# Ratcheting Up *vir* Gene Expression in *Agrobacterium tumefaciens*: Coiled Coils in Histidine Kinase Signal Transduction

Yulei Wang,[a] Rong Gao,[a] and David G. Lynn*[a, b]

*The transmembrane histidine kinase VirA is responsible for the recognition of information from several plant-derived xenognostic signals that control gene transfer between* Agrobacterium tumefaciens *and its eukaryotic host. As with other histidine autokinases, VirA appears to exist as a homodimer within the inner membrane of the bacterium. In this study, we identify the putative homodimeric coiled-coil-like motifs Helix TM2 (amino acids (aa) 259 – 288) and Helix C (aa 293 – 327) within the previously assigned signal input domain. The functional importance of these coiled-coil interactions in signal-mediated VirA activation is investigated by the construction of fusion proteins with the leucine zipper domain of the transcription factor GCN4. Replacement of the membrane-spanning and periplasmic domains of VirA with the GCN4 leucine zipper gave functional proteins with increased signal-induced* vir *gene expression. When the GCN4 fusion was used to conformationally bias the interface of the Helix C coiled coil, constitutively active chimeras were created. The activity of these constructs was dependent on the interface of the Helix C coiled coil, and a ratchet model is proposed in which VirA activation is achieved by signal-induced switching of the interfaces of the homodimer. Since VirA functions as a transducer and integrates various host cues indirectly, these data highlight its role as an "antenna" for the tumor-inducing (Ti) plasmid, able to monitor the host proteome so as to select for successful xenognostic signaling strategies.*

**KEYWORDS:**

pathogenesis · protein engineering · signal transduction · two-component system · VirA

## Introduction

Sophisticated cellular signaling mechanisms have evolved to detect environmental conditions and elicit appropriate adaptive responses. The histidine kinase sensor/transducers and aspartic acid receiver/response regulators form two-component regulatory systems which are the primary mediators of these signal transduction events in prokaryotes.[1–3] When harbored within *Agrobacterium tumefaciens*, the tumor-inducing (Ti) plasmid expresses a typical two-component system in which the integral membrane histidine kinase VirA catalyzes the transfer of phosphate to the cytoplasmic aspartic acid response regulator VirG, which is the transcriptional activator of the virulence regulon of the plasmid.[4] In response to activation, specific genes are laterally transferred between the bacterial pathogen and its eukaryotic host.

Located within the bacterial inner membrane, VirA is thought to function as a homodimer;[5] each monomer consists of two transmembrane domains (TM1 and TM2), a periplasmic domain, and a cytoplasmic domain that contains linker (L), kinase (K), and receiver (R) subdomains.[6, 7] At least four separate xenognosin signals are integrated through VirA to regulate virulence (*vir*) gene expression, including phenolics (for example, acetosyringone (AS)), monosaccharides (such as glucose and arabinose), pH value, and temperature.[4, 8] The phenolics are absolutely required for *vir* gene activation, and the other signals play synergistic roles to enhance the sensitivity, specificity, and maximal response.

VirA is distinguished from other transducers by its apparent ability to perceive this wide spectrum of input signals. With these input signals now defined, VirA provides a unique opportunity for mechanistic investigation. Based on the finding that coiled coil motifs are conserved in the linker domain of many histidine kinases, Singh et al.[9] proposed that these motifs may play an important role in the regulation of histidine kinase activity. Here we identify two coiled-coil-like motifs in the VirA linker, the domain that connects the second of the two transmembrane domains to the cytoplasmic kinase domain.[7] Data from fusion proteins designed to favor coiled-coil quaternary structure within the homodimer correlate the kinase activity of VirA with a specific interface of the coiled coil. These results allow us to propose a "ratchet model" for phenol-induced VirA activation in which specific interfaces of the coiled coil are required for kinase output activity. This mechanism would allow multiple signals to

[a] *Prof. D. G. Lynn, Dr. Y. Wang, R. Gao*
*Department of Chemistry, University of Chicago*
*5735, S. Ellis Avenue, Chicago, IL 60637 (USA)*

[b] *Prof. D. G. Lynn*
*Address for correspondence:*
*Department of Chemistry and Biology*
*Emory University*
*Atlanta, GA 30322 (USA)*
*Fax: (+1) 404-727-6586*
*E-mail: dlynn2@emory.edu*

synergistically up-regulate *vir* expression and additionally suggests a strategy for the integration and coevolution of the Ti plasmid within the *Agrobacterium* host.

## Results

### Secondary structural analysis of the VirA linker domain

Figure 1 shows the amino acid sequences of the second transmembrane (TM2) and linker (L) domains of VirA from the wide-host-range Ti plasmids VirA (pTiA6), VirA (pTi15955), and VirA (pTiC58), and a limited-host-range plasmid, VirA (pTiAg162), aligned by the pile-up program of the GCG package. The secondary structure predictions shown below the alignment were obtained from PHDsec algorithms and indicate 60.4% $\alpha$ helix, 9.8% $\beta$ strand, and 29.9% loops across this region.

Closer visual inspection revealed the characteristic heptad patterns of coiled-coil-like motifs in at least two helical regions: Helix TM2 (amino acids (aa) 259 – 288), which includes the membrane-spanning domain, and Helix C (aa 293 – 327), a relatively long stretch located just on the C-terminal side of Helix TM2. In contrast to typical dimeric coiled coils in which interhelical salt bridges are present between flanking residues at the ***e*** or ***g*** positions;[10] these two positions in Helix C are dominated by hydrophobic residues (I, L, A), aromatic residues (F), or small polar side chains (S, T). The absence of charged residues alters positional stability, but not overall association.[10] Interestingly, the four amino acid gap (aa 312 – 315) in some of the sequences will alter the registry of the coiled coil interface at the C-terminal end, but not dimerization. Therefore, these analyses predict the existence of two "coiled-coil-like" motifs within the linker domain.

### GCN4 chimeras for dimerization

Cross-linking experiments have indicated that VirA, as in the case of many other histidine autokinases,[1–3] exists as a preformed homodimer and is localized in the inner membrane of the *Agrobacterium* cell.[5, 11] To determine the dimerization stability of the identified coiled-coil-like motifs in the VirA linker domain, the linker (aa 285 – 471) was put behind a constitutive $P_{N25}$ promoter (Figure 2, pRG15)[12] and expressed in *Agrobacterium*. Under mild denaturing conditions, no apparent dimerization of the linker

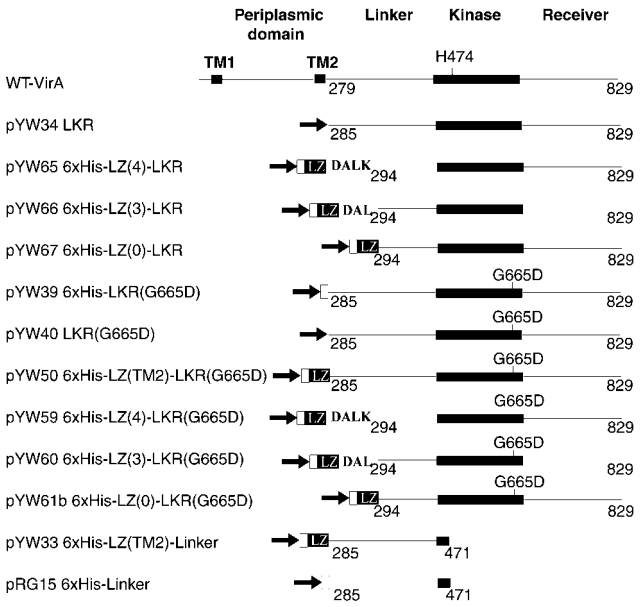


***Figure 2.*** *Schematic representation of the domain organization of VirA and the VirA fragments. The domains are indicted at the top. Black arrows: phage $P_{N25}$ promoter; unfilled block: 6 × His tag; LZ: leucine zipper of GCN4. The inserted amino acid sequences are displayed as bold letters.*

domain was detected by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis (Figure 3, lane 2), which indicates that the linker coiled-coil-like motifs are not particularly stable dimers and are not very prone to oligomerization. To enhance dimerization, an LZ(TM2) – linker fusion protein was constructed in which the VirA linker domain was fused with the leucine zipper (LZ) peptide of GCN4 (Figure 2, pYW33). This LZ(TM2) – linker dimer is more readily detected than dimers of the linker alone (Figure 3, lane 1), which suggests that such a fusion strategy successfully stabilizes dimerization.

To evaluate the functional significance of enhanced dimerization, LZ(TM2) – LKR constructs were prepared. A point mutation, G665D, within the kinase domain provides a detectable base level of *vir* gene induction, but still responds to acetosyringone (AS).[13] We elected to use this mutant as it provided a positive control for both basal and induced activities. In the LZ(TM2) – LKR(G665D) construct, the leucine zipper of GCN4 is fused with the LKR(G665D) protein (aa 285 – 829) as in the linker construct,

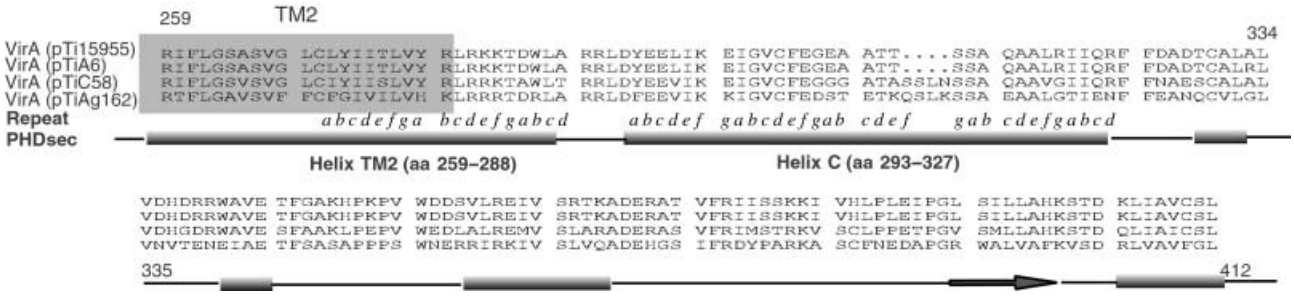


***Figure 1.*** *Sequence alignment of the four VirA linker sequences. The alignment and the secondary structure prediction were obtained as described in the Experimental Section. The numbering corresponds to the VirA (pTiA6) sequence. The heptad repeat motif is reported on the line labeled "Repeat". The secondary structure prediction is reported on the line labeled "PHDsec". Dots represent gaps. Cylinders stand for $\alpha$ helices, arrows for $\beta$ strands, and straight lines for unpredicted structure.*

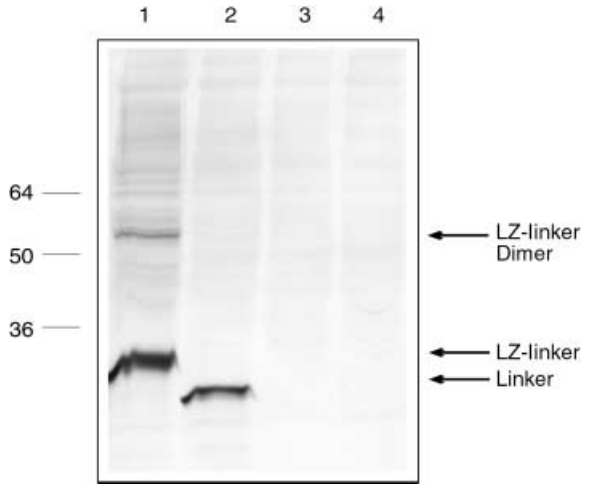


***Figure 3.*** *Immunoblot detection of LZ(TM2) – linker and VirA linker proteins. Cell lysates from* A. tumefaciens *carrying different plasmids were incubated with SDS sample buffer (Novex; 2% β-mercaptoethanol, 4% SDS) at 37°C for 10 mins, followed by SDS-PAGE (14%) and Western blot analysis with anti-RGSHis antibodies (Qiagen). Lanes: 1: A348/pYW33, LZ(TM2) – linker; 2: A348/pRG15, VirA linker; 3: A348; 4: A348 – 3/pYW15b, cloning vector.*

and replaces both the membrane-spanning and the periplasmic domains of VirA (Figure 2, pYW50). LZ(TM2) – LKR(G665D) shows significantly increased basal *vir* gene expression (without AS induction) and a maximal *vir* gene expression (with 100 μM AS) that is approximately 2.5 times greater than the LKR(G665D) construct control (Figure 4). These results provide the first functional evidence that homodimerization of VirA is important for its in vivo activity. More importantly, the presence of the LZ dimerization domain does not change the sensitivity to AS; both LZ(TM2) – LKR(G665D) and LKR(G665D) show similar dose responses (Figure 4). If AS-mediated VirA activation occurred through triggering the dimerization of VirA, LZ(TM2) – LKR(G665D) would be expected to show an induction level different from that of LKR(G665D). However, while the level of *vir* gene expression of the LZ(TM2) – LKR(G665D) construct can be further induced with AS to approximately three to four times its basal activity, the same relative level of induction occurs without the LZ dimerization domain (Figure 4). These observations, particularly when taken together with the previous cross-linking evidence that VirA exists as a dimer even in the absence of AS,[5] suggest that AS-mediated VirA activation is unlikely to be the result of an induction of dimerization.

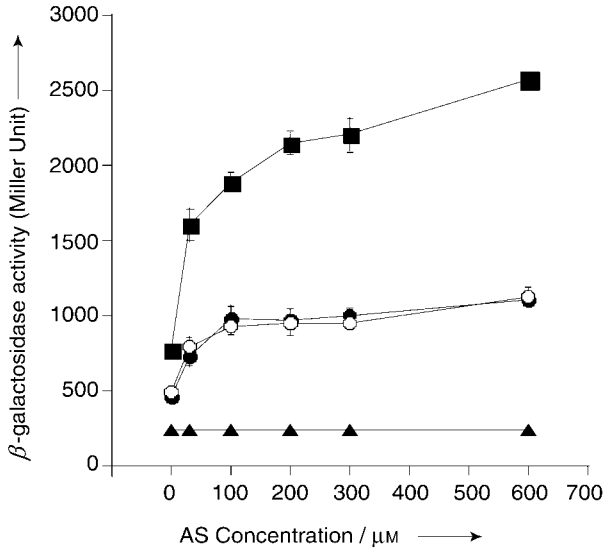


***Figure 4.*** *Induction of the* virB:lacZ *gene by LZ(TM2) – LKR(G665D). Derivatives of* A. tumefaciens *A348 – 3/pSW209Ω carrying different vVirA constructs were cultured and assayed for* vir *gene expression as described. (▲) = pYW15a, negative control; (●) = pYW39, 6 × His – LKR(G665D); (○) = pYW40, LKR(G665D); (■) = pYW50, 6 × His – LZ(TM2) – LKR(G665D)*

### Construction of LZ(0/3/4) – LKR chimera

The linker coiled-coil-like motifs appear not to be a particularly stable dimerization domain (Figure 3), or to be primarily responsible for the induction of *vir* gene expression by enforced VirA homodimerization (Figure 4). In particular, the absence of interhelical salt bridges suggests that Helix C can adopt coiled-coil interactions through several potential hydrophobic interfaces. Interface I uses the typical ***a*** and ***d*** residues to form the dimer interface, while interface II and interface III position either the ***a*** and ***e*** residues or the ***d*** and ***g*** residues, respectively, at the dimer interface. As shown in Figure 5, interfaces II and III can each be achieved by a conrotatory 51° turn of each helix from its interface I position. Relative to interface I, interface II is the result of a conrotatory clockwise rotation of both helices, whereas interface III is the result of a similar counterclockwise rotation of the two helices. This observation suggests that the relative orientation of the two subunits might not be fixed, and that Helix C may be able to adopt several flexible interfaces. The relative phase of each interface may be critical for signal transmission.

To test the functional role of the redundant interfaces within Helix C, the leucine zipper of GCN4 was fused directly with LKR(G665D) (aa 293 – 829; Figure 5). A four-residue deletion (aa 312 – 315) is seen in some of the wild-type strains, therefore a flexible adapter with either 0, 3 (DAL), or 4 (DALK) amino acids was inserted at the fusion junction to give LZ(0), LZ(3), and LZ(4), respectively. As shown in Figure 5, these adapter sequences enforced the alignment of the hydrophobic interface of the GCN4 leucine zipper peptide with each of the three Helix C coiled-coil interfaces. The large folding free energy ($> 10$ kcal mol$^{-1}$) of the GCN4 leucine zipper peptide dimer[14] has been shown to distort coiled-coil dimers to alternate interface preferences in the region of the fusion junction,[15] and the same is predicted for this linker dimerization domain. Therefore, in LZ(0), the hydrophobic interface I of Helix C would be biased towards the phase of the LZ dimer interface. Correspondingly, in LZ(3) and LZ(4), interface II and interface III of Helix C should be aligned correspondingly with the GCN4 dimer interface.

The $P_{N25}$-based broad-host-range expression vector pYW15a[12] was used to construct and express different VirA truncation mutants in *Agrobacterium*, as outlined in Figure 2. Western analysis with either antiVirA or anti-RGSHis antibodies confirmed that the VirA mutants were the expected size and occurred at

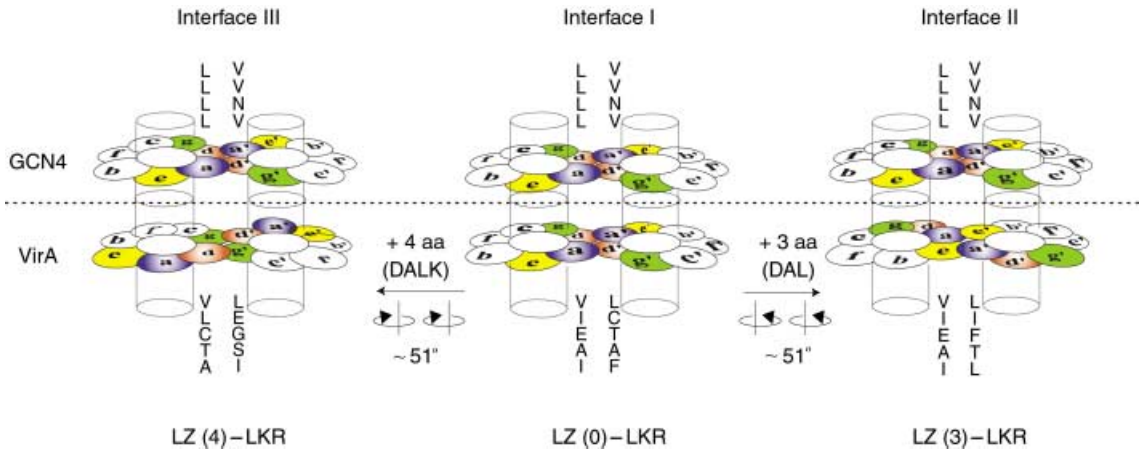


***Figure 5.*** *Helical wheel representations of LZ(0) – LKR, LZ(3) – LKR, LZ(4) – LKR fused with GCN4. The seven-residue (heptad) repeats are represented by the letters* ***a*** *to* ***g****, and the dimer designations, LZ(0), LZ(3), LZ(4), indicate the number of amino acids of the adapter inserted between the GCN4 and VirA LKR sequences. Three potential interfaces could be adopted in a continuous helix dimer. Interface II and Interface III are achieved by conrotatory rotation of the helices of Interface I by one seventh of a turn counterclockwise or clockwise, respectively.*

approximately equivalent levels (Figure 6). Similar results are observed for the expression of LZ(0/3/4) – LKR$^{WT}$ (WT = wild type) constructs (data not shown). Preliminary analysis found all of these VirA constructs fractionated with the *Agrobacterium* inner membranes.

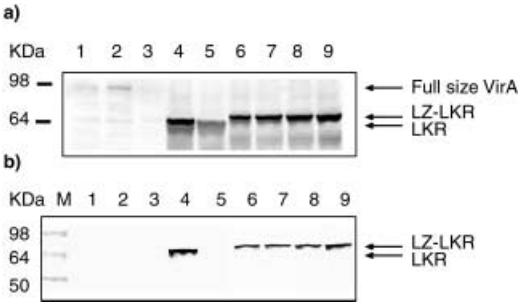


***Figure 6.*** *Expression of various VirA constructs in* A. tumefaciens *detected by antiVirA (A), and by anti-RGSHis (B). Lanes: 1: A348, VirA$^{WT}$; 2: A348 – 3/pMutA G665D, VirA(G665D); 3: A348 – 3, VirA deletion; 4 – 9: A348 – 3 carrying various VirA constructs. 4: pYW39, 6 × His – LKR(G665D); 5: pYW40, LKR(G665D); 6: pYW50, LZ(TM2) – LKR(G665D); 7: pYW59, LZ(4) – LKR(G665D); 8: pYW60, LZ(3) – LKR(G665D); 9: pYW61b, LZ(0) – LKR(G665D).*

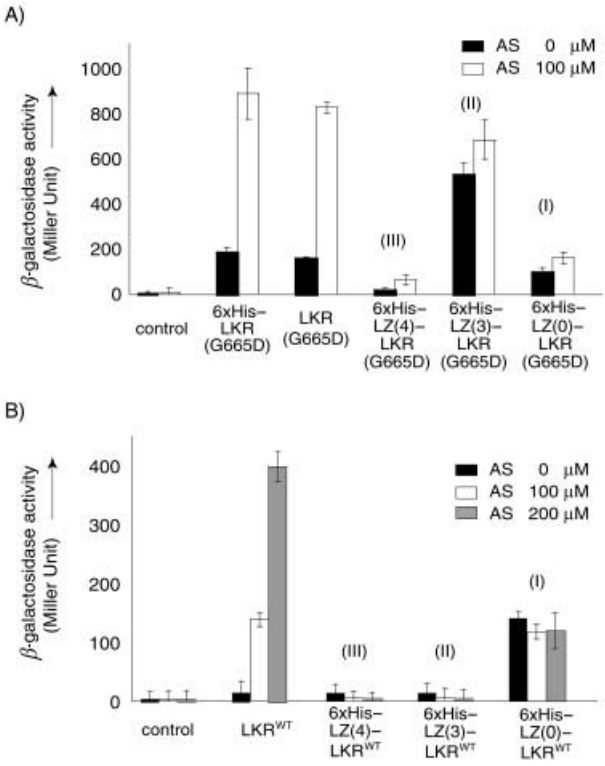


***Figure 7.*** *Induction of the* virB:lacZ *gene. Derivatives of* A. tumefaciens *A348 – 3/pSW209Ω carrying different VirA constructs were cultured in induction medium with either no AS or 100 μM AS and assayed for* vir *gene expression as described. A) LZ(0/3/4) – LKR(G665D): Strains: pYW15a, VirA deletion, control; pYW39, 6 × His-LKR(G665D); pYW40, LKR(G665D); pYW59, 6 × His – LZ(4) – LKR(G665D); pYW60, 6 × His – LZ(3) – LKR(G665D); pYW61b, 6 × His – LZ(0) – LKR(G665D). B) LZ(0/3/4) – LKR$^{WT}$: Strains: pYW15a, VirA deletion, control; pYW34, LKR$^{WT}$; pYW65, 6 × His – LZ(4) – LKR$^{WT}$; pYW66, 6 × His – LZ(3) – LKR$^{WT}$; pYW67, 6 × His – LZ(0) – LKR$^{WT}$.*

### *vir* gene activation by LZ(0/3/4) – LKR fusions

In the absence of AS, the base level of *vir* gene expression by LZ(3) – LKR(G665D) is about three times higher than that by either LKR(G665D) or 6 × His – LKR(G665D) (Figure 7 A). LZ(0) – LKR(G665D) gives intermediate expression, with activity similar to LKR(G665D), whereas LZ(4) – LKR(G665D) is essentially inactive. These results are consistent with the proposal that Helix C is both critical for *vir* gene activation and is sensitive to distortions of the homodimer interface. More importantly, as compared to the nonengineered LKR(G665D) whose activity can be further increased by approximately fourfold by 100 μM AS, these fusion proteins no longer respond to AS induction. It appears, therefore, that interface II (represented by the LZ(3) – LKR(G665D)) is able to mimic the highly activated conformation of Helix C, while interfaces I and III (represented by LZ(0) – LKR(G665D) and LZ(4) – LKR(G665D) respectively) preorganize conformations with intermediate activity and noninducing conformations.

The $LKR^{WT}$ constructs show a similar phenotype. As shown in Figure 7B, LZ(3) – $LKR^{WT}$ and LZ(4) – $LKR^{WT}$ are inactive, while LZ(0) – $LKR^{WT}$ shows constitutive *vir* gene expression. In the absence of AS, this fusion protein showed approximately 10 times higher basal activity than $LKR^{WT}$ alone and did not respond to AS. The active interface in the wild type, LZ(0), is therefore different from the active interface (LZ(3)) in the construct that contains the single point mutation G665D in the kinase domain.

## Discussion

The VirA/VirG two-component signal transduction elements set into motion the complex cascade of events necessary to simultaneously coordinate hosts from two separate kingdoms.[16, 17] This system is unusual in that a number of small molecule mediators are known to regulate activity and the histidine kinase VirA is assigned as the central transducer of all the input signals.

The linker domain of VirA, which connects the transmembrane elements with the kinase domain, has been implicated repeatedly as critical for signal perception. Deletion of the N-terminal portion of the linker, aa 283-304,[18] and introduction of five independent VirA mutations within this domain[19] led to defective *vir* gene activation. As assigned in our conformational analyses, three of these mutations, G268D, T284M, and R289Q, are within the Helix TM2 region (aa 262 – 288). The other two are within Helix C (A318T; aa 293 – 327) and just outside Helix C (S333L).

Although few structures of the histidine kinase family are available, a combination of computational methods and intuitive comparisons has allowed several structures to be successfully predicted from the primary sequence.[20, 21] Our methods identified Helix TM2 as an inner-membrane-spanning amphipathic helix conserved in all chemoreceptors and a large number of sensor histidine kinases, including VirA.[19] This conservation, together with the mutagenesis data mentioned above, implicated Helix TM2 in signal input.[19] However, results from this study suggest that a critical functional role of Helix TM2 is to anchor homodimerization. The VirA truncation mutant LKR(G665D) lacks Helix TM2 and is still AS responsive, but its enforced dimerization with the soluble GCN4 leucine zipper domain substantially increases *vir* gene expression. This fusion protein retained a level of induction similar to LKR(G665D), which indicates that dimerization of VirA is important for VirA activation but dimerization alone is not sufficient to fully activate VirA.

Sequence analysis of Helix C suggested three separate energetically accessible coiled-coil interfaces. Three fusion proteins, LZ(0) – LKR, LZ(3) – LKR, and LZ(4) – LKR, were engineered to bias the relative stability of each interface. Determination of the extent to which the conformations were successfully biased will require detailed structural analysis, but the observed phenotypes of the chimeras are consistent with some degree of bias. The LKR(G665D) construct showed constitutive and AS-inducible *vir* gene expression, whilst the bias in favor of interface I (LZ(0)) gave moderate activity that was no longer AS responsive. When the bias was shifted to interface II (LZ(3)) or interface III (LZ(4)) by conrotatory displacement through one seventh of a whole helical rotation, clockwise or counterclockwise respectively, highly activating and nonactivating constructs resulted. Most importantly, the level of *vir* gene activation by all three fusion proteins was independent of AS induction.

The LZ – $LKR^{WT}$ fusion proteins show lower relative activity than the mutants, but the same constitutive phenotype and lack of AS responsiveness. More interestingly, the active construct possesses stability biased towards interface I, the interface that shows moderate activity in the LZ – LKR(G665D) constructs. Since LKR(G665D) carries a mutation that resides just outside the ATP-binding G2 block, the local conformation of the catalytic domain appears to be altered and sufficiently so to give partial constitutive activity. Clearly the kinase and linker domains are intimately coupled. This coupling appears to be transmitted through the phase of the coiled-coil interface. Therefore, mutations in the kinase domain alter the linker interface required for activation of VirA.

Combined, these data allow us to propose the ratchet model for VirA activation shown in Figure 8. In this model, VirA exists as a homodimer in three conformations, "OFF", "PARTIALLY ON", and "ON". The off form of VirA, interface III, probably serves as the ground state. AS perception allows a shift from interface III to interface I (partially on), interface II (fully on), or possibly any one of the three disrotatory interfaces that have not yet been constructed. This conformational shift propagates through the

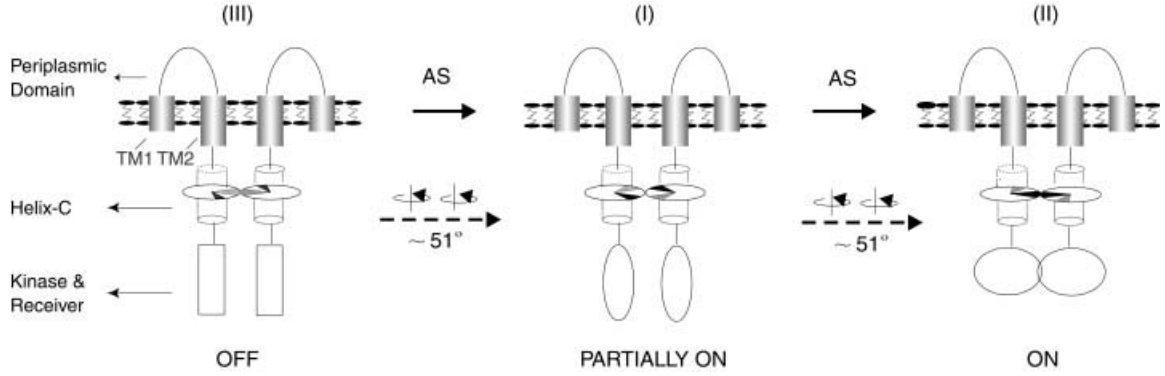


***Figure 8.*** *Ratchet model for signal-mediated VirA activation. The Helix C interfaces III, I, and II are coupled to the kinase domain to give OFF, PARTIALLY ON, or ON functional activity, respectively.*

linker to the kinase domain, by rotation and/or by a positional shift, to regulate output activity. While several mechanisms have been proposed for signal activation in the homologous chemoreceptors,[22] a similar interface shift in coiled coils has been implicated in at least one possible model for the bacterial Tar receptor[15] as well as for myosin heptad dynamics.[23] Such interface shifting has also been observed directly in the crystal structure of the low-pH form of influenza hemaglutinin, $TBHA_2$.[24]

As stated above, signal perception events through VirA integrate at least four separate signals. First, VirA operates only within a limited temperature range.[25] Second, monosaccharides, including glucose and arabinose, are perceived through the periplasmic domain of VirA but recognition requires the bacterial sugar-binding protein, ChvE.[11, 26] The ChvE/sugar complex affects the sensitivity, specificity, and maximal response of the system, and point mutations and/or deletions in the periplasmic domain of VirA, both in the putative ChvE binding domain and distal to it, can abolish the effect of sugars.[6, 11, 18, 26–28] Third, the expression of *virG* itself can be induced by low pH values, a phenotype partially controlled by the bacterial protein ChvD.[29] However, expression of *virG* from constitutive promoters does not create a pH independent inducing system,[12, 30] which indicates that other pH-related inputs into the system must occur. Truncations of VirA suggest that the cytoplasmic domain alone is still responsive to pH value,[18] which further suggests that other cytoplasmic proteins are required for perception. Finally, the most critical xenognosins are plant phenolics (for example, AS). Despite the recent experiment results that suggest that VirA may be a direct phenol sensor,[31] affinity labeling,[32] structural analyses of inducing phenolics[33] and recent genetic evidence have suggested that phenol activation also requires bacterial chromosomal proteins.[34] Given that the same cytoplasmic versions of VirA that respond to pH value also respond to phenols,[18] such phenol-sensing proteins are likely to be cytoplasmic.

With the accumulating evidence that several bacterial proteins may be involved in mediating signal perception by VirA, it is possible that VirA has evolved to interrogate these proteins for signal transmission. In the context of the "ratchet model" proposed here, such interrogation may be accomplished by direct binding to the coiled-coil interface of the VirA linker, and can thus modulate both interface stability and the resulting kinase activity. The unique ability of the Ti plasmid to simultaneously pathogenize hosts from different kingdoms would therefore be built on the evolutionary adaptability afforded by coiled coils which allow elements from the *Agrobacterium* proteome to be selected for their ability to bind to and modulate the very first and most critical stage of pathogenic integration. Such a model would allow the Ti plasmid signaling events to function uniquely within, and even coevolve with, its host proteome.

## Experimental Section

**Strains, plasmids, and oligonucleotides**: *A. tumefaciens* strain A348-3 containing pTiA6NC with *virA* deleted[32] was used to express various *virA* alleles. *A. tumefaciens* strains were grown in Luria broth (LB) medium or AB minimal medium[35] at 28 °C. *E. coli* strain XL1-Blue (Stratagene) was grown in LB medium at 37 °C and used as the cloning host. Plasmids constructed in this study are listed in Figure 2.

**Construction of plasmids**: To construct the LZ – VirA chimera, a DNA cassette encoding the leucine zipper of GCN4 and an N-terminal GGCGG linker was constructed by mutually primed DNA synthesis. The amino acid sequence of this cassette was: GGCGG KVKQ L EDKVEE L LSKNYH L ENEVAR L KKLV. A flexible linker that contained amino acids Gly-Gly-Cys-Gly-Gly was incorporated into the protein at the N terminus to enable disulfide cross-linking for in vitro analysis. 15 pmol of each oligo were mixed and amplified by polymerase chain reaction (PCR): GCN4a, 5′-GGA GGT TGC GGA GGT AAG GTC AAG CAG CTG GAG GAC AAG GTC GAG GAG CTG CTG TCG AAG AAC TAC CAT C-3′; GCN4b, 5′-GAC CAG CTT CTT CAG TCT TGC GAC CTC GTT CTC CAG ATG GTA GTT CTT CGA CAG CAG CTC CTC GAC CTT G-3′; GCN4c, 5′-GC *GAG CTC* GGA GGT TGC GGA GGT AAG-3′ (Sac I); GCN4d, 5′-GC *GCA TGC* GAC CAG CTT CTT CAG TCT-3′ (Sph I). The expected approximately 110-base-pair (110-bp) PCR product was TA cloned (cloned into a vector containing complementary 3′-thymidine overhangs) to give pYW26. In order to minimize the possible error that could be introduced into the fusion constructs by PCR amplification, a LZ(TM2) – linker fusion was first constructed by fusing the GCN4 leucine zipper cassette with the VirA linker domain (aa 285 – 471). Recombinant PCR was used to construct the fusion proteins. Two primary PCR reactions were performed separately. Reaction A used plasmid pYW26, which contained the GCN4 leucine zipper cassette, as the template. The sense and antisense primers were, respectively, XMN, 5′-GC *GAAGATATTC* GA *GAGCTC* GGA GGT TGC-3′ (Xmn I; Sac I) and FS0, 5′-AAA CGC CGC GCT AAC CAA TCG ACC AGC TTC TTC AG-3′, which has 15 bp complementary to the C-terminal sequence of the GCN4 cassette and 20 bp complementary to the N-terminal sequence of the VirA linker. Reaction B used plasmid pVRA8 which contains the $virA^{WT}$ gene as the template. The sense and antisense primers were, respectively, FS0a, 5′-CAA GAC TGA AGA AGC TGG TCG ATT GGT TAG CGC GG-3′ with 15 bp complementary to the N-terminal sequence of LKR and 20 bp complementary to the C-terminal sequence of GCN4, and NAE, 5′-GC *GGTACC* TCC GCC GGC AAG TGT AC-3′ (Kpn I). The PCR products from reactions A and B were an approximately 130-bp fragment encoding the GCN4 cassette and a roughly 580-bp fragment encoding the linker domain of VirA. The mixture of the two primary PCR products was used as the template (1/20 (v/v) each) in the secondary PCR reaction. Two outside primers XMN and NAE were used to amplify the fusion. The resulting PCR product (about 700 bp) which encodes the LZ – linker fusion protein, was subsequently digested with Sac I and Kpn I and cloned into pYW15b[12] to give pYW33, which contains the $P_{N25}$—6 × His – LZ(TM2) – linker fusion protein. The complete LZ(TM2) – LKR(G665D) fusion product was then constructed by swapping restriction fragments that contained the rest of the LKR(G665D) into this LZ – linker construct. An approximately 800-bp EcoR I/Kpn I fragment of pYW33 containing the $P_{N25}$—6 × His – LZ(TM2) – linker fusion construct was first moved into pBluescript II KS (+/-) (in which BstE II is a unique site) to give plasmid pYW42. The construction of the LZ(TM2) – LKR(G665D) fusion protein was achieved by cloning a BstE II/Kpn I fragment (roughly 2 kb) from pMutA G665D into the BstE II/Kpn I sites of pYW42 to give pYW49. Finally, a EcoR I/Kpn I fragment (approximately 2.7 kb) that contained $P_{N25}$—6 × His – LZ(TM2) – LKR(G665D) was cloned into pYW15a to give pYW50.

Construction of the VirA-Linker (VirA, aa 285 – 471) was achieved by a domain swap into pYW42. An EcoR I/Kpn I fragment (about 1.6 kb) of

pYW39[12] was moved into pYW49 to give pRG13. The Sac II/BstE II fragment (about 500 bp) of pRG13 was swapped into pYW42 to give pRG14. An approximately 550-bp Kpn I fragment containing VirA – Linker (VirA aa 285 – 471) from pRG14 was cloned into pYW15b[12] to give pRG15, which contained the $P_{N25}$—6 × His—linker construct.

Construction of LZ – LKR with 0, 3, or 4 amino acid adapter sequences at the fusion junction, designated LZ(0/3/4) – LKR(G665D) herein, was achieved by a domain swap into pYW49. Specifically, LZ(0/3/4) – LKR(G665D) was achieved by swapping LZ(0/3/4) – linker fragments with the LZ – linker fragment of pYW49. LZ(0/3/4) – linker was constructed by recombinant PCR as described above. The fusion junction in these constructs was from aa 294 of the VirA linker instead of aa 285 as for the LZ(TM2) – linker construct. The primers used in the primary and secondary PCR for construction of LZ(0/3/4)—linker were FS0A, FS0S, FS3A, FS3S, FS4A, FS4S, XMN, and NAE. The templates for the primary PCRs were pYW49 and pVRA8. For each fusion construct, the two primary PCR products, a roughly 130-bp fragment and an approximately 540-bp fragment, were gel-purified to remove excess primary PCR primers and then used as templates for the secondary PCR reactions. The three secondary PCR products (about 660 bp) were TA cloned into pCR2.1 or digested with Sac II and BstE II and cloned into pYW49 to give plasmids pYW56, pYW57b, and pYW58. An EcoR I/Kpn I fragment (about 2.7 kb) that contained $P_{N25}$—6 × His – LZ(0/3/4) – LKR(G665D) were cloned into pYW15a to give pYW61b, pYW60, and pYW59.

Construction of LZ(0/3/4) – $LKR^{WT}$ was achieved by swapping a BstE II/Kpn I fragment (about 1.3 Kb; encoding the C-terminal portion of $VirA^{WT}$) from pYW21 into pYW58, pYW57b, and pYW56 respectively, to replace the corresponding fragment of VirA(G665D). A EcoR I/Kpn I fragment (about 1.7 kb) that contained the $P_{N25}$—6 × His – LZ(0/3/4) – $LKR^{WT}$ constructs was then moved into pYW15a to give pYW67, pYW66, and pYW65. All construct sequences were confirmed with the ABI PRISM Dye Primer Cycle Sequencing kit (Perkin Elmer).

**Immunoblot analysis**: SDS-PAGE was conducted in 14% or 8 – 16% polyacrylamide gels (Novex), and the proteins were electroblotted (Electro-blotter, Bio-rad) onto nitrocellulose membrane (Bio-rad). The VirA fragments were either probed with polyclonal antibody raised against the VirA kinase domain or with monoclonal anti-RGSHis purchased from Qiagen. The proteins were either visualized by incubation with alkaline phosphatase-conjugated goat antirabbit secondary antibody (Bio-rad; 1:1500, for antiVirA) or antimouse secondary antibody (Qiagen; 1:1000, for anti-RGSHis) followed by development with 1-Step NBT/BCIP (Pierce) as the substrate.

***vir* gene induction assay**: *Agrobacterium* strains were grown in LB medium (30 mL) supplemented with appropriate antibiotics to an optical density of $OD_{600} = 0.4 – 0.6$. The bacteria were harvested by centrifugation at 6000 rpm (GSA rotor) for 15 minutes at 4 °C. After washing with phosphate-buffered saline (PBS) buffer twice, the pellets were stored at − 70 °C. As needed, cells were thawed, diluted to $OD_{600} \sim 0.1$ in the induction medium[32] (1 ml) with glycerol (1.0%), and cultured at 20 °C with shaking at 225 rpm. The β-galactosidase activity was assayed essentially as described by Miller[36] after a 20 hr induction or at the indicated time intervals.

**Computational analysis of VirA sequences**: Multiple sequence alignments were analyzed with the GCG pile-up program (University of Wisconsin). Secondary structure predictions were made by using the PHDsec program.[37]

[1] J. S. Parkinson, E. C. Kofoid, *Annu. Rev. Genet.* **1992**, *26*, 71 – 112.
[2] M. C. Pirrung, *Chem. Biol.* **1999**, *6*, R167 – 175.
[3] M. M. McEvoy, F. W. Dahlquist, *Curr. Opin. Struct. Biol.* **1997**, *7*, 793 – 797.
[4] S. C. Winans, *Microbiol. Rev.* **1992**, *56*, 12 – 31.
[5] S. Q. Pan, T. Charles, S. Jin, Z. L. Wu, E. W. Nester, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 9939 – 9943.
[6] L. S. Melchers, T. J. Regensburg-Tuink, R. B. Bourret, N. J. Sedee, R. A. Schilperoort, P. J. Hooykaas, *EMBO J.* **1989**, *8*, 1919 – 1925.
[7] C. H. Chang, S. C. Winans, *J. Bacteriol.* **1992**, *174*, 7033 – 7039.
[8] J. D. Heath, T. C. Charles, E. W. Nester in *Two-Component Signal Transduction* (Eds.: J. A. Hoch, T. J. Silhavy), ASM Press, Washington D.C., **1995**, pp. 367 – 385.
[9] M. Singh, B. Berger, P. S. Kim, J. M. Berger, A. G. Cochran, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 2738 – 2743.
[10] P. B. Harbury, T. Zhang, P. S. Kim, T. Alber, *Science* **1993**, *262*, 1401 – 1407.
[11] N. Shimoda, A. Toyoda-Yamamoto, J. N. S. Usami, M. Katayama, Y. Sakagami, Y. Machida, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 6684 – 6688.
[12] Y. Wang, A. Mukhopadhyay, V. R. Howitz, A. N. Binns, D. G. Lynn, *Gene* **2000**, *242*, 105 – 114.
[13] B. G. McLean, E. A. Greene, P. C. Zambryski, *J. Biol. Chem.* **1994**, *269*, 2645 – 2651.
[14] J. A. Zitzewitz, O. Bilsel, J. Luo, B. E. Jones, C. R. Matthews, *Biochemistry* **1995**, *34*, 12812 – 12819.
[15] A. G. Cochran, P. S. Kim, *Science* **1996**, *271*, 1113 – 1116.
[16] T. Kunik, T. Tzfira, Y. Kapulnik, Y. Gafni, C. Dingwall, V. Citovsky, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 1871 – 1876.
[17] G. Hansen, M. D. Chilton, *Curr. Top. Microbiol. Immunol.* **1999**, *240*, 21 – 57.
[18] S. C. Turk, R. P. van Lange, T. J. Regensburg-Tuink, P. J. Hooykaas, *Plant Mol. Biol.* **1994**, *25*, 899 – 907.
[19] S. L. Doty, M. C. Yu, J. I. Lundin, J. D. Heath, E. W. Nester, *J. Bacteriol.* **1996**, *178*, 961 – 970.
[20] B. Rost, C. Sander, *J. Mol. Biol.* **1993**, *232*, 584 – 599.
[21] S. A. Benner, D. L. Gerloff, *FEBS Lett.* **1993**, *325*, 29 – 33.
[22] W. G. Scott, B. L. Stoddard, *Structure* **1994**, *2*, 877 – 887.
[23] J. H. Brown, C. Cohen, D. A. Parry, *Proteins* **1996**, *26*, 134 – 145.
[24] P. A. Bullough, F. M. Hughson, J. J. Skehel, D. C. Wiley, *Nature* **1994**, *371*, 37 – 43.
[25] S. Jin, Y. N. Song, W. Y. Deng, M. P. Gordon, E. W. Nester, *J. Bacteriol.* **1993**, *175*, 6830 – 6835.
[26] G. A. Cangelosi, G. Martinetti, E. W. Nester, *J. Bacteriol.* **1990**, *172*, 2172 – 2174.
[27] L. M. Banta, R. D. Joerger, V. R. Howitz, A. M. Campbell, A. N. Binns, *J. Bacteriol.* **1994**, *176*, 3242 – 3249.
[28] S. C. Turk, R. P. van Lange, E. Sonneveld, P. J. Hooykaas, *J. Bacteriol.* **1993**, *175*, 5706 – 5709.
[29] S. C. Winans, R. A. Kerstetter, E. W. Nester, *J. Bacteriol.* **1988**, *170*, 4047 – 4054.
[30] C. Y. Chen, S. C. Winans, *J. Bacteriol.* **1991**, *173*, 1139 – 1144.
[31] S. M. Lohrke, H. Yang, S. Jin, *J. Bacteriol.* **2001**, *183*, 3704 – 3711.
[32] K. Lee, M. W. Dudley, K. M. Hess, D. G. Lynn, R. D. Joerger, A. N. Binns, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 8666 – 8670.
[33] M. E. Duban, K. Lee, D. G. Lynn, *Mol. Microbiol.* **1993**, *7*, 637 – 645.
[34] A. M. Campbell, J. B. Tok, J. Zhang, Y. Wang, M. Stein, D. G. Lynn, A. N. Binns, *Chem. Biol.* **2000**, *7*, 65 – 76.
[35] M. D. Chilton, T. C. Currier, S. K. Farrand, A. J. Bendich, M. P. Gordon, E. W. Nester, *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 3672 – 3676.
[36] J. H. Miller, *Experiments in molecular genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, **1972**.
[37] B. Rost, C. Sander, *Proteins: Struct. Funct. Genet.* **1994**, *19*, 55 – 72.

Received: September 18, 2001 [F 297]