162, respectively. These values corresponded to 90%, 75%, 75%, and 54%, respectively, with the amino-acid sequences. The identities between all other combinations are summarized in vertical rows at the end of Fig. 3. Two hydrophobic regions identified in the NH₂-terminal half of pTi VirA [13–15] were satisfactorily conserved in pRiA4 VirA (amino-acid residues 18–39 and 260–280), and thus every VirA should have similar topology: these two regions are membrane-spanning, while their inside and outside regions are periplasmic and cytoplasmic, respectively [11,12]. The sequence comparison of VirA further supported the previous proposal on the phylogenetic relationships among these plasmids, that a set of the *vir* genes have evolved from a common ancestral set and that pTiC58 has the greatest similarity with pRiA4 (see the second row at the end of Fig. 3); nevertheless these two plasmids belong to different incompatibility groups [5,18]. However, pTiC58 *virA* carried an extra

12-bp sequence between nt 1401 and 1402 of the pRiA4 *virA* sequence. Since pTi15955 *virA* and pTiA6 *virA* also contain no such sequence, this 12-bp sequence seemed to be added after phylogenetic separation between pRiA4 and pTiC58.

In contrast to the *virA*-coding sequences, their 5'-flanking regions were comparatively less conserved. Furthermore, the conserved portions were generally limited within a short stretch including the inducible promoter region, and only in the case between pRiA4 and pTiC58, the homologous regions extended to the farthest region that was sequenced. Within the conserved 5'-flanking regions, the signal sequences for the translation initiation [17] and for VirG-binding [7–9] were particularly close to one another except for pTiAg162.

As to the 3'-flanking regions, almost no similarity was found among the five plasmids, except between pRiA4 and pTiC58. In addition, the homologous

```

VirA      ■
ETIVALVEPDV-LREAYEDKIAALGYEPVGFKTCAELCNWMSKGEQADLVLDQSSLSDNQSATALHAAFK--TASIIIGG--SDLKMSLSRDDMTSALFLPKPISSRTHAYAIRTKIKA
VirG
MKHVLVIDDVAMRHILIVEYLTIHAFKVTAVA-DSKQFNRVLSSETVDVAVVDLNLGREDGLEIVRTLATKS-DVPMIIISGDRLEEADKVVALELGATDFIAKPPGTREFLAIRIRVALRE
NtrC
AGSILVADDDTAIRTVLNQALSRAGYEVRLTC-NAATLWRWWSQEGEDLVITDVVMPDENAFDLLPRIKKMRPNIPVIVMSA-QNTFMTAIRPSERGAYEYLKPPFDLKELITIVGRALAE

```

Fig. 5. Amino-acid sequence similarity between pRiA4 VirA (715–829 residues) and the positive regulators, pRiA4 VirG (1–118 residues) [18] and *Bradyrhizobium parasponiae* NtrC (3–121 residues) [10]. The three residues that are always conserved within the phosphorylation target domain of positive regulators are marked with black squares. Identical residues or conserved replacements are represented by an underline. Grouping for conserved replacements are as follows: (A, G, S, T, P), (L, I, V, M), (D, E, N, Q), (K, R, H), and (F, Y, W).

region exceptionally found between pRiA4 and pTiC58 contained an intervening 900-bp sequence on pRiA4 (nt 3179–4078). The right end of this 900-bp sequence resembled IS66 (see Fig. 2), an insertion sequence that has been found in *A. tumefaciens* strain A66 [19]. The upstream region of pTi15955 *virA* but not of the other *virA* genes shows significant similarity to IS66 [14]. These facts suggest that DNA rearrangements occurred frequently around the *virA* regions after phylogenetic separation.

3.3. Functional domain common to *VirA* and *VirG*

It has recently been shown that pTi *VirA* is autophosphorylated, presumably at His-474 in the middle portion [20], the amino-acid sequence around which is highly conserved among pRi and pTi (see Fig. 3). The phosphorylated *VirA*, as in the case of other sensors such as NtrB, EnvZ, and PhoR, appears to activate *VirG* by transferring its phosphoryl group [10,21]. The phosphorylation target domain of positive regulators for respective sensors (*VirG*, NtrC, OmpR, PhoB, respectively) is composed of about 100 amino acids, the central Asp residue (Asp-52 for *VirG*) of which seems to receive the phosphoryl group [10]. A similar domain structure was found at the COOH-terminal region of *VirA* (amino-acid residues 715–829), as shown in Fig. 5. Since the amino-acid residues within this region are similar to one another among the five plasmids, particularly the three important residues marked with black squares in Fig. 5 are completely conserved, this possible phosphorylation target domain may contribute to regulation of signal transduction including the phosphotransfer from *VirA* His-474 to *VirG* Asp-52.

REFERENCES

- [1] Nester, E.W., Gordon, M.P., Amasino, R.M. and Yanofsky, M.F. (1984) *Annu. Rev. Plant Physiol.* 35, 387–413.
- [2] Hooykaas, P.J.J. (1989) *Plant Mol. Biol.* 13, 327–336.
- [3] Stachel, S.E. and Nester, E.W. (1986) *EMBO J.* 5, 1445–1454.
- [4] Stachel, S.E. and Zambryski, P.C. (1986) *Cell* 46, 325–333.
- [5] Hirayama, T., Muranaka, T., Ohkawa, H. and Oka, A. (1988) *Mol. Gen. Genet.* 213, 229–237.
- [6] Aoyama, T., Takamami, M. and Oka, A. (1989) *Nucleic Acids Res.* 17, 8711–8725.
- [7] Jin, S., Roitsch, T., Christie, P.J. and Nester, E.W. (1990) *J. Bacteriol.* 172, 531–537.
- [8] Pazour, G.J. and Das, A. (1990) *J. Bacteriol.* 172, 1241–1249.
- [9] Tamamoto, S., Aoyama, T., Takamami, M. and Oka, A. (1990) *J. Mol. Biol.* (in press).
- [10] Stock, J.B., Stock, A.M. and Mottonen, J.M. (1990) *Nature* 344, 395–400.
- [11] Melcher, L.S., Regensburg-Tuink, T.J.G., Bourret, R.B., Sedee, N.J.A., Schilperoort, R.A. and Hooykaas, P.J.J. (1989) *EMBO J.* 7, 1919–1925.
- [12] Winans, S.C., Kerstetter, R.A., Ward, J.E. and Nester, E.W. (1989) *J. Bacteriol.* 171, 1616–1622.
- [13] Leroux, B., Yanofsky, M.F., Winans, S.C., Ward, J.E., Ziegler, S.F. and Nester, E.W. (1987) *EMBO J.* 6, 849–856.
- [14] Melchers, L.S., Thompson, D.V., Idler, K.B., Neuteboom, S.T.C., De Maagd, R.A., Schilperoort, R.A. and Hooykaas, P.J.J. (1988) *Plant Mol. Biol.* 11, 227–237.
- [15] Morel, P., Powell, B.S., Rogowsky, P.M. and Kado, C.I. (1989) *Mol. Microbiol.* 3, 1237–1246.
- [16] R  ther, U. and M  ller-Hill, B. (1983) *EMBO J.* 2, 1791–1794.
- [17] Shinc, J. and Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* 71, 1342–1346.
- [18] Aoyama, T., Hirayama, T., Tamamoto, S. and Oka, A. (1989) *Gene* 78, 173–178.
- [19] Machida, Y., Sakurai, M., Kiyokawa, S., Ubasawa, A., Suzuki, Y. and Ikeda, J. (1984) *Proc. Natl. Acad. Sci. USA* 81, 7495–7499.
- [20] Jin, S., Roitsch, T., Ankenbauer, R.G., Gordon, M.P. and Nester, E.W. (1990) *J. Bacteriol.* 172, 525–530.
- [21] Jin, S., Roitsch, T., Ankenbauer, R., Gordon, M. and Nester, E. (1989) Abstracts of Fallen Leaf Lake Conference 'Molecular Biology of Bacterial Plant Pathogens' (South Lake Tahoe, California), p. 42.