Molecular characterization of the virulence gene virA of the Agrobacterium tumefaciens octopine Ti plasmid

Melchers, L.S.; Thompson, D.V.; Idler, K.B.; Neuteboom, S.T.C.; Maagd, R.A. de; Schilperoort, R.A.; Hooykaas, P.J.J.

Plant Molecular Biology 9: 635–645 (1987)

Molecular characterization of the virulence gene *vir*A of the *Agrobacterium tumefaciens* octopine Ti plasmid

Leo S. Melchers,[1] Dave V. Thompson,[2] Ken B. Idler,[3] Saskia T. C. Neuteboom,[1] Ruud A. de Maagd,[4] Rob A. Schilperoort[1] and Paul J. J. Hooykaas[1]

[1]*Biochemistry Laboratory, Department of Plant Molecular Biology, Leiden University, Wassenaarseweg 64, 2333 AL Leiden, Netherlands;* [2]*Agrigenetics Corporation, Advanced Research Division, 5649 East Buckeye Road, Madison, WI 53716, USA;* [3]*Abbott Laboratories, Abbott Park, IL 60064, USA;* [4]*Botanical Laboratory, Department of Plant Molecular Biology, Leiden University, Nonnensteeg 3, 2311 VJ Leiden, Netherlands*

Received 30 June 1987; accepted in revised form 17 September 1987

*Key words: Agrobacterium vir*A sequence, gene regulation, membrane protein, fusion proteins

Abstract

The virulence loci play an essential role in tumor formation by *Agrobacterium tumefaciens.* Induction of *vir* gene expression by plant signal molecules is solely dependent on the virulence loci *vir*A and *vir*G. This study focused on the *vir*A locus of the octopine type Ti plasmid pTi15955. The nucleic acid sequence of a 5.7-kilobase fragment encompassing *vir*A was determined. Genetic analysis of this region revealed that *vir*A contains one open reading frame coding for a protein of 91 639 daltons. Immunodetection with antibodies raised against a 35-kDa VirA fusion protein produced in *E. coli* identified by the VirA product in wild-type *Agrobacterium* cells. Moreover, it is shown that the VirA protein is located in the cytoplasmic membrane fraction of *Agrobacterium.* These data confirm the proposed regulatory function of VirA whereby VirA acts as a membrane sensor protein to identify plant signal molecules in the environment. The proposed sensory function of VirA strikingly resembles the function of the chemotaxis receptor proteins of *E. coli.*

Introduction

Signal exchange in plant-bacteria interactions is a process which induces the development of specialized parasitic or symbiotic interactions between the plant host and the bacterium [11].

The soil phytopathogen *Agrobacterium tumefaciens* induces tumors in plants by transferring and integrating a specific DNA fragment, the T-region, into the nuclear genome of plant cells at infection sites (see for reviews: [2, 18, 36]). The T-region transfer process is activated via the induction of the expression of the Ti plasmid virulence (*vir*) genes by plant signal molecules [39, 42]. The *vir* gene products act in *trans* to mobilize the T-region from the bacterial Ti plasmid [6, 12, 15]. In the Ti plasmid the T-region is flanked by 24-bp direct repeats [1, 42], which are the only sequences on the element essential for transfer [26]. The Ti plasmid virulence region contains at least seven operons of which some (*vir*A, *vir*B, *vir*D and *vir*G) are absolutely essential for tumor induction, whereas other (*vir*C, *vir*E, *vir*F) are only necessary for tumor induction on certain plant species [17, 29, 43, 49]. Activation of the *vir*-genes is the direct result of the recognition by *Agrobacterium* of signal moleculer produced by wounded

plants [39, 42]. A number of phenolic compounds including acetosyringone and hydroxyacetosyringone have been identified as the active plant signal molecules which trigger *vir* gene expression [42].

Induction of *vir*-gene expression is regulated by *vir*A and *vir*G ([44, 46], our unpublished results). The *vir*A locus is transcribed constitutively and is non-inducible, while *vir*G is transcribed constitutively and is plant-inducible also [44]. The VirG protein is closely related to a number of positive regulatory proteins of *E. coli* viz. OmpR, PhoB and Dye [35, 46]. The similarity of the VirG protein to these positive regulatory proteins strongly suggests that it might function similarly to activate *vir* expression.

To aid our understanding of the role of *vir*A in the regulation of expression of the *vir*-regulon we have sequenced the *vir*A locus of the octopine type plasmid pTi15955.

In this study we show that the *vir*A locus codes for one protein of 829 amino acids. Furthermore, our data clearly indicate that the VirA protein is located in the cytoplasmic membrane of *Agrobacterium tumefaciens* by immunodetection with VirA specific antibodies. Therefore, we propose that the VirA protein present in the *Agrobacterium* inner membrane transduces or transports the plant signal, whereby the VirG protein becomes activated and subsequently triggers *vir* gene expression.

Materials and methods

Materials

Restriction endonucleases were purchased from either Promega Biotech. or New England Biolabs and used according to the supplier's recommendations. T_4 polynucleotide kinase, unlabeled nucleotides, and dideoxynucleotides were obtained from Pharmacia P.L. Biochemicals. (γ-^{32}P)ATP was purchased from New England Nuclear.

Strains and plasmids

E. coli strains JM101 and KMBL1001 (wild-type K-12) were used for propagation of plasmid constructs and were grown in LC-medium [16]. *Agrobacterium tumefaciens* strain LBA969 contains plasmid pAL969, which is a cointegrate of the octopine Ti plasmid pTiB6 and R772 [13]. Plasmids pTiB6 and pTi15955 are almost identical [9]. Strain LBA1010 (wild-type) contains plasmid pTiB6. The avirulent *Agrobacterium* strain LBA2505 (*vir*A mutant) was constructed as follows. The BamHI fragment 3 of pTi15955 (Fig. 1) was subcloned onto pBR322 to yield pRAL3220. The plasmid clone pRAL3220 was digested with KpnI to delete fragment KpnI-10 followed by insertion of a 3-kb fragment containing the spectinomycin resistance (Sp^r) marker of Tn7 to produce plasmid pRAL3250. This KpnI-10 deletion was subsequently introduced into the Ti plasmid pAL969 by exchange of the wild-type sequence for the Sp^r marker via homologous recombination in *E. coli* KMBL1164 (Δ (*lac-pro*) XIII, *thi, sup*E). The resultant Ti plasmid pAL2505 was conjugally transferred to *Agrobacterium* to give strain LBA2505 (Rif^r, Sp^r). Southern analysis of *Agrobacterium* LBA2505 total DNA digested with KpnI or BamHI was used to check the position of the deletion.

Agrobacterium tumefaciens strain 15955 (LBA8255) was used for isolating octopine Ti plasmid DNA [22]. The pTi15955 fragment BamHI-3 of plasmid pRAL3220 was subcloned onto vectors pIC19R and pUC19. Subclones pRAL3235 and pRAL3282 both contain a 2147-base pair (bp) PstI fragment (2100–4246)* but in different orientations in vector pIC19R [32]. Plasmid pRAL3236 consists of HindIII fragment 21b (981–3186)* cloned into pIC19R; pRAL3254 contains KpnI fragment 10 cloned into pUC19. Plasmid pRAL3251 was derived from pRAL3220 by deletion of internal SstI fragments with SstI and ligation.

Recombinant Ti plasmids were constructed via homologous recombination of different Vir-region clones (pRAL3235, pRAL3236, pRAL3251 and pRAL3254) into the site of the deleted *vir*A locus of Ti plasmid pAL2505. Homologous recombination was performed upon introduction of a 3-kb Sp^r-marker (Fig. 3). The recombinant Ti plasmids were checked via Southern blotting of *Agrobacterium* to-

* Map position on *vir*-sequence according to sequence of Fig. 2.

tal DNA. Plasmid conjugal transfer was as described by Hooykaas *et al.* [16], plasmid isolation was by the method of Birnboim and Doly [3], and standard recombinant DNA procedures were according to Maniatis *et al.* [31].

Nucleotide sequencing

DNA sequence reactions were conducted according to the method of Maxam and Gilbert [33], as modified by Barker *et al.* [1]. Both strands were sequenced over their entire length. Sequence data were analyzed using the programs of Martinez and Devereux *et al.* [8].

Construction of inducible expression plasmids

E. coli expression plasmids pRAL3286 and pRAL3287 carry segments of the pTi15955 *vir*A coding region fused in frame with the *lac*Z α-peptide coding region of vector pIC19R which contains the *E. coli lac*-promoter. The expression plasmid pRAL3286 encodes a 65.5-kDa VirA fusion protein and was constructed as follows. Plasmid pRAL3236 was digested with SstI to remove the internal 661-bp SstI-fragment, ligated and transformed into *E. coli*. The resulting plasmid pRAL3286 contains the coding information of *vir*A from amino acid residue number 86 to 604 fused in frame to the coding information of the N-terminal 13 amino acids of the α-peptide. At the C-terminus an extra 58 amino acids are added before translation is terminated (Fig. 4).

A 35-kDa VirA fusion protein is encoded by expression plasmid pRAL3287. This plasmid was obtained by deletion of an internal 862-bp BglII-fragment from plasmid pRAL3282, followed by Klenow treatment to blunt-end-fill the BglII-site and ligated to produce plasmid pRAL3287. The fusion protein encoded by pRAL3287 contains N-terminal 11 amino acids derived from polylinker coding sequence plus an additional serine residue, as a result of BglII site filling-in reaction, fused to the C-terminal 309 amino acids of the VirA protein.

Protein purification and immunization

Strain KMBL1001 (pRAL3287) was induced with 100 μM isopropyl-β-D-thiogalactopyranoside (IPTG) and grown to an optical density at 600 nm of 0.8. The fusion proteins accumulate inside the *E. coli* cells to large amounts (1 to 10 mg fusion protein per liter of culture). Cells were disrupted by sonication and subsequently centrifuged at 12000 rpm for 1 h in a Sorvall HB4-rotor. VirA fusion proteins were recovered by dissolving the protein pellet in Laemmli sample buffer [25].

Fusion proteins were prepared by electrophoresis in preparative 10% polyacrylamide gels. Protein positions were visualized by faint staining with Coomassie brilliant blue R-250. VirA fusion proteins of about 35 kDa were isolated from excised gel strips by electrophoresis and subsequently analyzed by sodium dodecyl sulfate polyacrylamide gele electrophoresis (SDS-PAGE) for concentration and purity. Purified fusion proteins were emulsified in complete Freund's adjuvant: initial subcutaneous injection of 2 kg New Zealand White rabbits contained 300 μg of protein. Rabbits were boosted every 2 weeks with 100 μg of protein in incomplete Freund's adjuvant (during a period of 2 months) and serum was collected 7 days later.

Gel electrophoresis and immunoblotting

SDS-PAGE was performed according to Laemmli [25]. Visualization of the VirA protein bands by immunoblotting (Western blot) was performed essentially as described by Burnette [4]. (Serum was diluted 1:400). The bands were stained using the alkaline phosphatase reaction according to the Protoblot Immunoscreening system (Promega Biotech).

Cell fractionation

Agrobacterium cells of strain LBA969 were separated into four distinct fractions (total cell lysate, cytoplasmic proteins, outer membrane and inner membrane fractions) using the procedure described by De Maagd and Lugtenberg [7]. Membranes were

isolated in the presence of 0.1 M KCL. Fractions of 0.5 ml were collected from the bottom of the gradient and the relative protein content was determined by measurement of A_{280}. 2-Keto-3-deoxy-octonate content was estimated by using the thiobarbituric acid assay [19] after precipitation of fraction samples with 10% trichloroacetic acid. NADH-oxidase was measured by the method of Osborn *et al.* [40].

Results

Nucleotide sequence analysis of the virA *locus*

The genetic organization of the approximately 40-kb Vir-region of the octopine Ti plasmid is shown in Fig. 1 [14, 17, 20, 21, 43]. The *vir*A complementation group is located on the most leftward end of the Vir-region within BamHI-fragment 3 of the octopine Ti plasmid. We determined the nucleotide sequence of a 5.7-kb fragment of the pTi15955 Vir-region which encompasses the *vir*A locus (Fig. 1). The VirA-region clones pRAL3220, pRAL3235, pRAL3236 and pRAL3254 were used to determine the DNA sequences of both strands. The nucleotide sequence presented in Fig. 2 revealed two open reading frames in the analyzed 5.7-kb region of the Ti plasmid. One large open reading frame (ORF) reading from left to right was found within the interval of the KpnI fragment 10, and encodes a protein of 829 amino acids. The second ORF encoding a protein of 255 amino acids is located about 600 bp downstream of the former open reading frame and ends within restriction fragment KpnI-4. In order to determine the role that each of these two coding regions plays in virulence, we constructed a deletion mutant which lacks both ORF's. To this end the restriction fragment KpnI-10 of Ti plasmid pAL969 was replaced by the Sp^r-marker from plasmid pRAL3250 via double crossing over to yield strain LBA2505. *Agrobacterium* strain LBA2505 was avirulent on *Kalanchoe daigremontiana, Kalanchoe tubiflora, Nicotiana glauca* and tomato. Complementation of *Agrobacterium* LBA2505 was performed in *cis* to determine the smallest restriction fragment capable of complementing the *vir*A mutation (Fig. 3; for construction recombinant Ti plasmids see Materials and methods).

Agrobacterium recombinant strains were tested for virulence on three different plant species (Table 1). Strains LBA2505, LBA2507, LBA2508 and LBA2510 were avirulent on all test plants. Thus, a functional *vir*A locus was not present on plasmid pRAL3235, pRAL3236 or pRAL3251 as these *vir*-clones could not complement the *vir*A-mutation of strain LBA2505. The virulence of strain LBA2505 was fully restored upon introduction of pRAL3254 in Ti plasmid pAL2505, which implicates that a complete *vir*A locus is present on pTi15955 fragment KpnI-10.

These data revealed that the *vir*A transcriptional unit consists of a single gene capable of coding a 829-amino acid protein. Furthermore, the 255-codon

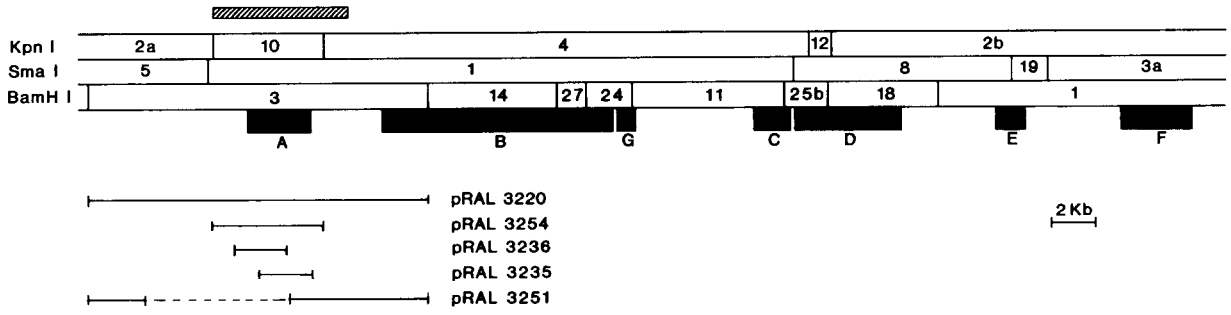

Fig. 1. Physical map of the octopine plasmid pTi15955 virulence region. Map positions of the seven different *vir*-loci are shown. The bottom of this figure shows the fragments of plasmids pRAL3320, pRAL3235, pRAL3236, pRAL3251 and pRAL3254 used for sequence analysis and complementation experiments. The dashed line indicates a deleted region present in plasmid pRAL3251. A hatched bar at the top represents the sequenced region (see Fig. 2).

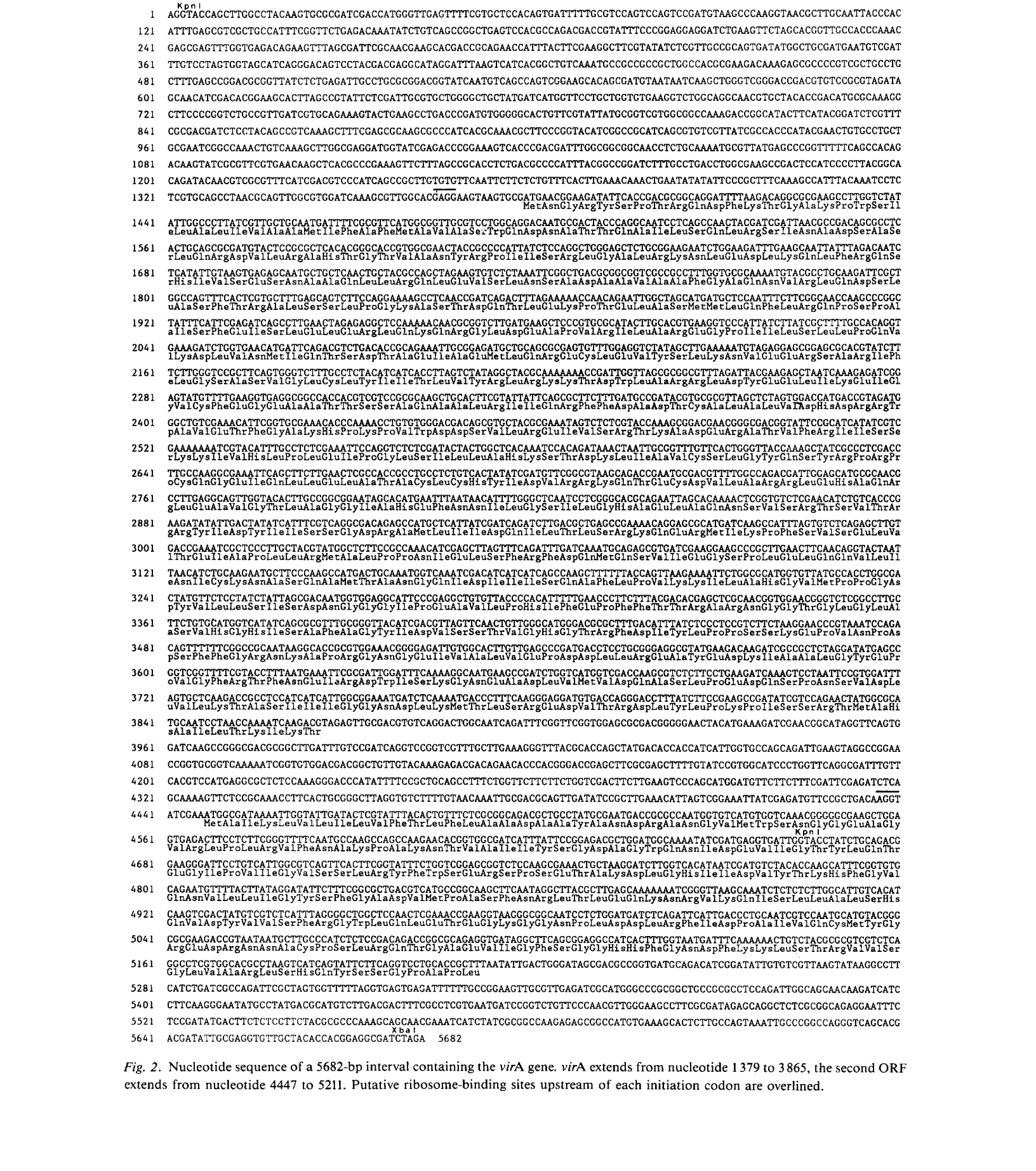

Fig. 2. Nucleotide sequence of a 5682-bp interval containing the *vir*A gene. *vir*A extends from nucleotide 1379 to 3865, the second ORF extends from nucleotide 4447 to 5211. Putative ribosome-binding sites upstream of each initiation codon are overlined.

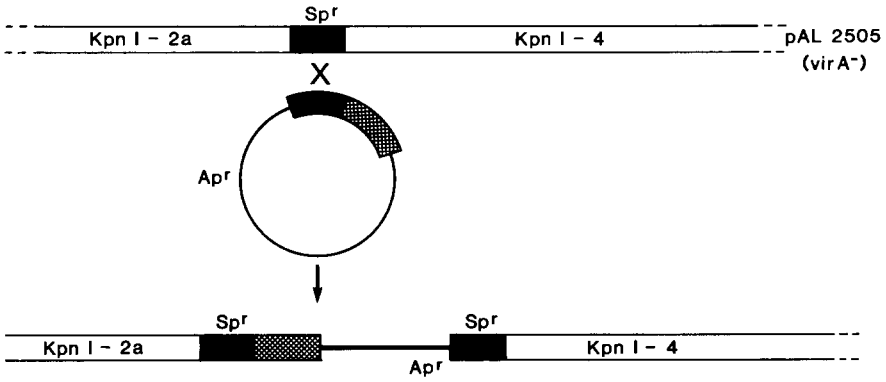

Fig. 3. Construction of recombinant Ti plasmids. Schematic drawing shows the homologous recombination process used for the introduction of Vir-region clones into the Ti plasmid pAL2505. Upon homologous recombination between both Sp^r-fragments different pTi15955 DNA fragments (indicated by a hatched region) were introduced into plasmid pAL2505. Complementation experiments in *cis* were performed with plasmids pRAL3235, pRAL3236, pRAL3251 and pRAL3254 (see Table 1).

Table 1. Plant tumor induction tests.

Strain	Plasmid	Tumors on test plant*
LBA1010	Cr, pTiB6	+ +
LBA969	Cr, R772::pTiB6	+ +
LBA2505	Cr, pAL2505 (virA$^-$)	−
LBA2507	Cr, pAL2505::pRAL3236	−
LBA2508	Cr, pAL2505::pRAL3235	−
LBA2509	Cr, pAL2505::pRAL3254	+ +
LBA2510	Cr, pAL2505::pRAL3251	−

* *Nicotiana glauca, Kalanchoe tubiflora* and tomato.

In all these *Agrobacterium* strains a large cryptic plasmid (Cr) is present. Tumor formation was classified as wild-type size + + or negative −.

ORF does not encode an essential virulence function as this region can be deleted from the Ti plasmid without affecting virulence of *Agrobacterium.*

Analysis of the 5′-end region of the *vir*A gene showed no sequence identical to the -10 or -35 consensus sequence of *E. coli* promoter regions [34]. A sequence similar to the *E. coli* consensus ribosome-binding site sequence (TAAGGAGGTG (5−9 bp) ATG) [41] was observed preceding the reading frame of *vir*A (GCACGAGGAA (9 bp) ATG). The predicted VirA polypeptide has a molecular weight of 91639 daltons and a net charge of -2. The Lipman and Pearson FASTP program [28] was used to search the NBRF protein database for proteins homologous to the VirA product. No homologous proteins were found. However, a striking nucleotide sequence homology was found between a 185-bp region of the pTi15955 *vir*-sequence and the T-center region (T_c-region) of the same pTi15955 plasmid. A 185-bp element of KpnI-10 (at position 4115−4299; Fig. 2) turned out to be present in direct repeat in the T_c-region (at position 14234−14418; [1]) with only 14 mismatches (82% homology). This element was located in the sequence between the coding regions of *vir*A and the 255-codon ORF located downstream of *vir*A.

Figure 4 shows the distribution of hydrophilic and hydrophobic amino acid residues of the VirA protein using an algorithm developed by Kyte and Doolittle [24]. The VirA protein contains two hydrophobic regions (amino acid residue 18 to 39 and 260 to 278) and closer examination of the amino acid sequence at the N-terminus revealed that the VirA protein contains a putative signal sequence. The N-terminus has a hydrophobic region of 22 amino acids followed immediately by 5 positively charged amino acids and contains a putative cleavage site according to the "(−3, −1) rule" proposed by Von Heyne [45] (amino acid residues 35 and 37 are Ala). These results suggested that VirA may be an exported or membrane-associated protein. To analyse this hypothesis we raised antibodies against the VirA protein in order to determine the cellular location of this protein.

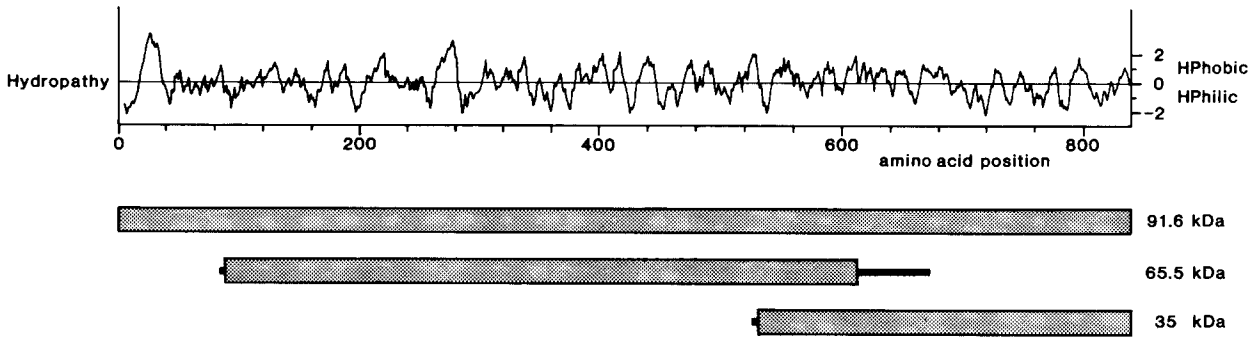

Fig. 4. Hydropathy profile of VirA plotted by the method of Kyte and Doolittle (values averaged over nine amino acids; [24]). Horizontal bars below represent the 91.6-kDa VirA protein of *Agrobacterium* and the *E. coli* fusion proteins of 65.5 kDa and 35 kDa, which are encoded by expression plasmids pRAL3286 and pRAL3287, respectively. The polypeptides encoded by vector sequences are indicated with a black bar.

Induction of VirA fusion proteins

In order to obtain VirA-specific antibodies we constructed inducible expression plasmids pRAL3286 and pRAL3287 for the overproduction of VirA fusion proteins in *E. coli*. To this end the *vir*A coding region, indicated in Fig. 4, was fused to the *lacZ* α-peptide coding sequence under control of the *E. coli lac*-promoter of vector pIC19R ([32]; constructions see Materials and methods). *E. coli* cells containing pRAL3286 or pRAL3287 produced the expected fusion proteins upon induction of the *lac*-promoter with IPTG (Fig. 5A). The fusion proteins determined by pRAL3286 and pRAL3287 were examined by SDS-PAGE and migrate close to the predicted molecular weights of approximately 65.5 kDa and 35 kDA respectively (Fig. 5). These induced proteins were not observed in *E. coli* cells which only contained the vector pIC19R. The 35-kDA fusion protein which embraces the C-terminal part of VirA was purified (see Fig. 5B) and was used to immunize rabbits as described in Materials and methods. After immunization serum was obtained and used to detect the VirA protein in *Agrobacterium*. The specificity of the antiserum directed against the 35-kDa fusion protein was shown in Western immunoblotting experiments (Fig. 6A). Proteins from IPTG-induced *E. coli* cells containing pRAL3286 or pRAL3287 were separated on SDS-polyacrylamide gels and transferred to nitrocellulose. The antiserum reacted specifically with the 35-kDA fusion protein (Fig. 6A, lane 6), but not with the 65.5-kDa fusion protein which consists of the N-terminal part of VirA. The results of this experiment show the specificity of the antibodies raised against the 35-kDa fusion protein and indicate that the antibodies are able to recognize denatured proteins.

The VirA protein is probably produced in relatively small amounts in wild-type *Agrobacterium* strains, as it could not be identified by Coomassie brilliant blue staining. Therefore, different bacterial fractions (total cell lysate, cytoplasmic proteins and cell membranes) were run on SDS-PAGE and analyzed by Western blotting using the anti-35-kDA serum. The immunodetection of the VirA protein in *Agrobacterium tumefaciens* with antiserum is shown in Fig. 6B. The VirA protein was visualized in *Agrobacterium* strains LBA1010 and LBA969 which are wild-type for *vir*A, but was absent in the *Agrobacterium vir*A deletion mutant LBA2505. The apparent molecular weight of the protein was in agreement with the size of 91.6 kDa as predicted by the *vir*A coding sequence. The VirA-signal was relatively weak in the total cell lysate, but a significantly stronger signal was detected in the fractions containing the cell envelope proteins of LBA1010 or LBA969. In the cytoplasmic protein fractions, however, no VirA protein could be visualized by immunodetection.

In order to establish the precise cellular location of the VirA protein, membranes of *Agrobacterium*

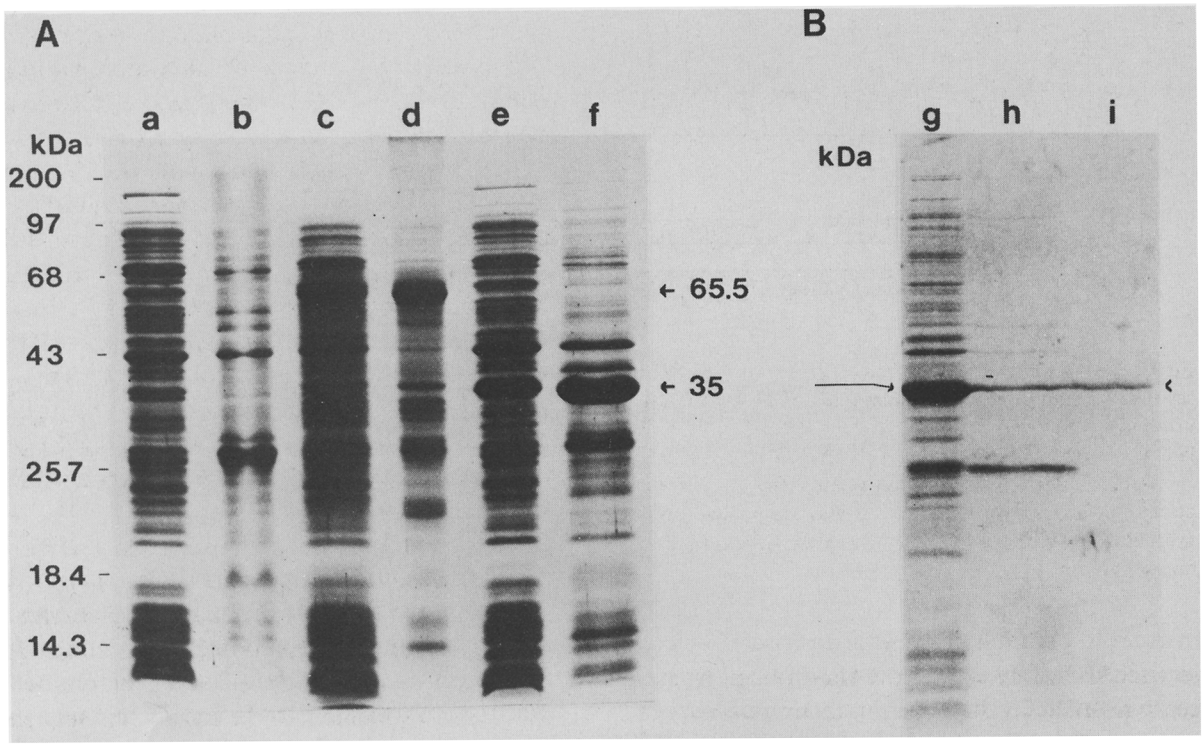

Fig. 5. Fusion proteins produced *in vivo* in *E. coli*. Proteins were stained with Coomassie brilliant blue R. (A) Protein patterns of IPTG-induced *E. coli* cells containing plasmids pIC19R (lanes a, b), pRAL3286 (lanes c, d) or pRAL3287 (lanes e, f). Lanes a, c and e total cellular protein; lanes b, d and f protein pellet fraction isolated from sonicated cells by centrifugation. (B) Protein profiles during purification of a 35-kDa VirA fusion protein. The 35-kDa fusion protein was produced upon induction of KMBL1001 (pRAL3287) with IPTG, and was visualized among total cellular protein (lane g). Lane h, protein pellet obtained by centrifugation of sonicated cells. Lane i, purified 35-kDa VirA fusion protein.

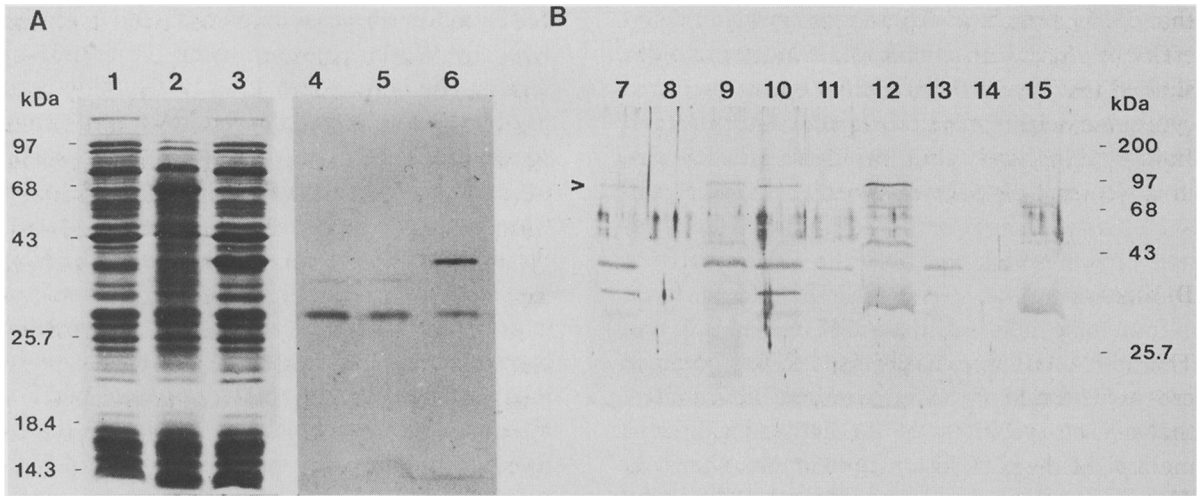

Fig. 6. Western immunoblot analysis of proteins produced in *E. coli* (lanes 4, 5 and 6) and *Agrobacterium* (lanes 7–15) using antibodies raised against the 35-kDa VirA fusion protein. (A) Total cellular protein of *E. coli* strains induced with IPTG. Lanes 1 and 4, pIC19R; 2 and 5, pRAL3286; 3 and 6 pRAL3287. Duplicate gels were used for Coomassie brilliant blue staining (lanes 1–3) and immunodetection (lanes 4–6). (B) Immunodetection of different *Agrobacterium* cell fractions. Strain LBA1010 (lanes 7–9), LBA969 (lanes 10–12) and LBA2505 (lanes 13–15) were analyzed. Lanes 7, 10 and 13, total cell lysate; 8, 11 and 14, cytoplasmic proteins; 9, 12 and 15, unseparated cell membranes.

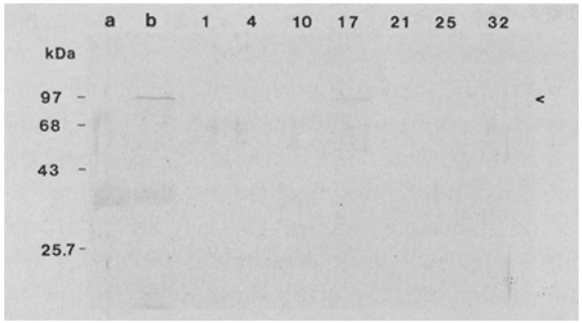

Fig. 7. Western immunoblot analysis of gradient fractions from separated *Agrobacterium* cell membranes of LBA969. Lane a, unseparated cell membranes of strain LBA2505 (*vir*A$^-$) and lane b, unseparated cell envelopes of strain LBA969. Representative fractions of isolated outer membranes (number 4), intermediate fractions (numbers 1, 10, 21, 25 and 32) and of isolated cytoplasmic membranes (number 17) were analyzed with 35-kDa antiserum.

strain LBA969 were separated using density gradient centrifugation [7]. Fractions of 0.5 ml were collected and the relative protein content of the fractions was determined. The A_{280} pattern of the gradient fractions showed two major bands for which the peak fractions had buoyant densities of 1.229 and 1.169 g/ml. The distribution of the outer membrane marker 2-keto-3-deoxy-octonate and the cytoplasmic membrane marker NADH-oxidase indicated that the separation of the two membranes was quite efficient (data not shown). Immunodetection showed that the VirA protein fractionated with the cytoplasmic membrane of *Agrobacterium* (Fig. 7; fraction 17, ρ 1.169 g/ml); no VirA-signal was seen in the outer membrane fractions.

Discussion

The *vir*A locus is essential for the formation of crown gall tumors by *Agrobacterium*. We show here that a functional *vir*A locus is present on KpnI fragment 10 of the pTi15955 octopine-type Ti plasmid. This observation corresponds with the map position of the pTiA6 *vir*A locus, determined by a detailed analysis of the genetic and transcriptional organization of the pTiA6 *vir*-region [43]. We have determined the nucleotide sequence of *vir*A, and found that the *vir*A locus contains one open reading frame that is capable of coding for a 829-amino acid protein of about 91.6 kDa. The promoter regions and transcription initiation sites of six *vir*-genes of octopine Ti plasmid pTiA6, which is almost identical to pTi15955 [9], were determined previously [5]. The nucleotide sequence of the pTi15955 *vir*A upstream region turned out to be almost identical to the corresponding *vir*A region of plasmid pTiA6. A comparison of both *vir*A upstream regions (178 bp) shows that they differ by four nucleotides only: the nucleotide sequence in Fig. 2 differs at position 1228 (T instead of A), position 1309 (T for C), position 1372 (A for T) and contains an extra nucleotide (T) at position 1253.

During the preparation of this manuscript the complete nucleotide sequence of the pTiA6 *vir*A gene was published [27]. Comparison of the nucleotide sequence of the VirA coding regions of pTi15955 and pTiA6 showed a total of 15 differences which results in 8 differences in the amino acid sequences (positions 67; 74; 75; 128; 333; 648; 786 and 787). Thus, the VirA protein encoded by pTi15955 (MW 91639; net charge -2) is somewhat different from that of the pTiA6 VirA protein (MW 91797; net charge $+1$). Our data imply that the difference in net charge of the pTiA6 ($+1$) and pTiAg162 (-2) VirA proteins mentioned in ref. [27] is not a distinguishing characteristic between wide host range and limited host range VirA proteins.

A 185-bp sequence of KpnI fragment 10 located 250 bp downstream of the *vir*A ORF turned out to be also present in direct repeat in the T_c-region of plasmid pTi15955 (only 14 mismatches). Earlier, sequence homology was found between a 531-bp region of IS51 of *Pseudomonas savastanoi* and a portion of the T_c-region in the plasmid pTi15955 [48]. The 185-bp region of KpnI-10 is not homologous to either of the transposable elements IS51 or IS52, which were found in *Pseudomonas savastanoi*. The homology found by Machida *et al.* [30] between IS66 and the virulence region restriction fragment BamHI-3 was located in the sequenced region upstream of the *vir*A gene. The nucleotide sequence of BamHI-3 (at positions 530–831, Fig. 2) showed significant homology (about 77%) with 300 bp of the IS66 DNA sequence (positions 386–692, [30]).

Two VirA fusion proteins were visualized in *E. coli* cells upon induction of expression vectors pRAL3286 and pRAL3287 with IPTG. The fusion proteins of 35 kDa and 65.5 kDa corresponded in size to the proteins predicted by the DNA sequence. The 35-kDa *vir*A fusion protein was purified using a 2-step purification procedure and used to raise specific antiserum. This antiserum was used to localize the VirA protein in *Agrobacterium* by Western immunoblot analysis.

The 91.6-kDa VirA protein could be localized in *Agrobacterium* whole cell extracts and specifically in the membrane fraction but not in the cytoplasmic protein fraction. The sequence data revealed that the VirA protein contains two long hydrophobic stretches. In addition the N-terminal amino acid sequence contains a putative signal sequence and hypothetical cleavage site according to the prediction rules of Von Heyne [45]. Hence, it was likely that the *Agrobacterium* VirA protein was an exported protein or a membrane protein. From fractioning the *Agrobacterium* membrane in outer and inner membrane it became clear that VirA is a cytoplasmic membrane protein. Identical results were reported very recently for the VirA protein of pTiA6 by Leroux and coworkers [27].

The VirA protein resembles *E. coli* protein EnvZ and various other transmembrane chemoreceptor proteins [23] in that it contains two hydrophobic regions which anchor the protein in the inner membrane, a periplasmic domain and a cytoplasmic domain. The C-terminal regions of the proteins determined by the *ntr*B gene of *K. pneumoniae,* the *envZ*, *cpx*A, and *pho*R genes of *E. coli* and the *vir*A gene of *A. tumefaciens* show conservation in their amino acid sequences [27, 38].

Next to VirA a second regulatory component VirG is essential for *vir*-expression. The VirG protein is likely to be a positive transcriptional regulator for *vir*-promoter sequences [35, 44, 46]. Our results support a model whereby the VirA protein, present in the cytoplasmic membrane, functions in the initial recognition of the plant signal molecule and transduces or transports this extracellular signal; this step leads to the activation of the VirG protein which in turn activates transcription of the rest of the *vir*-regulon.

A similar two-component regulatory system has been proposed to be present in a variety of prokaryotic species. For example, the *E. coli* genes *envZ* and *omp*R are involved in the regulation of the genes *omp*F and *omp*C in response to altered osmolarity [10]. Similar regulatory systems which respond to their environment (nutrient limitation) are *ntr*B/*ntr*C and *pho*R/*pho*B [38]. Recently it was shown that NtrB regulates the activity of NtrC by phosphorylation or dephosphorylation, whereby NtrC-phosphate activates transcription from nitrogen-regulated promoters [37]. Future work will show how the VirA protein regulates the activity of VirG and whether VirG is covalently modified in response to environmental signals.

Acknowledgements

Dr P. van de Putte kindly provided the *E. coli* strains KMBL1001 and KMBL1164. We thank Drs A. Beijersbergen and Dr C. W. Rodenburg for critical reading of the manuscript and Mrs M. J. G. Bergenhenegouwen for typing it. This work was supported in part by the Netherlands Foundation of Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

References

1. Barker RF, Idler KB, Thompson DV, Kemp JD: Plant Mol Biol 2: 335–350 (1983).
2. Bevan MW, Chilton M-D: Ann Rev Genet 16: 357–384 (1982).
3. Birnboim HC, Doly J: Nucleic Acids Res 7: 1513–1523 (1979).
4. Burnette WM: Anal Biochem 112: 195–203 (1981).
5. Das A, Stachel S, Ebert P, Allenza P, Montoya A, Nester E: Nucleic Acids Res 14: 1355–1364 (1986).
6. De Framond AJ, Barton KA, Chilton M-D: Biotechnology 1: 262–269 (1983).
7. De Maagd RA, Lugtenberg B: J Bacteriol 167: 1083–1085 (1986).
8. Devereux JR, Haeberli D, Smithies O: Nucleic Acids Res 12: 387–395 (1984).
9. De Vos G, De Beuckeleer M, Van Montagu M, Schell J: Plasmid 6: 249–253 (1981).
10. Hall MN, Silhavy TJ: J Mol Biol 151: 1–15 (1981).

11. Halverson LJ, Stacey G: Microbiol Rev 50: 193–225 (1986).
12. Hille J, Klasen I, Schilperoort RA: Plasmid 7: 107–116 (1982).
13. Hille J, Van Kan J, Klasen I, Schilperoort RA: J Bacteriol 154: 693–701 (1983).
14. Hille J, Van Kan J, Schilperoort RA: J Bacteriol 158: 754–756 (1984).
15. Hoekema A, Hirsch PR, Hooykaas PJJ, Schilperoort RA: Nature 303: 179–180 (1983).
16. Hooykaas PJJ, Klapwijk PM, Nuti PM, Schilperoort RA, Rörsch A: J Gen Microbiol 98: 477–484 (1977).
17. Hooykaas PJJ, Hofker M, Den Dulk-Ras H, Schilperoort RA: Plasmid 11: 195–205 (1984).
18. Hooykaas PJJ, Schilperoort RA: Adv Genet 22: 210–283 (1984).
19. Keleti G, Lederer WH: Handbook of Micromethods for the Biological Sciences. Van Nostrand Reinhold Co., New York (1974) pp 74–75.
20. Klee H, White FF, Iyer VN, Gordon MP, Nester EW: J Bacteriol 153: 878–883 (1983).
21. Knauf V, Yanofsky M, Monotya A, Nester E: J Bacteriol 160: 564–568 (1984).
22. Koekman BP, Hooykaas PJJ, Schilperoort RA: Plasmid 4: 184–195 (1980).
23. Krikos A, Mutoh N, Boyd A, Simon MI: Cell 33: 615–622 (1985).
24. Kyte J, Doolittle RF: J Mol Biol 157: 105–132 (1982).
25. Laemmli UK: Nature 227: 680–685 (1970).
26. Leemans J, Deblaere R, Willmitzer L, De Greve H, Hernalsteens JP, Van Montagu M, Schell J: EMBO J 1: 147–152 (1982).
27. Leroux B, Yanofsky MF, Winans SC, Ward JE, Ziegler SF, Nester EW: EMBO J 6: 849–856 (1987).
28. Lipman DJ, Pearson WR: Science 227: 1435–1441 (1985).
29. Lundquist RC, Close TJ, Kado CI: Mol Gen Genet 193: 1–7 (1984).
30. Machida Y, Sakurai M, Kiyokawa S, Ubasawa A, Suzuki Y, Ikeda J-E: Proc Natl Acad Sci USA 81: 7495–7499 (1984).
31. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
32. Marsch JW, Erfle M, Wykes EJ: Gene 32: 481–485 (1984).
33. Maxam AM, Gilbert W: Methods Enzymol 65: 499–560 (1980).
34. McClure WR: Ann Rev Biochem 54: 171–204 (1985).
35. Melchers LS, Thompson DV, Idler KB, Schilperoort RA, Hooykaas PJJ: Nucleic Acids Res 14: 9933–9942 (1986).
36. Nester EW, Gordon MP, Amasino RM, Yanofsky MF: Ann Rev Plant Physiol 35: 387–413 (1984).
37. Ninfa AJ, Magasanik B: Proc Natl Acad Sci USA 83: 5909–5913 (1986).
38. Nixon BT, Ronson CW, Ausubel FM: Proc Natl Acad Sci USA 83: 7850–7854 (1986).
39. Okker RJH, Spaink H, Hille J, Van Brussel TAN, Lugtenberg B, Schilperoort RA: Nature 312: 564–566 (1984).
40. Osborn MJ, Gander JE, Parisi E, Carson J: J Biol Chem 247: 3962–3972 (1972).
41. Shine J, Dalgarno L: Proc Natl Acad Sci USA 77: 7117–7121 (1974).
42. Stachel SE, Messens E, Van Montagu M, Zambryski P: Nature 318: 624–629 (1985).
43. Stachel SE, Nester EW: EMBO J 5: 1445–1454 (1986).
44. Stachel SE, Zambryski PC: Cell 46: 325–333 (1986).
45. Von Heyne G: J Mol Biol 173: 243–251 (1984).
46. Winans SC, Ebert PR, Stachel SE, Gordon MP, Nester EW: Proc Natl Acad Sci USA 83: 8278–8282 (1986).
47. Yadav NS, Van der Leyden J, Bennett DR, Barnes WM, Chilton M-D: Proc Natl Acad Sci USA 79: 6322–6326 (1982).
48. Yamada T, Lee P-D, Kosuge T: Proc Natl Acad Sci USA 83: 8263–8267 (1986).
49. Yanofsky MF, Nester EW: J Bacteriol 168: 244–250 (1986).