Mutational Analysis of the VirG Protein, a Transcriptional Activator of Agrobacterium tumefaciens Virulence Genes

THOMAS ROITSCH, HAO WANG, SHOUGUANG JIN, AND EUGENE W. NESTER*

Department of Microbiology, SC-42, University of Washington, Seattle, Washington 98195

Received 6 June 1990/Accepted 19 July 1990

The VirG protein of Agrobacterium tumefaciens is required in conjunction with the VirA protein for transcriptional activation of the virulence (vir) genes in response to plant phenolic compounds. These proteins are members of a family of two component regulatory systems. vir genes are activated via a cascade of phosphorylation reactions involving a specific aspartic acid residue of the VirG protein. We have conducted a mutational analysis of the VirG protein. By mutating conserved and nonconserved aspartic acid residues in the N-terminal domain, we demonstrated that two of three conserved aspartic acid residues located in two different regions are important for the phosphorylation of VirG by VirA phosphate. A third conserved N-terminal region was also shown to be critical for the biological function of VirG as a transcriptional activator. The identification of phosphorylatable but biologically inactive mutated VirG proteins suggests that not only phosphorylation but also a conformational change is necessary for its activity. We further demonstrated that phosphorylation is not required for sequence-specific binding to a vir gene regulatory sequence (vir box) and that the C-terminal domain is sufficient for DNA binding. The data support the model of a two-domain structure for the VirG protein and demonstrate that the sequence homologies to other two-component regulatory systems reflect both functional and structural homologies.

Agrobacterium tumefaciens is a phytopathogenic soil bacterium which genetically transforms susceptible plant cells by integrating bacterial DNA into the plant genome (reviewed in references 20 and 40). The pathogenic activity of A. tumefaciens requires the expression of the virulence (vir) genes, which are encoded by the large (ca. 200-kilodalton) tumor-inducing (Ti) plasmid. The vir genes are transcriptionally activated in response to phenolic compounds such as acetosyringone and certain sugars which are released by wounded plant tissue (29; G. A. Cangelosi, R. G. Ankenbauer, and E. W. Nester, Proc. Natl. Acad. Sci. USA, in press).

Two of the vir genes, virA and virG, are required for the expression of all vir genes (30). Previous studies have demonstrated that the corresponding gene products are highly homologous to members of a family of two-component regulatory systems (17, 19, 38). In those bacterial systems one protein senses a specific environmental signal and transfers the information to a second protein via phosphorylation, which in turn regulates the cellular response (for a review see reference 33).

The VirA protein was shown to be a membrane-spanning protein (18, 39), which agrees with the model of its being a transmembrane receptor sensing the plant signal. The periplasmic region interacts with the galactose-binding protein (ChE), which plays a key role in the recognition of sugar molecules which activate vir gene transcription (10; Cangelosi et al., in press). VirA is autophosphorylated in vitro at a histidine residue (11, 14), which is required for its biological activity in vivo (14).

The VirG protein is a sequence-specific binding protein, which recognizes the vir box, a 12-base-pair cis-acting regulatory sequence required for vir gene induction (7, 15, 23). The VirG protein is phosphorylated by the autophosphorylated VirA protein on an aspartic acid residue, which is conserved among all members of the response regulator proteins (13).

Although a large number of proteins have been identified as members of the sensor/regulator protein family, only a few systems have been biochemically and genetically characterized to determine the mechanism of signal transduction. Signal transfer via phosphorylation of the response regulator protein has previously been shown for CheA/CheY (9), NtrB/NtrC (21, 35), EnvZ/OmpR (2), and PhoR/PhoB (Makino et al., as cited in reference 22). All regulator proteins share an N-terminal domain, and subfamilies share C-terminal domains.

A mutational analysis of the VirG protein was carried out to elucidate the function of highly conserved amino acids in VirG. This paper presents these data.

MATERIALS AND METHODS

Bacterial strains and plasmids. A. tumefaciens Mx321 and Mx358 have Tn5Hohol insertions in the virG and virE gene, respectively, of strain A348. Escherichia coli TGI was supplied with the oligonucleotide mutagenesis kit from Amersham Corp., Arlington Heights, Ill., and was used for all cloning procedures.

The native virG gene clone in the phagemid pTZ18R, called pSW167, and the VirG protein overproducing plasmid pPC401 were described previously (15). The broad-host-range vector pUCD2 has been described by Close et al. (6).

Plasmid constructions. The virG gene on pSW167 was subjected to site-directed mutagenesis by the method of Taylor et al. (34) with a kit provided by Amersham Corp. We used 34- to 36-mer oligonucleotide primers to create the following amino acid changes: D8/N, D9/N, D10/N, D33/N, D47/N, V50/ALE, V51/D52-L53/EDE, D52/N, D59/N, D72/N, D81/N, D87/N, and D98/N. All mutations created in virG were confirmed by DNA sequencing (27).

The wild-type virG gene and the mutated virG genes were subcloned as 1.2-kilobase PvuII-Asp718 fragments in the PvuII-Asp718-cut broad-host-range vector pUCD2. To over-
produce the mutated virG gene in E. coli, fragments of the wild-type gene in pPC401 were replaced by the corresponding sequences of the mutated gene. Mutations D1/N, D2/N, and D3/N were subcloned as AflIII fragments. All other mutations were subcloned as SacI-Fst1 fragments. All plasmid constructs with point mutations in virG are listed in Table 1.

To overproduce the N-terminal domain of the VirG protein in E. coli, the virG gene in pPC401 was truncated at the 3' end after codon 156. A BspMl-HindIII fragment was deleted from pPC401, and the plasmid was religated after treatment with T4 DNA polymerase, resulting in pSG683. To overproduce the C-terminal domain of virG, an internal fragment corresponding to amino acids 11 to 112 was deleted. A Tth1111-BsmI fragment was deleted from pPC401, and the remaining plasmid was religated after T4 DNA polymerase treatment to create an in-frame fusion of the virG gene in pSG684.

Phosphorylation of VirG. The assay for phosphorylation of the purified VirG protein by VirA phosphate has been described elsewhere (13).

Other methods. The following methods were carried out by published procedures: overproduction, purification, and renaturation of VirG and mutated VirG proteins (15); DNA-binding assays (15); DNA footprinting assays (15); Western immunoblot analysis (15); virulence assays (37); and the β-galactosidase assay (28). Plasmids were introduced into A. tumefaciens by triparental mating, using the helper plasmid pRK2073 (8).

RESULTS

Biological consequences of mutagenesis of aspartic acid residues. The phosphorylated amino acid residue of the VirG protein, at position 52 (13), is one of three conserved aspartic acid residues in the N-terminal domain of response regulator proteins. To further understand the roles of aspartic acid residues in VirG phosphorylation and biological function, we replaced all conserved and nonconserved aspartic acid residues in the N-terminal domain by the nonphosphorylatable amino acid asparagine, which has a size similar to aspartic acid.

To determine the effect of the various aspartic acid-to-asparagine changes on the biological activity of the VirG protein, we tested their ability to complement a virG mutant of A. tumefaciens for acetylsyringone-mediated induction of the vir genes and tumorigenicity. Broad-host-range plasmids containing the wild-type virG and the various D/N mutants were mobilized into strain Mx321, containing a Tn3Hoh1 insertion in the virG gene. The inducibility of the virE gene was assayed by immunoblotting the acetylsyringone-induced cell extracts with VirE2 antibody. In addition to the mutation at the site of phosphorylation (D52/N), several other D-to-N mutations interfered with VirG function (Fig. 1). The mutations D47/N and D87/N resulted in a reduced induction level, and the mutations D8/N, D59/N, and D98/N completely abolished induction of the virE operon. The mutated amino acids which affect vir gene induction are either conserved among all response regulator proteins (D8) or close to such residues. These mutations are clustered in three regions: at the very N terminus with two conserved aspartic acid residues (D8 and D9), at the conserved phosphorylation site (D52), and close to a conserved lysine residue at position 102 (32).

The effect of the D/N mutations on vir gene induction was reflected in the virulence of these strains (Table 2). Agrobacterium strains containing VirG mutations which did not abolish vir gene induction were virulent on Kalanchee leaves. The mutations D47/N and D87/N, which resulted in reduced levels of induction, were attenuated in virulence (D47/N) or had lost virulence (D87/N).

Effect of the mutations on phosphorylation. In further experiments we tested whether the loss of biological activity of the various mutated VirG proteins was due to a lack of phosphorylation. The mutated virG genes were overproduced in E. coli under the control of the lac promoter, purified, and incubated with autophosphorylated VirA protein. A distinct difference exists between mutations at conserved and nonconserved aspartic acid residues (Fig. 2; Table 2). The mutation of the conserved aspartic acid D8 results in a nonphosphorylatable protein, as shown previously for the mutation of the conserved phosphorylation site D52 (13). The biologically inactive mutants VirGD59/N and VirGD98/N were still phosphorylatable by VirA phosphate in vitro. These results demonstrate that the alteration of the biological activity in aspartic acid-to-asparagine mutations at D8 and D52 was due to the lack of phosphorylation. The mutations D59/N and D98/N probably interfere with the ability of the protein to assume an activated configuration. All virG mutants which complemented the virG mutation are

MUTATIONAL ANALYSIS OF VIRG FUNCTION

FIG. 1. virG gene induction assay. A total of 10 μg of induced Agrobacterium cell extracts was run on a sodium dodecyl sulfate–12% polyacrylamide gel and immunoblotted with antibody against VirE2 protein. The results are shown for strain Mx321 with the following plasmids: pRS1500 (wild-type VirG) (lane 1); pRS5505 (VirGD8/N) (lane 2); pRS5607 (VirGD3/N) (lane 3); pRS5801 (VirGD10/N) (lane 4); pRS7001 (VirGD33/N) (lane 5); pRS7101 (VirGD74/N) (lane 6); pRS2606 (VirGV50/AVE) (lane 7); pSG688 (VirGV51DL/EEE) (lane 8); pSG687 (VirGD25/N) (lane 9); pRS7202 (VirGD/N) (lane 10); pRS5702 (VirGD72/N) (lane 11); pRS5901 (VirGD81/N) (lane 12); pRS7301 (VirGD87/N) (lane 13); and pRS7401 (VirGD98/N) (lane 14). kDa, Kilodaltons.

TABLE 1. virG gene mutations and relevant plasmid designations

<table>
<thead>
<tr>
<th>virG mutation</th>
<th>virG clone in pTZ18R</th>
<th>virG clone in pUCD2</th>
<th>lacp-virG fusion clone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>pSW167</td>
<td>pRS1500</td>
<td>pPC401</td>
</tr>
<tr>
<td>D8/N</td>
<td>pRS5002</td>
<td>pRS5505</td>
<td>pRS7505</td>
</tr>
<tr>
<td>D9/N</td>
<td>pRS5114</td>
<td>pRS5607</td>
<td>pRS7605</td>
</tr>
<tr>
<td>D10/N</td>
<td>pRS5205</td>
<td>pRS5801</td>
<td>pRS7703</td>
</tr>
<tr>
<td>D33/N</td>
<td>pRS6508</td>
<td>pRS7001</td>
<td></td>
</tr>
<tr>
<td>D47/N</td>
<td>pRS6602</td>
<td>pRS7101</td>
<td></td>
</tr>
<tr>
<td>V50/AVE</td>
<td>pRS1903</td>
<td>pRS2606</td>
<td>pRS2714</td>
</tr>
<tr>
<td>V51/EEE</td>
<td>pRS2904</td>
<td>pSG688</td>
<td>pSG686</td>
</tr>
<tr>
<td>D52/N</td>
<td>pRS2805</td>
<td>pSG687</td>
<td>pSG685</td>
</tr>
<tr>
<td>D59/N</td>
<td>pRS6703</td>
<td>pRS7202</td>
<td>pRS7803</td>
</tr>
<tr>
<td>D72/N</td>
<td>pRS5305</td>
<td>pRS7502</td>
<td>pRS6020</td>
</tr>
<tr>
<td>D81/N</td>
<td>pRS5414</td>
<td>pRS5901</td>
<td>pRS6108</td>
</tr>
<tr>
<td>D87/N</td>
<td>pRS6812</td>
<td>pRS7301</td>
<td></td>
</tr>
<tr>
<td>D98/N</td>
<td>pRS6905</td>
<td>pRS7401</td>
<td>pRS7904</td>
</tr>
</tbody>
</table>
phosphorylated by VirA. These data demonstrate that the phosphorylation of VirG is required but not sufficient for its activity as a transcriptional activator of the vir genes and that activation probably requires a transient conformational change.

Analysis of other mutations at the phosphorylation site. To further understand the mechanism of VirG activation, we tested the hypothesis that the negative charge is the major effect of phosphorylation of the VirG protein. To mimic this charge effect, we substituted the phosphorylatable aspartic acid and its two neighboring amino acids with negatively charged glutamic acid residues. The mutated protein is inactive as a transcriptional activator of the virulence genes, and the corresponding Agrobacterium strain is avirulent (Fig. 1; Table 2). Moreover, purified VirGV51DL/EEE was not phosphorylated by VirA (data not shown). This result shows that charged amino acids cannot substitute for phosphorylation in VirG activation. Although glutamic acid may also be phosphorylated to form an acyl-phosphate, the VirA protein apparently transfers its phosphate only to aspartic acid.

The OmpR protein belongs to the same subclass of response regulator proteins as VirG (4). Both proteins share homologous N- and C-terminal regions, and both transcriptionally activate other genes. A mutation in a conserved N-terminal amino acid sequence was previously found to abolish OmpR function (12). The mutation consisted of the substitution of a valine residue by the tripeptide sequence alanine-leucine-glutamate (V/ALE). The corresponding valine residue in VirG is located 2 amino acids N-terminal of the phosphorylatable aspartic acid residue. To test whether the V/ALE mutation affects phosphorylation we introduced the V50/ALE mutation into the virG gene. This resulted in a nonphosphorylatable gene product which could not complement wild-type VirG function in vivo (Fig. 1; Table 2).

Phosphorylation is not required for sequence-specific binding. Gel retardation assays were carried out to analyze whether binding of VirG to the cis-acting regulatory sequence depended on its being phosphorylated. The wild-type protein, as well as phosphorylatable and nonphosphorylatable mutated proteins, specifically retarded DNA fragments containing the virE promoter in the presence of a 1,000-fold excess of competitor (calf thymus) DNA (Fig. 3). Under the same conditions, no retardation was observed when the labeled DNA fragment came from the virE structural gene (data not shown). These results show that phosphorylation is not required for sequence-specific binding of the VirG protein.

The virG mutations are recessive in a wild-type background. Although the nonphosphorylatable VirG mutants still bind to the vir box, phosphorylation may regulate virG function by helping the protein adopt an active conformation which has an enhanced affinity for DNA. To address this question, we analyzed whether the virG mutations are dominant or recessive to a wild-type virG gene, i.e., whether a mutated and biologically inactive VirG protein competes with the same affinity as the wild-type protein for DNA-binding sites. The phosphorylation-negative mutants D8/N and D52/N, as well
MUTATIONAL ANALYSIS OF VirG FUNCTION

FIG. 4. Overproduction of the N- and C-terminal domains of VirG. (A) Schematic presentation of plasmid constructions for overproduction of the N-terminal (pSG683) and C-terminal (pSG684) domains of VirG under the control of the lac promoter. (B) Coomassie brilliant blue staining. Lanes: 1, gel-purified native-size VirG; 2, insoluble protein from isopropyl-β-D-thiogalactopyranoside (IPTG)-induced DH5α(pSG683); 3, gel-purified N-terminal domain; 4, insoluble protein from IPTG-induced DH5α(pSG684); 5, gel-purified C-terminal domain.

as the wild-type gene, were introduced into strain Mx358, and the induction of a virE-lacZ fusion was measured after induction by acetylsyringone. The mean values of β-galactosidase activity for three independent experiments are as follows: VirGD8/N [Mx358(pRS505)], 1,321 U; VirGD52/N [Mx358(pSG687)], 1,298 U; and wild-type virG [Mx358(pRS1500)], 1,348 U. These data demonstrate that the mutant and inactive VirG protein does not affect the expression of the virE operon. The mutations which result in a nonphosphorylated and inactive VirG protein are recessive in a wild-type background, although the gene copy number is fourfold higher than the Ti plasmid-encoded wild-type gene. This suggests that the nonphosphorylatable VirG protein has a lower affinity to the vir box than the wild-type protein does.

The C-terminal domain is sufficient for specific binding to the vir box sequence. To further prove the two-domain structure of the VirG protein and to test whether these two domains can function independently, both domains were overproduced in E. coli. To overproduce the N-terminal domain of the VirG protein, the virG gene under control of the lac promoter in pPC401 was truncated at the 3′ end deleting about 35% of the gene (Fig. 4). The C-terminal domain was overproduced by deleting an internal virG fragment in the virG-overproducing plasmid pPC401 corresponding to amino acids 11 to 112 (Fig. 4). The deleted N-terminal fragment corresponds to 43% of the gene. The purified N- and C-terminal domains are shown in Fig. 4.

Gel retardation assays were carried out to test the ability of the N- and C-terminal domain of the VirG protein to bind specifically to a vir gene promoter. The overproduced domains of the VirG protein were gel purified and renatured as described for the native-size VirG (15). The C-terminal domain specifically retarded a virB promoter fragment in the presence of a 1,000-fold excess of competitor DNA (Fig. 5). The N-terminal domain did not bind. The labeled fragment from the virB structural gene was not retarded (data not shown).

To determine whether the binding sequence of the C-terminal domain was the same as the VirG protein, we performed a DNase I footprinting analysis by using the retarded band shown in Fig. 5, lane 4. The C-terminal domain specifically protected about 80 nucleotides on both strands (Fig. 6). The protected area covered both vir boxes of the virB promoter, and the two vir boxes are located in the center of the protected area. This pattern is identical to that obtained previously with the native-size VirG protein (15).

To test the phosphorylation of the N-terminal and C-terminal domains, we incubated both with autophosphorylated VirA protein. Neither of the overproduced domains was phosphorylated by VirA as indicated by the lack of radioactive label at the position of the two low-molecular-weight proteins (data not shown). This result has been confirmed with several independent preparations of the truncated proteins.

DISCUSSION

A growing number of proteins have been identified as members of a family of two-component regulatory systems (32) which play a key role in cells responding to their environment. However, only a few systems have been studied in detail, and so far data on the effect of specific amino acid substitutions on the biological activity of response regulator proteins are available only for the CheY protein of the chemotaxis system (5, 26). We have conducted a mutational analysis of the VirG protein, which belongs to a different subfamily from that of CheY (4).

By mutating all of the aspartic acid residues, not only those that are conserved, we hoped to demonstrate the specificity of the expected effects at the conserved amino acids. In addition, this approach introduced point mutations evenly throughout the N-terminal domain. The presented data reveal that two of three conserved aspartic acid resi-
inactive mutated VirG proteins suggests that a conformational change is also necessary for its activity. Since data from the chemotaxis system lead to the same conclusion (5), response regulator proteins in general may adopt an active conformation upon phosphorylation. The similar effect of a specific mutation in a conserved N-terminal sequence in OmpR and VirG (V50/ALE) also demonstrates that the sequence homologies of the regulator proteins reflect functional homologies. Our data indicate that the substitution of a valine at position 50 by a tripeptide sequence interferes with the phosphorylation of a conserved aspartic acid residue C terminal of this position.

Although previous reports have suggested that phosphorylation is not required for sequence-specific binding of VirG (15, 23), we proved this unequivocally by using nonphosphorylatable mutants. This finding may explain previous reports, which have suggested that the VirG protein is biologically active without modification if it is present in large enough quantities. Increasing the copy number of virG (25) or increasing the translational efficiency of VirG (23) led to a partial activation of the vir genes in the absence of acetylsyringone and a functional VirA protein. A similar observation has been reported for another response regulator protein, the uhpA gene product (36). Although phosphorylation is not required for DNA binding of VirG, the competition experiments in a wild-type background suggest that phosphorylation may regulate VirG function by increasing its affinity for the binding site. This agrees with the findings for OmpR and NtrC, which can bind to the corresponding cis-acting regulatory sequences without phosphorylation (16, 21). Phosphorylation of OmpR stimulates its DNA-binding ability and enhances transcription (1, 3), whereas phosphorylation of NtrC is a prerequisite for transcriptional activation (21).

The presented data support the two-domain structure of the VirG protein, which was based on protein homology profiles. The effect of mutations in the N-terminal domain suggests that at least two different regions of this domain cooperate to form an active site necessary for phosphorylation. A similar situation was found in the CheY protein (5). Although the site of phosphorylation was shown to be located in the N-terminal domain, we were not able to demonstrate the phosphorylation of the N-terminal domain. This may be due to the improper folding of the specific truncated protein used for this experiment, since phosphorylation of the overproduced N-terminal domain of OmpR by the corresponding histidine kinase could be demonstrated (16). Alternatively, it may reflect a fundamental difference between VirG and OmpR. The C-terminal domain of the VirG protein was shown to be sufficient for sequence-specific binding to the cis-acting regulatory sequence. A helix-turn-helix motif has been identified previously in this region as a putative DNA-binding site (24). The DNA-binding ability of the corresponding domain of OmpR was also shown to be independent of the N-terminal domain (16). The finding that none of the mutations in the N-terminal domain of VirG interfered with DNA binding also indicates that this protein has two functional domains which are independent of one another.

This report provides additional support for the notion that the sequence homologies of regulatory proteins involved in adaptive responses reflect both functional and structural homologies. Although a common mechanism via phosphoryl group transfer has been established, additional work is necessary to understand the whole signal transduction chain in the various systems. For response regulator proteins like VirG it will be necessary to elucidate the molecular basis for
the activation by phosphorylation, the mechanism of inactivation of these regulatory proteins, and the interactions with other proteins for transcriptional activation. The development of an in vitro transcription system for A. tumefaciens should facilitate future studies aimed at understanding the function of the VirG protein at the molecular level.

ACKNOWLEDGMENTS

We thank X. Pu for excellent technical assistance, D. Parkhurst for synthesizing the oligonucleotides, and G. Cangelosi for helpful comments on the manuscript.

This work was supported by Public Health Service grant 5R01GM32618-18 from the National Institutes of Health (to E.W.N.). S.J. was supported by PMIF graduate training grants from the University of Washington. T.R. was supported by research fellowship R0798-2-2 from the Deutsche Forschungsgemeinschaft.

LITERATURE CITED

34. Taylor, J. J., O. Ott, and F. Eckstein. 1985. The rapid generation of


