THE GENETICS OF TCV RESISTANCE

A Thesis

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Master of Science in

Biochemistry

By

Katrina Emilee Vaitkunas

Date: 30 April 2003

Approved:

Dr. Krisin K. Wobbe, Advisor

1. **Turnip Crinkle Virus**
2. **Arabidopsis thaliana**
3. **Plant resistance**

Abstract:

Most plants are capable of mounting resistance responses to various pathogen attacks. For a hypersensitive response (HR) to occur, a dominant or semi-dominant resistance (R) plant gene is required to recognize a dominant avirulence (Avr) factor of the pathogen. Three types of *Arabidopsis thaliana*, Dijon-17 (Di-17), Dijon-3 (Di-3), and Columbia-0 (Col-0), are significant in understanding the genetics of *Turnip crinkle virus* (TCV) resistance. It has been shown that three genes are needed for successful resistance to TCV in *A. thaliana*: the dominant R gene *HRT*, the recessive gene *rrt*, and a third gene, *TIP*. Crosses of Di-17 and Di-3 plants, and crosses of Di-3 and Col-0 plants are being analyzed to determine the genotype of the F_1 progeny. Using cleaved amplified polymorphic sequence (CAPS) markers, it is possible to determine the genotype of the progeny compared to the wild-type parents at the *HRT* and *TIP* loci. Additionally, protein analysis tools will be employed to compare the Di-3 and Di-17 *TIP* alleles to determine if there are any significant differences in the protein.

Acknowledgements:

I would like to thank Dr. Kristin K. Wobbe for giving me the opportunity to work in her research group. Her guidance and support over the past few years have been an invaluable part of my progress and education, and I would never have been able to accomplish so much with out her.

I would like to thank Dr. Craig Fairchild for teaching me the art of crossing plants, which was an important skill for this research project.

I would like to thank Dr. Fred Souret and Tom Hammond for providing some of the basis of this project, and for helping to set me up when I first moved into the lab.

I would like to thank Ye Zhan for teaching me numerous lab techniques, including viral inoculation.

I would like to thank my assistants, Anne Skulas, Anne St. Martin, and Anthony Scanzani, for helping me complete essential aspects of my project.

I would like to thank the members of GH 06 for all of their suggestions, trouble-shooting, friendship, and support.

I would like to thank Alexander Clifford for his advice and assistance with the digital camera and troubleshooting computer problems.

I would like to thank my friends and family for their support and patience during this project.

Table of Contents **Page**

List of Figures **Page**

List of Tables **Page**

Introduction:

Plants and Pathogens

In the wild, plants are typically challenged by pathogen attacks caused by a variety of viruses, bacteria, fungi, and nematodes (Baker et al., 1997). There are close to 11,000 plant diseases that have been observed; of these, 120 have been attributed to fungi, 30 have been attributed to viruses, and 8 have been attributed to bacteria (Montesinos et al., 2002). It is common for plants to be resistant to many of the pathogens that they encounter. Pathogens that are capable of causing disease are "the exception, not the rule" (Staskawicz, 2001). Although plants are generally resistant to pathogens, approximately 13% of the world's crop production is lost to disease, creating severe losses in food production (Table 1) (Montesinos et al., 2002).

Host/Disease	Pathogen	Crop loss due to pathogen
Rice	All	15% ($33 billion) loss worldwide between 1988-1990
Rice-nematodes	*Meloidogyne* spp.	10 to 50% losses in China, 1984
Maize-southern corn leaf blight	*Cochliobolus heterostrophus*	$1 billion loss in U.S. epidemic, 1970
Cotton	All	10.5% ($4.3 billion) loss worldwide between 1988-1990
Coffee	All	14.8% ($2.8 billion) loss worldwide between 1988-1990
All major crops	All	13.3% (76.9 billion) loss worldwide between 1988-1990

Table 1. **Crop losses caused by plant disease.** Examples of the severity of crop damage and loss due to pathogen infection. (Adapted from Baker et al., 1997).

In order to resist the pathogen attacks that plants face, a system of defense responses necessary for survival have been developed by plants. These defense responses can be grouped into two main categories: constitutive or inducible (active). Constitutive responses are pre-existing features of a plant that are constantly "turned on," whereas inducible responses are

activated by a pathogenic challenge (Fig. 1) (Lucas, 1998).

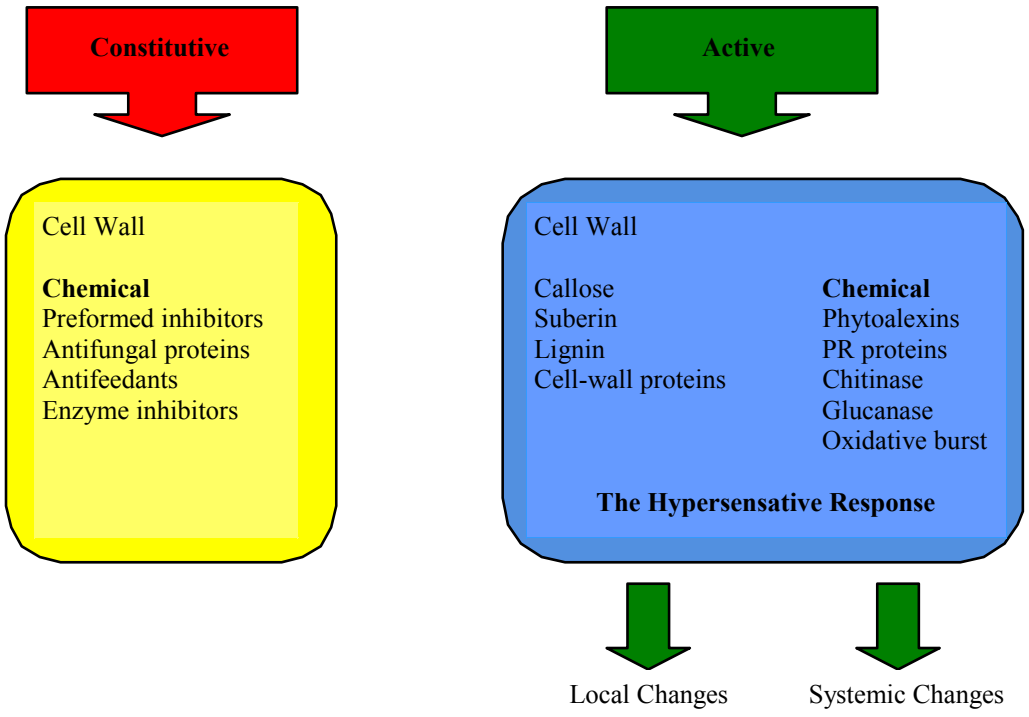


Fig. 1. **Diagram of defense features** that plants possess based on existing (constitutive) or induced (active) defenses. (Adapted from Lucas, 1998).

Most plants are capable of mounting a hypersensitive response (HR) to a pathogen attack, during which necrotic lesions form at the infection site. This programmed cell death is part of an active defense response that prevents the pathogen from systemically spreading throughout the plant (Matthews, 1991). Lesions are able to form on infected leaves within hours or days of the initial challenge (Yi et al., 1998). For instance, *Turnip crinkle virus* lesions are faintly visible at two days post inoculation, and clearly visible at three days post inoculation. In conjunction with the lesion formation during HR, plants also generate signaling molecules to warn nearby cells, strengthen cell walls, and supply anti-microbial enzymes to the defense system (Yi et al., 1998).

The HR is typically triggered by plant recognition of the pathogen. For this recognition process to occur, a dominant or semi-dominant resistance (*R*) gene, corresponding to a dominant avirulence (Avr) factor of the pathogen, must be present in the plant. If either the plant or the pathogen lacks one of the resistance gene products, the plant will be unable to

mount an active defense response, and will eventually become diseased (Flor, 1971). This gene-for-gene relationship, hypothesized by Flor, holds true for many pathogen-plant interactions (Fig.2). Resistance (*R*) genes are typically very specific for the pathogen Avr that they target (Glazebrook, 1999). For example, it has been shown that there is a gene-for-gene relationship between the *avrPto* gene of *Pseudomonas syringae* pv. *tomato* and the *Pto* gene of tomato plants (Martin et al., 1993). However, there are some cases in which *R* gene "dual recognition" has been identified. For instance, *RPP8/HRT* alleles recognize both oomycete parasites and a virus, while the tomato *Mi* gene provides resistance to both nematodes and aphids (Dangl and Jones, 2001).

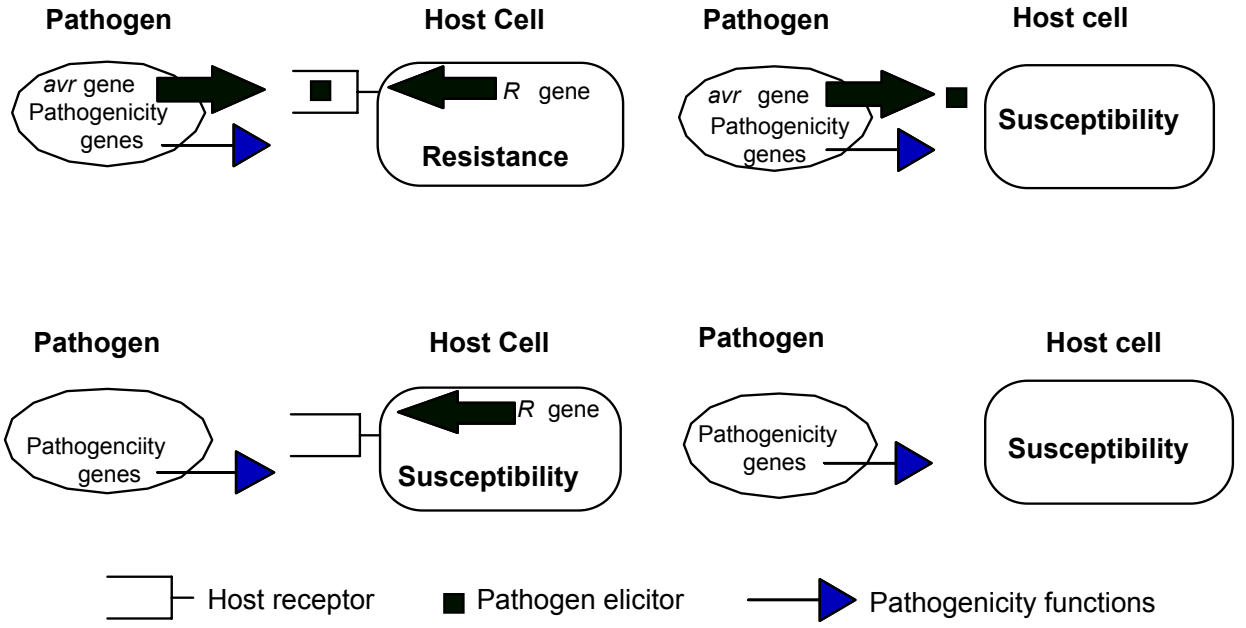


Fig. 2. **Host-pathogen gene-for-gene specificity model.** (Adapted from Lucas, 1998)

Pathogens that elicit an HR activate the systemic acquired resistance (SAR) response (Glazebrook, 1999). Common features of SAR include increased levels of salicylic acid (SA), expression of pathogenesis-related (*PR*) genes, a rapid oxidative burst, cellular decompartmentalizaion, and strengthened plant resistance (Staskawicz et al., 1995, Glazebrook, 1999). It is important that the plant is able to activate SAR, as the plant is unable to store the necessary resistance factors to prevent pathogen attack over an extended period of time (Glazebrook, 1999).

For instance, the gene *nahG* encodes an enzyme that degrades SA. Accumulation of SA

is often observed in plants responding to pathogenic challenge (Cohn et al., 2001). It has been shown that in transgenic plants expressing the bacterial *nahG* gene, the plants do not accumulate SA after pathogen exposure and were more susceptible to all pathogens: both those that would normally cause a resistance response and those that would typically cause disease (Glazebrook, 1999, Cohn et al., 2001). It was also found that these transgenic *nahG* plants did not express *PR* genes in uninoculated plants, nor did they develop SAR (Cohn et al., 2001). Thus, it seems that SA plays an important role in SAR (Cohn et al., 2001).

However, it has been observed that in some cases, gene-for-gene resistance has been able to occur without the presence of HR cell death, and in cases in which the HR is delayed (Yu et al., 1998, Lam et al., 2001). This is the case with *dnd1* mutants. When challenged with avirulent *P. syringae* pathogens, *dnd1* mutants do not demonstrate an HR, although they demonstrate the SAR phenotype, and resistance levels similar to levels presented in wild-type plants (Yu et al., 1998, Glazebrook, 1999). In some cases, as noted with the resistance conferred by the barley *mlo* gene to the fungus *Erysiphe graminis* f.sp. *hordei*, the induction of visible HR follows other resistance factor expression (Lam et al., 2001).

R genes encode proteins that are specific for the Avr proteins, or elicitors, that they target (Nimchuk et al., 2001, Ji et al., 1998). There are five classes of *R* genes: 1) intracellular protein kinases; 2) receptor-like protein kinases with an extracellular leucine-rich repeat (LRR) domain: 3) intracellular LRR proteins with a nucleotide binding site (NBS) and a leucine zipper (LZ) motif; 4) intracellular NBS-LRR proteins with some similarity to Toll and interleukin-1 receptor (TIR) proteins (found in insects and mammals, respectfully) or a coiled-coil (CC) domain; and 5) LRR proteins that code for membrane-bound extracellular proteins (Fig. 3) (Martin, 1999, Baker et al., 1997, Dangl and Jones, 2001). The genes that potentially code for the LRR proteins have been sought out in many plants; the *Arabidopsis* genome contains approximately 150 sequences identified as *NBS-LRR* genes, whereas in a recent study it was found that *Lycopersicon esculentum* Moll (tomato) contains approximately 75 (Nimchuk et al., 2001, Pan et al., 2000).

LRR structures in *R* genes are implicated in mediating protein-protein interactions, and are thought to be responsible for Avr recognition specificity (Fluhr, 2001). Evidence of this interaction between Avr factors and the LRR domain of the *R* genes has been demonstrated. A single amino acid change in the LRR domain of the fungal *Magnapporthe grisea* rice *R* gene

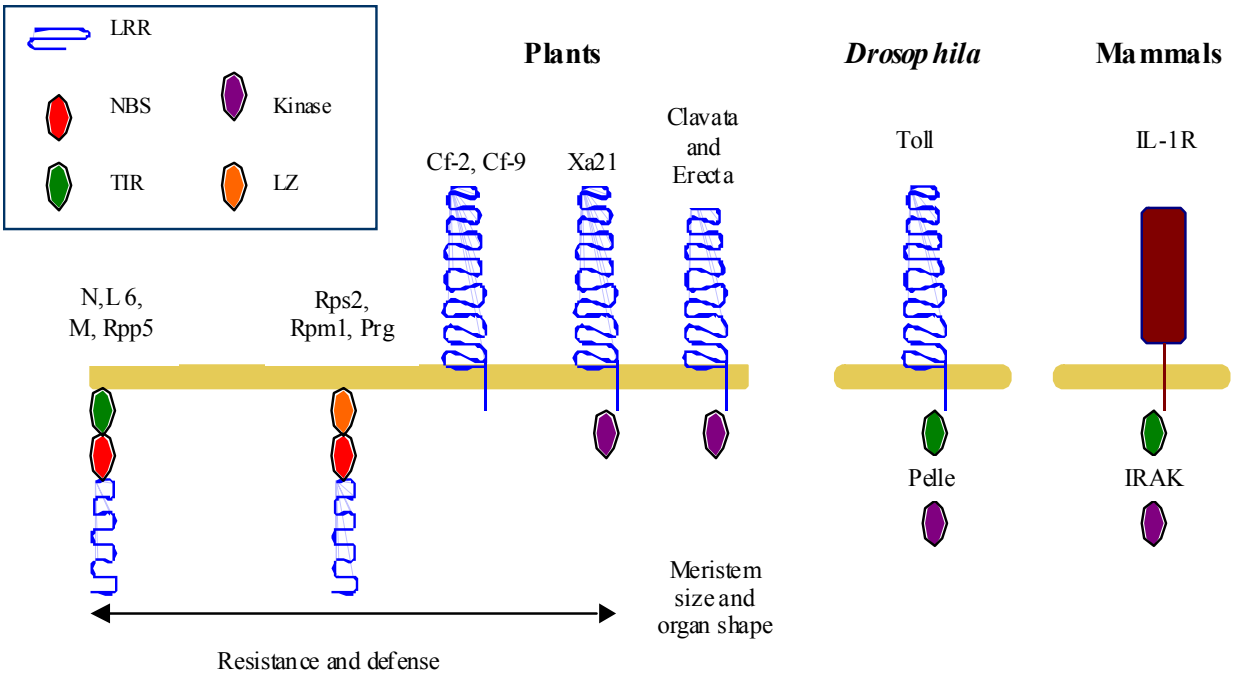


Fig. 3. **Defense and development pathways of plants, insects and mammals.** Genes encoding LRR-NBS motif proteins include *RPS2, PRM1,* and *RPP5* from Arabidopsis, *Prf* from tomato, *N* from tobacco, and *L6* and *M* from flax. The amino terminal domains of *N, L6* and *RPP5* have homology with Toll and IL-1R domains. *Pto* encodes a serine-threonine kinase that shares homology to RAF, IRAK and Pelle kinases. *Cf-2* and *Cf-9* encode large LRR domains that may act as transmembrane receptors. Xa21 encodes an extracellular LRR domain and an intracellular serine-threonine kinase domain. **TIR:** Toll-IL-1R homology domain; **kinase:** serine-threonine kinase; **LZ:** leucine zipper (also known as a coiled-coiled domain (CC) . (Adapted from Baker et al., 1997)

was responsible for the difference between susceptible and resistant alleles (Fluhr, 2001, Bryan et al., 2000). A yeast (*Saccharomyces cerevisiae)* two-hybrid system was subsequently used to demonstrate that the recombinant LRR domain of the resistant allele could interact with the Avr factor, whereas the susceptible allele was only capable of weak Avr factor interactions (Fluhr, 2001, Jia et al., 2000).

The variability in LRR regions of different *R* genes might aid gene-for-gene specificity. Interestingly, the LRR region of an *Arabidopsis R* protein, *RPS5*, appears to be necessary for signaling events after the Avr factors are perceived (Cohn et al., 2001). Two inactive *RPS5* alleles were found to have mutations in their LRR regions (Cohn et al., 2001). One mutant was found to affect the function of other *R* genes, suggesting that *R* genes may also be involved in signaling (Cohn et al., 2001). Since not all *R* genes possess LRR domains, it is not likely that LRR regions are the sole factors controlling gene-for-gene specificity (Cohn et al., 2001).

NBS-LRR *R* genes are the largest class of *R* genes (Dangl and Jones, 2001). These

proteins have a variable number of carboxy-terminal LRRs, and the domains are found in diverse proteins where they are thought to play a role in protein-protein interactions, peptide-ligand binding, and protein-carbohydrate interaction (Dangl and Jones, 2001). NBS-LRR proteins can be further divided based upon their proposed N-terminal structure: one type contains a domain similar to a Toll-interleukin receptor-like domain (TIR-NB-LRR), while others contain proposed coiled-coil domains (CC-NB-LRR) (Dangl and Jones, 2001). Of the two NBS-LRR *R* types, approximately 60% have been identified as TIR, whereas the other 40% or so have been identified as CC (Dangl and Jones, 2001). TIR domains are important to the function of certain *R* genes in both plant and animal immunity, although TIR domains are located at the N terminus of NBS-LRR *R* genes in plants, while TIR domains are located at the C terminus of a single-pass transmembrane receptor in animals (Fluhr, 2001). Plant NBS-LRR domains seem to be an important aspect of *R* gene specificity, as Luck et al. (2000) demonstrated. After recombining LRR and TIR regions of different alleles in the *L* locus in flax, which correspond to the resistance specificities to flax rust, it was determined that the phenotypes of the recombined alleles were unpredictable, most likely due to the "coadaption between different polymorphic regions of the gene (Luck et al., 2000)." Although the phenotypes were often unexpected, they could be classified into three groups: recombinants that conferred resistance identical to that of the LRR donor allele, resistant recombinants that possessed specificity unlike either parent allele, and non-resistant recombinants (Luck et al., 2000). Thus, it seems that specific recombination events have provided plants with the ability to resist the challenge of diverse pathogens.

Serine/threonine protein kinases are another class of *R* genes. This small group of cytoplasmic proteins contain eleven subdomains and as well as a myristylation motif at the amino terminus (Cohn et al., 2001). This myristylation motif, known to be important for subcellular protein localization, does not appear to be required for disease resistance (Cohn et al., 2001). *Pto* encodes a serine/threonine protein kinase. It has been shown that Pto phosphorylates another protein kinase, Pti1, upon Avr recognition, and that a third LRR class protein, Prf, is required for *Pto* function. It is hypothesized that these proteins are members of a signal cascade triggered by Pto (Ji et al., 1998).

The two remaining *R* gene classes both encode proteins that contain extracytoplasmic domains. Tomato Cf proteins, which confer resistance to *Cladosporium fulvum* with

appropriate *avr* genes, are extracellular membrane-bound LRR proteins (Cohn et al., 2001). Cf proteins are glycoproteins with extracytoplasmic regions of LRRs that are joined with a transmembrane region and a cytoplasmic tail (Cohn et al., 2001). The other extracytoplasmic protein class contains an LRR domain, a single transmembrane region, and serine/threonine protein kinase domain. These proteins are known as receptor-like kinases (RLKs) (Cohn et al., 2001). Studies have shown that the LRR regions of RLK proteins are important for ligand binding, whereas it is likely that the kinase domain of the RLK protein is important for the initiation of downstream signaling (Cohn et al., 2001).

It appears that many Avr proteins display little to no protein homology, and many have no apparent function, excepting a few that are virally encoded. Thus, it seems that the R proteins encoded by the plants are the receptors and the Avr proteins encoded by the pathogen are the ligands in the R-Avr interactions (Nimchuk et al., 2001). It has been proposed that the maintenance of Avr proteins by pathogens may be due the role that the proteins play as the virulence factor of pathogens on susceptible plants (Nimchuk et al., 2001). Nimchuk et al. (2001) further hypothesize that perhaps the host R proteins are an intricate defense system designed to prevent the Avr proteins from inhibiting basal and/or specific defense systems put forth by challenged plants. Thus far, only one R-Avr protein interaction has been demonstrated *in vitro* (Holt et al., 2003). This is an interesting finding, as it has been shown that *R* gene function is necessary for plant resistance. The lack of R-Avr direct interaction can be explained by the guard hypothesis (Fig. 4).

The guard hypothesis is based upon the assumptions that there is specificity between the R protein and host target of the Avr protein, and that the interaction between the Avr and host target in the absence of the R protein would enable the pathogen to be virulent (Fig. 4) (Holt et al., 2003). The Pto-AvrPto relationship is a good example. In this relationship, the guard hypothesis suggests that the Prf protein detects and monitors the Pto-AvrPto interaction (Holt et al., 2003). If AvrPto is found to be a virulence factor, somehow targeting the Pto kinase and preventing its ability to activate a defense pathway, Prf might be activated during this interaction. Holt et al. (2003) describe the interaction as follows: “Prf may act similar to a fishing pole with Pto as the bait and AvrPto as the trophy catch.”

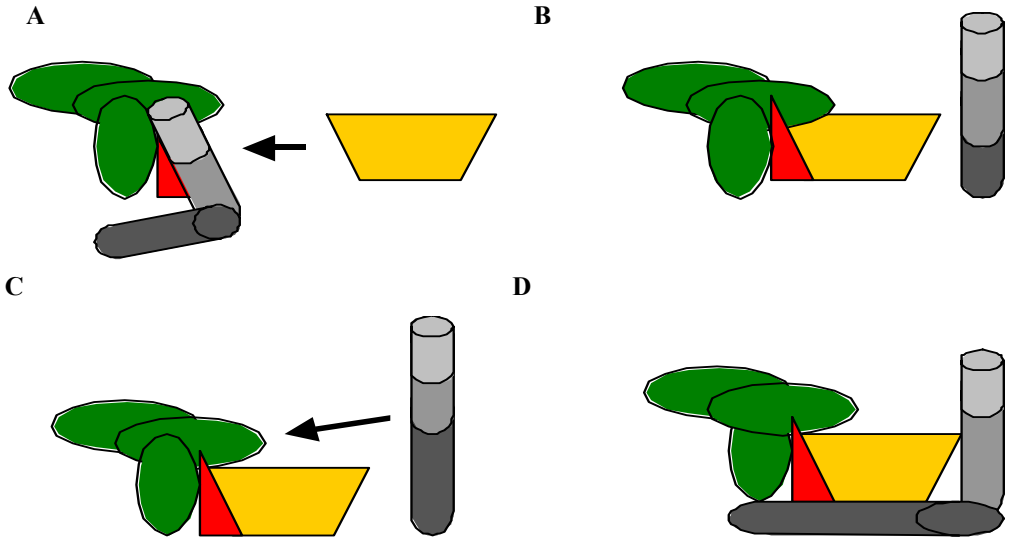


Fig. 4. **The guard hypothesis. A.** A complex of cellular proteins (green), a 'guardee' molecule (red), and an NBD-LRR protein (shaded gray from the N-terminus to the NB and the LRR domains) is targeted by a disease effector (orange). **B.** The pathogen effector binds to its target, resulting in the disassociation and activation of the NBD-LRR protein, causing disease resistance. **C.** Another scenario where the NBD-LRR protein joins the complex only after effector binding. **D.** Activation of the NBD-LRR protein after the entire guard-effector complex is formed. (Adapted from Dangl and Jones, 2001)

Arabidopsis thaliana

Arabidopsis thaliana, commonly known as thale-cress or mouse-ear cress, is a member of the Brassicaceae (mustard) family. *A. thaliana* has been coined the "Drosophila" of plants, due to the fact that it is a useful model system for plants. The *A. thaliana* genome is fairly small in size: 125Mb contained in five haploid chromosomes (Baker et al., 1997, Leutwiler et al., 1984). In 2000, the entire genomic sequence was completed, and all five chromosomes currently have detailed and broad genetic maps (www.arabidopsis.org, 2003). The genome contains a low amount of non-coding DNA, making *A. thaliana* useful for genetic analysis.

Fig. 5. ***Arabidopsis thaliana.*** This picture was taken of an *A. thaliana* plant approximately four weeks old.

Additionally, *A. thaliana* has a rapid life cycle, going from seed to flower in approximately six weeks. Since *Arabidopsis* plants produce many seeds, it is also easy to maintain a stock of seeds. These plants are easily cross-pollinated or

self-pollinated, and given their small physical size (about a tenth of the size of a rice plant), large crops may be grown at one time for lab experiments (Baker et al., 1997, Leutwiler et al., 1984). The varied ecotypes of *A. thaliana,* comprised of over 150 wild isolates, present great opportunities for genetic and pathogenic studies by plant scientists (Baker et al, 1997). Although there are many varieties of *A. thaliana*, the three significant to this particular study are the Dijon-17 (Di-17), Dijon-3 (Di-3) and Columbia-0 (Col-0). These plants are important models in understanding plant resistance, as they exhibit different responses upon *Turnip crinkle virus* (TCV) inoculation.

There are many pathogens that are studied using the *Arabidopsis thaliana* plant model. A few examples of fungi, bacteria, and viruses that challenge Arabidopsis ecotypes are given in Table 2.

Disease	**Pathogen**
Fungal diseases	
Downy mildew	*Peronospora parasitica*
Dark leaf spot	*Alternaria brassicae*
Powdery mildew	*Erysiphe cruciferarum; E. cichoracearum*
Leaf mold and leaf spot	*Cladosporium* sp.
Bacterial diseases	
Black rot on cruicifers	*Xanthomonas campestris* pv. *Campestris*
Bacterial speck on cruicifers	*Pseudomonas syringae* pv. *Maculicula; P. syringae* pv. *tomato*
Viral diseases	
Mild stunting	*Tobacco mosaic virus*
Mild stunting and desiccation	***Turnip crinkle virus***
Vein clearing and chlorotic spots	*Cauliflower mosaic virus*

Table 2. ***Arabidopsis*** **pathosystems.** Examples of fungal, bacterial and viral diseases, and the pathogens that cause them. (Adapted from Baker et al., 1997)

Turnip Crinkle Virus

Carmoviruses are single-component, icosahedral, positive-sense genome RNA viruses. The genome is a little more than 4kb, making these RNA viruses some of the smallest known to

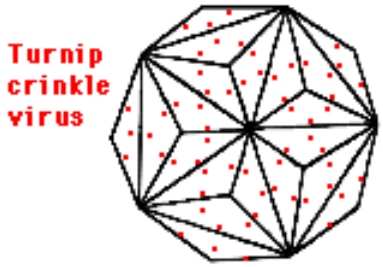


Fig. 6. **Structural representation of TCV.** This is a drawing of the icosahedral capsid containing 180 subunits of the 38kD protein. (From http://www.tulane.edu/~dmsander/WWW/335/335Structure.html, 2003)

infect eucaryotic hosts (Morris and Hacker, 1994). Members of the carmovirus family include *Carnation mottle virus* (CarMV), *Melon necrotic spot virus* (MNSV), *Cowpea mottle virus* (CMeV), *Saguaro cactus virus* (SCV), and *Turnip crinkle virus* (TCV) (Fig. 6) (Morris and Hacker, 1994). The carmoviruses can be found throughout the world, and are generally known to cause mild asymptomatic infections on a relatively small number of natural hosts (Morris and Hacker, 1994). The transmission vector for many of the carmoviruses are beetles, although irrigation water and soil have also been sited as possible viral carriers. *Turnip crinkle virus* has been associated with beetles as its viral vector, despite the lack of widespread infection in the wild. TCV is different from the other carmoviruses in that it has the ability to infect a wide range of experimental hosts, including *Arabidopsis* and *Brassica* (ie. mustard, cabbage, broccoli, cauliflower, kale, etc.) species. In the aforementioned species, it is possible to detect an accumulation of up to 0.5% of the fresh weight of plant tissue, making these hosts extremely important for TCV experimentation (Morris and Hacker, 1994).

The genome of *Turnip crinkle virus* contains 4,054 nucleotides. There are five proteins encoded in the TCV genome: proteins p28 and p88 are involved in viral replication, proteins p8 and p9 facilitate cell-to-cell movement, and protein p38, is the 38-kD capsid protein (Fig. 7) (Carrington et al., 1989, Zhao et al., 2000). The carmoviruses consist of a T=3 icosahedral capsid containing 180 subunits of the 38kD coat protein. The coat protein contains three distinct domains: R (internal N-terminal), S (shell), and P (projecting C-terminal) (Carrington et al., 1987, Morris and Hacker, 1994). R is a basic domain that extends into the virion particle, possibly interacting with the viral RNA, and is linked to the S domain by a connecting arm (Carrington et al., 1987, Morris and Hacker, 1994). The virion shell is made up of the S domain, which is attached to P by way of a hinge. P extends from the virion surface, projecting the carboxy-terminus (Morris and Hacker, 1994).

Of the carmoviruses, TCV is the only one that has been shown to support the replication of small subviral RNAs with a helper virus in the challenged plant. These small subviral RNAs

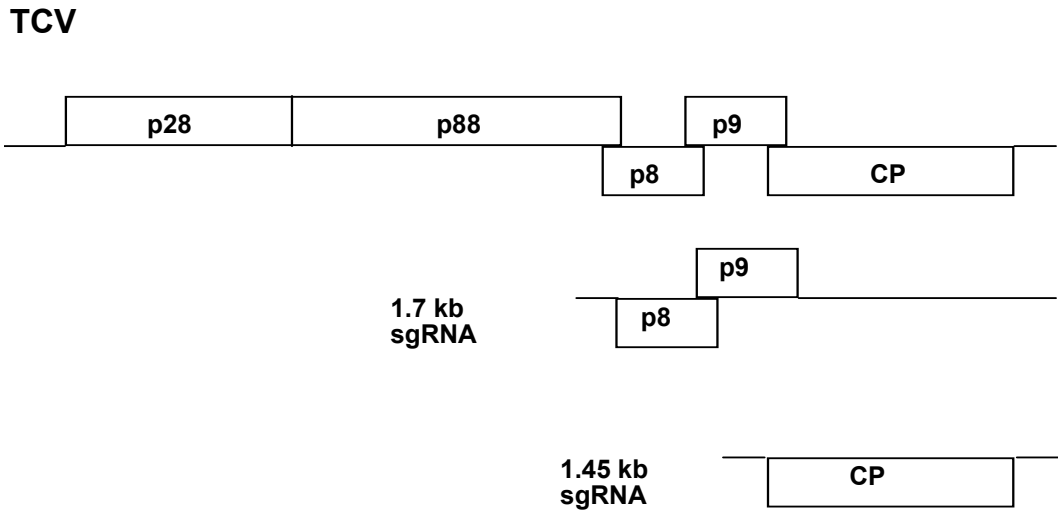


Fig. 7. **Genomic organization of TCV.** The open reading frames (ORFs) are denoted. (Adapted from Morris and Hacker, 1994)

can be placed into one of three categories: satellite RNAs, defective-interfering RNA, and chimeric RNA (Morris and Hacker, 1994). Satellite RNAs, RNAs D and F, are nonviral in origin. These small subviral RNAs do not appear to affect viral function or virulence. Defective-interfering RNA (RNA G) appears to come from the TCV sequence itself. RNA G has the ability to intensify viral symptoms, although it also affects the replication ability of the helper virus. The chimeric RNA (RNA C) is created from both viral and nonviral sequences. Research has suggested that RNA D and the TCV genome recombines by way of a template switch mediated by the viral RNA polymerase. Analysis has shown that the 3' half of RNA C greatly resembles the 3' end of the TCV genome, while the 5' half is almost identical to RNA D. RNA C, like RNA G, is capable of intensifying the symptoms brought on by TCV infection, although it does not seem to interfere with the helper virus replication activity (Morris and Hacker, 1994).

Arabidopsis and Turnip Crinkle Virus

Many ecotypes of *Arabidopsis* are susceptible to TCV, but in 1992, it was shown that Dijon-0 (Di-0) was resistant (Simon et al., 1992). Two isolates of the Di-0 line, Di-3 and Di-17, were later identified as useful ecotypes for research in plant-pathogen studies (Dempsey et al., 1993). Col-0 and Di-3 plants are completely susceptible to TCV infection. Challenge by TCV does not induce an HR, and although PR genes are expressed, they appear much later and

to a lesser extent than in resistant plants. Col-0 and Di-3 plants become systemically infected, showing severe crinkling and stunting, and are dead within three weeks. Conversely, Di-17 plants are capable of expressing an HR in response to TCV challenge. Di-17 plants developed necrotic lesions on the inoculated leaves, and PR genes are found to be present in both uninoculated and inoculated leaves (Dempsey et al., 1993). Furthermore, Di-17 plants typically display no further symptoms of viral infection.

The resistance ability of these three ectotypes of *Arabidopsis* plants is determined by three genes that have been suspected or proven to act in a gene-for-gene resistance manner to TCV: the dominant *R* gene, *HRT* (Hypersensitive response to TCV), a recessive gene, *rrt* (regulation of resistance to TCV), and *TIP* (TCV interacting protein). The three ecotypes, Di-17, Di-3 and Col-0, have different combinations of these genes, thereby causing the plants to have different resistance responses (Table 3).

Di-17	**Di-3**	**Col-0**
HRT	No HRT	No HRT
rrt	rrt?	RRT
TIP	TIP*	TIP

Table 3. **Resistance genes of *Arabidopsis thaliana* species.** * denotes a variation in the gene when compared to the genetic sequences of Di-17 and Col-0 *TIP*.

The dominant *R* gene to TCV is *HRT*. The *HRT* gene, located on Chromosome 5, encodes a 105 kD protein that includes a leucine zipper in the N-terminus, an NBS, and a LRR at the C-terminus. *HRT* is required for TCV resistance and formation of the HR, although it is incapable of conferring total resistance to the plant (Cooley et al., 2000). Progeny of crosses between resistant and susceptible plants were studied by Kachroo and colleagues. All of the HRT^- plants developed systemic infections, but some HRT^+ plants were also susceptible to infection (Kachroo et al., 2000). If *HRT* was the only gene necessary to confer TCV resistance to *Arabidopsis* plants, then all of the HRT^+ plants should have remained resistant. Statistical analysis of the resistant and susceptible progeny further suggested the requirement for a recessive gene. This genes was called *rrt* (regulation of resistance to TCV).

Although the structure and function of the *rrt* gene have not been determined, two different mechanisms have been proposed by which *rrt* might function to confer resistance to TCV infection. The *rrt* allele in Di-17 plants might encode a protein that cannot suppress an

active resistance pathway, or the *rrt* allele might encode a protein that has the ability to interact with HRT in some manner, be it directly or indirectly (Kachroo et al., 2000).

The third gene, *TIP*, is known to interact with the TCV coat protein and is hypothesized to trigger the HR. Mutants in the N terminus of the TCV CP do not induce HR and do produce systemic disease on plants that are resistant to the wild-type virus (Zhao et al., 2000, Ren et al., 2000). These same mutations have been shown to abolish interaction with TIP (Ren et al., 2000). These results suggest that TIP is a significant component in the resistance response. It has been suggested that TIP has a function similar to that of the tomato *Pto* R gene. Like *Pto*, TIP may function in a concerted manner with HRT to confer resistance to TCV (Ren et al., 2000). To study the specificity of resistance responses to TCV, a TCV virus has been created with a mutation in the amino terminus of the coat protein. This mutation, in which an aspartic acid is replaced by an asparagine at the fourth position (D4N) by site-directed mutagenesis, results in a virus that no longer elicits the HR and active defense response from Di-17 plants (Zhao et al., 2000).

It is also necessary to study the differences between the *TIP* genes of the plants. There are nucleotide differences that could account for changes at the protein level. Perhaps the nucleotides code for amino acids that could cause structural changes or charge differences that alter interaction dynamics to account for the different HR responses demonstrated by *A. thaliana* plants.

Although wild-type TCV is capable of eliciting an HR from Di-17 plants, the plants inoculated with the mutant D4N-TCV become systemically infected (Fig. 8). Di-3 plants display neither an HR nor resistance to wild-type TCV, however suprisingly, inoculation with the D4N mutant confers an HR (Fig. 8). In this case, the HR appears later, well after systemic disease symptoms are visible. Thus, the Di-3 plants are not resistant to D4N, though they appear capable of recognizing the D4N virus and initiating a partial resistance response. The Col-0 genotype does not mount an HR nor is it capable of resistance to wild-type TCV. Unpublished data has shown that although Col-0 develop necrosis and crinkling of infected leaves upon challenge with D4N, younger leaves appear to be healthy, and the severity of the symptoms of the virus seem to wane over time. Additionally, no clear lesions appeared on the Col-0 plants (Fig. 8).

It was unexpected to see the development of an HR when Di-3 plants were inoculated

Fig. 8. **D4N inoculated *Arabidopsis thaliana.*** **A.** Inoculated Di-3 plant, showing systemic infection with apparent necrosis and leaf crinkling. **B.** Inoculated Di-17 plant, with less severe necrosis. **C.** Inoculated Col-0 plant demonstrating necrotic and crinkled old leaves, with apparently healthy new leaves. **D.** Mock inoculated Col-0 plant, demonstrating a healthy phenotype.

with D4N. Di-3 plants are known to lack the *R* gene, *HRT* (Klessig, personal communication). Thus, the formation of an HR must be due to the action of another HR-inducing protein/protein interaction. In order to identify a genetic locus responsible for conferring the HR on Di-3 plants, we propose to cross them with Di-17 plants that do not respond to HR. Additionally, since the *rrt* status of Di-3 plants is unknown, we will be able to determine its status in the F_2 of the Di-3 X Di-17 cross by analyzing responses to wild-type TCV. We also propose to cross Col-0 X Di-3 plants. This cross will be important to determine and compare the importance for

the dominant and recessive alleles of the *rrt* gene. It will also help to determine the status of the *rrt* gene in Di-3 plants. Potentially, *rrt* will be required to get full resistance to D4N. In these crosses we can determine phenotype upon inoculation with D4N and the genotype at the *HRT* and *TIP* loci. Since *rrt* has not been mapped, we will not be able to follow the genotype at this locus.

Analysis of the F_1 and F_2 Generations

Given the information that we have about the parent plants, we are able to make some predictions as to their phenotype upon inoculation with D4N. For Di-17 and Col-0, we can determine that 75% of the F_1 generation and 37.5% of the F_2 generation will form lesions.

Key: For the following Punnet squares: H = *HRT*; T = *TIP*; and * = *HRT* gene for D4N.

F1	H*T	h*T
H*T	H*H*TT	H*h*TT
h*T	H*h*TT	h*h*TT

F2	H*T	H*T	h*T	h*T
H*T	H*H*TT	H*H*TT	H*h*TT	H*h*TT
H*T	H*H*TT	H*H*TT	H*h*TT	H*h*TT
h*T	H*h*TT	H*h*TT	h*h*TT	h*h*TT
h*T	H*h*TT	H*h*TT	h*h*TT	h*h*TT

For crosses of Di-17 and Di-3, then the following can be predicted.

1) IF either version of *TIP* is effective AND Di-3 is homozygous for the D4N resistance gene, then:
    * The F_1 generation will be genetically uniform.
    * 75% of the F_2 population will form lesions.

F1	H*	H*
h*	H*h*	H*h*
h*	H*h*	H*h*

F2	H*	h*
H*	H*H*	H*h*
h*	H*h*	h*h*

2) IF either version of *TIP* is effective AND Di-3 is heterozygous for its resistance

gene, then:

* The F_2 population will be mixed.
* 1/2 of the F_1s will give no lesion formers in F_2.
* 1/2 of the F_1s will give 75% lesion formers in F_2.

F1	H*	h*
h*	H*h*	h*h*
h*	H*h*	h*h*

F2	H*	h*
H*	H*H*	H*h*
h*	H*h*	h*h*

3) IF the Di-3 version of *TIP* is required AND Di-3 is homozygous for the D4N resistance gene, then:

* 56% of the F_2 population will form lesions.

F1	H*T	h*T
H*t	H*H*Tt	H*h*Tt
h*t	H*h*Tt	h*h*Tt

F2	H*	h*
H*	H*H*	H*h*
h*	H*h*	h*h*

4) IF the Di-3 version of *TIP* is required AND Di-3 is heterozygous for the D4N resistance gene, then:

* 50% of the F_1 generation will form lesions.
* Of the F_1 plants that form lesions, 56% of the F_2 generation will also form lesions.

F1	H*T	h*T
h*t	H*h*Tt	h*h*Tt
h*t	H*h*Tt	h*h*Tt

F2	H*T	h*T	H*t	h*t
H*T	H*H*TT	H*h*TT	H*H*Tt	H*h*Tt
h*T	H*h*TT	h*h*TT	H*h*Tt	h*h*Tt
H*t	H*H*Tt	H*h*Tt	H*H*tt	H*h*tt
h*t	H*h*Tt	h*h*Tt	H*h*tt	h*h*tt

Since inoculation of the F_1 generation would prevent the plants ability to reproduce, it is necessary to use our genotypic and phenotypic knowledge of the F_2 plants as the basis for our

understanding of *Arabidopsis*-TCV interactions. However, it is important to genetically screen F_1 plants to identify successful crosses in order to accurately determine the genetic-phenotypic relationships in the F_2 plants.

Results and Discussion:

CAPS Analysis: An Important Genetic Tool

Previously determined genetic markers can be used to identify the parental and successive generation genotypes. The two markers used in this experiment are specific for the *HRT* and *TIP* loci. Cleaved amplified polymorphic sequence (CAPS) analysis involves amplifying a region in or near a gene of interest. To view the genotype variations, restriction enzymes have been selected which differentially digest the amplified DNA fragment, providing a mechanism to compare the genotypes.

To determine the genotype of *A. thaliana* plants in relation to the *HRT* gene in CAPS analysis, we implemented the use of primers for another gene, *DFR*, which is very closely linked to *HRT* (Dempsey et al., 1997, Cooley et al., 2000). Because these genes are so closely linked, there is little chance that following *DFR* will result in miss-assigning the genotype at *HRT* due to the crossing over of genes on Chromosome 5. It was important to determine the genotypes of the wild-type parents in order to establish a standard with which to compare the crossed progeny. Genomic DNA was extracted from Di-17, Di-3, and Col-0 plant tissue, and CAPS analysis was performed.

Using the *DFR* primers, genomic DNA from all three ecotypes was amplified. The amplified band for each parental genotype is 1,200 bp in size. Two different enzymes must be used for genotype analysis in this particular CAPS analysis. A *Bsa*A I restriction digest is capable of providing a comparison between Di-3 and Col-0 or Di-3 and Di-17. As in Fig. 9, the 1,200 bp Col-0 and Di-17 bands in Lane 1 and Lane 5, respectively, are cut by the enzyme *Bsa*A I into two smaller bands of sizes 550 bp and 650 bp (Lanes 2 and 6, respectively). However, the 1,200 bp

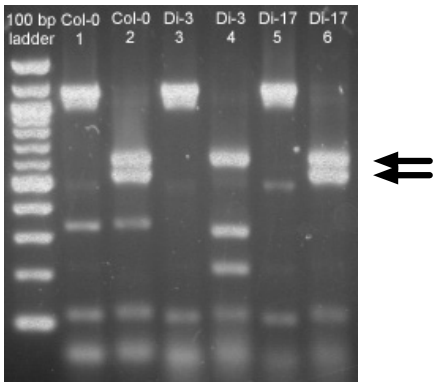


Fig. 9. **CAPS analysis of WT plants using *Bsa*A I.** Odd lanes are uncut, even lanes are digested with *Bsa*A I.

Di-3 band in Lane 3 is cut into three smaller bands by *Bsa*A I in Lane 4. These bands correspond to ~250 bp, ~300 bp, and 650 bp segments.

Given that these results make it impossible to distinguish between the Col-0 and Di-17 plants, it was necessary to use another enzyme. The second enzyme, *Dde* I, provides this important genotypic distinction (Fig. 10). Lanes 1, 3 and 5 show the uncut 1,200 bp Col-0, Di-3 and Di-17 bands, respectively. In Lanes 2 and 4, it is possible to see that the large DNA segment was cut into a smaller, 800 bp band by *Dde* I. However, Di-17 is not digested by *Dde* I, as shown in Lane 6. These results demonstrate that the three genotypes can be distinguished using two enzymes.

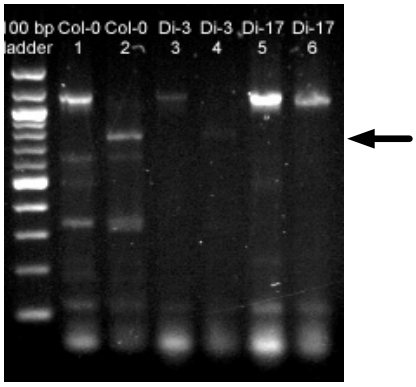


Fig. 10. **CAPS analysis of WT plants using *Dde* I.** Odd lanes are uncut, even lanes are digested with *Dde* I.

It is important to perform a similar CAPS analysis for the *TIP* gene. While the *TIP* genes for each line of *A. thaliana* had previously been cloned, it was still necessary to develop CAPS markers for the *TIP* gene. Specific primers were selected to amplify a segment of the *TIP* gene that would allow for the determination of each genotype by CAPS analysis. A segment of the gene which contained a nucleotide change was selected, as it was possible to identify an enzyme that would selectively restrict only one genomic sequence at that point. The sequence analysis revealed an *Xmn* I restriction site in the Di-17 and Col-0 *TIP* genes that was absent in the Di-3 *TIP* gene (unpublished data).

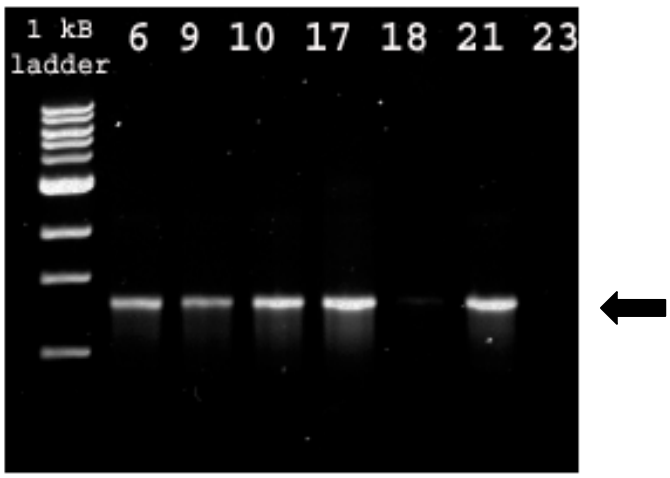


Fig. 11. **Amplification of *TIP*** using the CAP3 PCR program. The band is ~1.3 kB in size.

As the use of TIP sequences for CAPS analysis had not previously been done, it was necessary to optimize conditions for successful amplification using these primers. Initially, conditions similar to those used for *HRT* amplification were used, but these showed little, if any, amplification. Adjustments were made to several

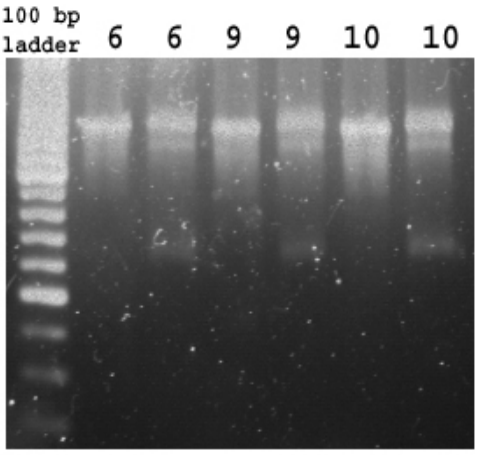


Fig. 12. **Restriction digest of *TIP*.** First lane for each number is uncut, the second lane for each number is digested with *Xmn* I. After restriction digest, the product bands remain unresolved. In an attempt to clarify these results, various parameters were modified. However, changes in the concentrations of the digestion reaction, time for the reaction and levels of sample loaded on the gel did not significantly alter these results.

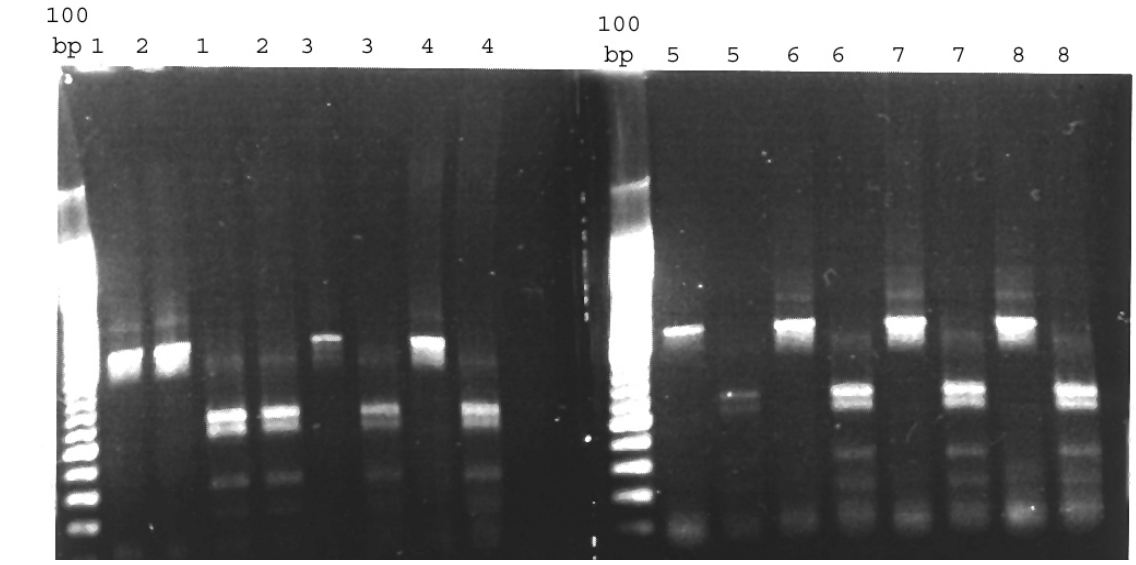
parameters, including the relative concentration of the genomic DNA and the number of elongation cycles for the PCR program. These parameter changes let to significant improvements in the amplification process (Fig. 11). However, there was still some difficulty in resolving the product bands of 500 bp and 600 bp after cleavage with the restriction enzyme *Xmn* I (Fig. 12).

Crossing Experiments

Crosses of both Di-3 and Di-17 plants, and Col-0 and Di-3 were performed, and seeds were collected and dried. The F_1 generations of both crosses were grown up, and CAPS analysis for *HRT* was performed on tissue samples taken from both types of plants (Fig. 13). This demonstrated that all crosses were successful since the F_1 contained one *HRT* locus from each parent. The plants that were identified as crosses at the *HRT* locus were grown up for seed. At this point, F_2 seeds from fifteen successful crosses of Di-3 and Di-17 and seven successful crosses of Col-0 and Di-3 have been harvested. These can be used to analyze the genotype and phenotype to D4N inoculation.

Initial Plant Experiments: The Soil Factor

During the course of analyzing F_2 progeny, the plants developed an unexpected range of symptoms. Plants that should have been susceptible to the viral species were not becoming necrotic and crinkled: they demonstrated purpling and severe growth variation (Fig. 14). Purpling usually occurs in a nutrient-deficient environment, and is typically identifiable by a



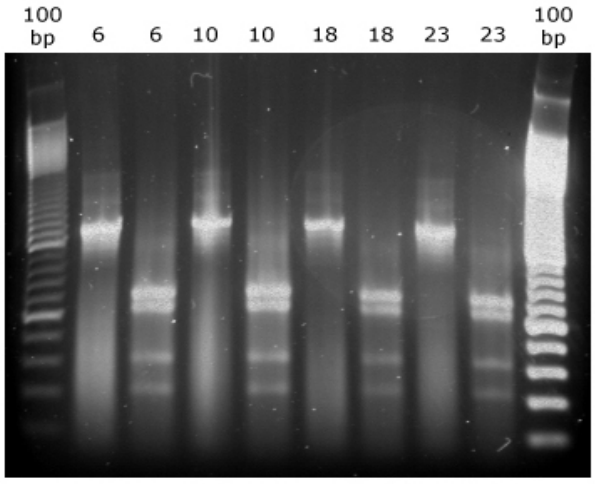


Fig. 13. **CAPS analysis of F_1 generations at the *HRT* locus.** Numbers at the top of the figures correspond to different plant samples. Samples were digested with *Bsa*A I. The second lane for each sample is the digested lane. Note the four bands in the digested lanes, corresponding to each of the bands seen in the parental digestion. **Top:** Di-17 X Di-3. For Di-17, digested bands are 550 and 650 bp. For Di-3, digested bands are 250, 300 and 650 bp. The 650 bp bands co-migrate, which is why only four total bands are observed. **Left:** Col-0 X Di-3. For Col-0, digested bands are 550 and 650 bp. For Di-3, digested bands are 250, 300 and 650 bp (as before).

dark purple color that appears to travel up the stem and bolt towards the tip of the leaf. Eventually, the entire leaf becomes a dark purple color, and the growth process of the plant slows down. This purpling was unexpected since the plants had been fertilized since the first watering, and the soil was not autoclaved. Autoclaving soil is often used to prevent bacterial and fungal growth, although some nutrients are destroyed in the process. Furthermore, some plants appeared to have stunted growth, whereas other plants in the same flat appeared to be abnormally large, with thick stems. The soil had significant nutrient heterogeneity with most

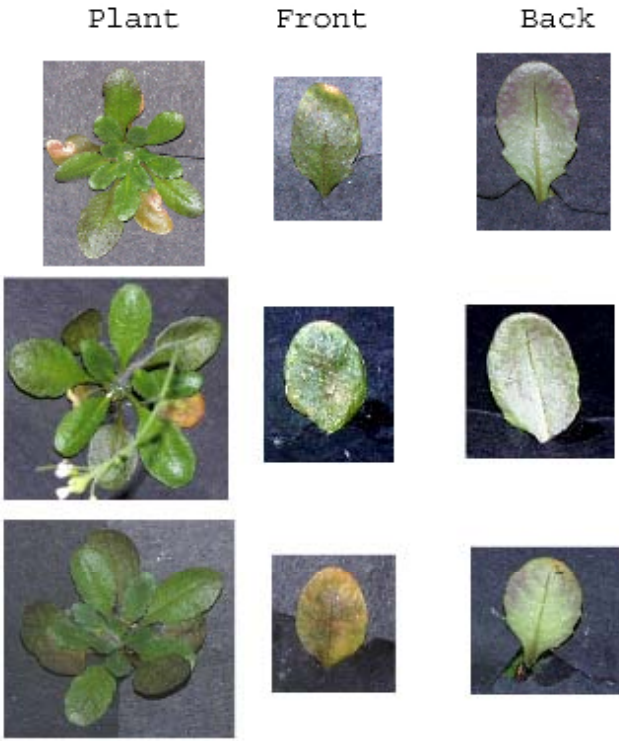


Fig. 14. **Plants demonstrating purpling and chlorosis.**

regions being deficient. Furthermore, previous growth experiments, performed in the same manner, but with a different lot of soil, had not proven to contain plants with these symptoms.

Eventually, it was determined that the only factor in the experiment that had been changed, and remained consistent amongst the plants, was the lot of soil that was being used. Plants rely on nutrients in soil for their health and well-being. Although nutrients are often supplemented later, as with fertilizers, it is still important for the soil to contain these essential nutrients for the initial stages of plant growth. Soil manufacturers mix soil and nutrients in order to optimize plant growth and health. This process is imperfect, and the nutrients can vary between lots.

Although the brand and mixture of the soil that was used had been implemented in growth experiments before, the lot was different. Soil (used and unused), water, and tissue samples were sent out to an analytical laboratory, Micro-Macro International, for nutrient identification. It was determined that the levels of some of the nutrients in the soil and plant tissue were well outside of the normal range (Table 4 and Table 5).

Elements such as magnesium, calcium, chloride, and potassium are important for maintaining ion balance in plants. Copper, iron, zinc, and manganese are essential for electron transport and serve as catalysts for enzymes. Consistent levels of elements promote a constructive growth environment, and plants are moderately capable of adjusting to excess metal-element levels. Plants can use selective ion uptake and decreased permeability of membranes to control uptake of elements contained in high levels in the soil. However, plants are unable to adapt well to decreased levels of nutrients in the soil. Nutrient deficiency can result in necrosis or death of a portion of a leaf, in addition to the purpling displayed on older

Media Analysis	Unused Media	Used Media	Normal Ranges (ppm)
	Sample (ppm)	Sample (ppm)	
Nitrate ($N0_3^-$)	5.84	0.00	40-200
Ammonium (NH_4^+)	5.88	0.77	0-30
Phosphorus (P)	47.43	11.12	5-30
Potassium (K)	92.27	32.81	40-200
Calcium (Ca)	76.16	23.47	40-200
Magnesium (Mg)	25.26	6.19	28-80
Iron (Fe)	1.03	0.06	0.30-3.00
Maganese (Mn)	0.23	0.00	0.10-3.00
Boron (B)	0.14	0.05	0.05-0.5
Copper (Cu)	0.02	0.01	0.01-0.30
Zinc (Zn)	2.01	0.41	0.10-0.30
Molybdenum (Mo)	0.01	0.00	0.01-0.10
Sodium (Na)	36.60	34.90	No data
Aluminum (Al)	0.40	0.06	No data

Table 4. **Media analysis.** Results of the media (soil) analysis performed by Micro-Macro International. Values highlighted in red indicate results lying outside of the normal range.

leaves due to anthocyanin accumulation (Jones, 1998). These symptoms are severe, and are capable of preventing accurate identification of viral infection responses in *A. thaliana*.

Nitrogen is an important nutrient for plants. It can be found in both organic and inorganic compounds in plants, including amino acids, nucleic acids, chlorophyll, purine bases and alkaloids. In soil, nitrogen is usually found in the nitrate ($N0_3^-$) anion, or in the ammonium (NH_4^+) cation. Although NO_2^- may also be present in the soil, elevated levels (<5 ppm) are toxic to plants (Jones, 1998). Soil analysis revealed that the nitrate levels in the soil were severely deficient in unused media; the level was nearly $1/8^{th}$ of the lower end of the normal range for soil. As would be expected, used media contained undetectable levels of nitrate.

Nitrogen plays a key role in plant metabolism. Plants lacking sufficient nitrogen tend to mature slower and are weaker than their healthy counterparts. In addition, the leaves of these plants tend to be lighter in color (Jones, 1998). These problems are significant from a virology standpoint in that it is likely that the resistance system of the plant would not be matured, and would be difficult to identify. A main determinant of infection of *A. thaliana* plants by TCV is

Tissue Analysis	%	Low	Medium	High	Sufficiency Ranges
Nitrogen (N)	Not available		Not available		1.50-4.60
Phosphorus (P)	0.56		X		0.11-0.67
Potassium (K)	2.12		X		0.11-0.69
Calcium (Ca)	2.09		X		0.30-2.6
Magnesium (Mg)	0.27		X		0.11-1.90
Sulfur (S)	Not available		Not available		Not available
Iron (Fe)	45.44		X		30-250
Manganese (Mn)	77.08		X		30-300
Boron (B)	48.42		X		14-175
Copper (Cu)	2.98	X			5-28
Zinc (Zn)	140.32			X	25-100
Molybdenum	0.46		X		0.2-5.0

Table 5. **Tissue analysis.** Nutrient analysis of plant tissue performed by Micro-Macro International. Values highlighted in red indicate nutrient levels that lie outside of the sufficiency range for plant tissue.

necrosis. However, if the plant leaves are not healthy to begin with, it is difficult to distinguish symptoms of viral infection from the background symptoms of plant malnutrition.

Levels of magnesium were also deficient in both the unused and used soil samples. These samples were found to contain only 90% and 22% of the lower ranges of normal values, respectively. However, magnesium levels were normal in the plant tissue. This further illustrates the remarkable ability of organisms to adapt to less than ideal circumstances. Magnesium is contained in chlorophyll, and serves as an enzyme cofactor in many phosphorylation reactions. Plants deficient in magnesium display chlorosis and yellowed leaves (Jones, 1998), similar to the effects of TCV infection. Since tissue levels of magnesium were normal, this is not likely to have been a problem for our plants.

Phosphorus can be found in plants in many forms: adenosine triphosphate (ATP), RNA, DNA, and phytin, to name a few. In the soil, phosphorus is available to plants in both inorganic and organic form. Dihydrogen phosphate ($H_2PO_4^-$) and monohydrogen phosphate (HPO_4^-) are two common anion forms in the soil. Additionally, aluminum, iron, and calcium phosphate are the major inorganic sources of phosphorus (Jones, 1998).

Soil sampling revealed that unused soil contained greater than 150% of the upper limit for the suggested amount of phosphorus. Phosphorus levels in used soil were within the

normal range. Tissue analysis results corresponded to the used soil data, providing information that the phosphorus concentration in the plant itself remained within the normal range for tissue. However, the levels of phosphorus in the tissue were close to the upper limit of normal ranges. Given that plant tolerances differ amongst species, it is possible that the high levels of phosphorus in the soil could affect the *A. thaliana* plants. Increased levels of phosphorus prove to be detrimental to plants. Excess phosphorus can affect the levels of nutrients such as iron and zinc. It has also been suggested that high levels of phosphorus could interfere with plant metabolism (Jones, 1998). It is possible that "pockets" of concentrated phosphorus were contained in the soil, causing the abnormal overgrowth of plants, appearing sporadically in various pots in the flats.

Copper plays many roles in plants: it is integral to plastocyanin, it functions as a link in the photosystem I/photosystem II electron transport system, it plays a role in nitrogen fixation, etc. Copper is usually found in the soil in low molecular weight proteins, although the cupric ion (Cu^{2+}) is present in small quantities (Jones, 1998). Plants lacking sufficient copper demonstrate symptoms such as stunted growth and necrosis (Jones, 1998), again making the presence of copper essential for the plant-virus experiments. Although the results showed that copper levels in both the used and unused media fell within the typical range, the levels were at the low end. Tissue analysis further revealed that the level of copper in the plant was lower than normal, which would further account for the necrosis seen on the plants.

Zinc was also found to be in excess in the soil samples: with a six fold higher level in ususued soil, and almost 1.5 times excess in used soil. As would be predicted from the large decrease in zinc levels in soil, sampled plant tissue contained 40% more zinc than in normal tissue. *In planta*, zinc is involved with carbonic anhydrase activity. In soil, zinc is found in cationic form (Zn^{2+}) and in organic complexes. An excess of zinc can affect plants that are sensitive to iron. When high levels of zinc are present, chlorosis is observable (Jones, 1998).

These results correspond to the symptoms of purpling and necrosis demonstrated in plants grown in this soil. As previously mentioned, some symptoms of TCV are similar, if not identical, to those demonstrated by nutrient excess and deficiency. Therefore, it was impossible to accurately determine the effects of TCV inoculation on *A. thaliana* plants. However, the soil problem was eventually rectified by supplying the plants with an appropriate fertilizer that provided the plants with the correct levels of nitrogen, until a new lot of soil was obtained.

TIP: Differences in the Alleles

As previously discussed, *TIP* is of interest since it interacts directly with the coat protein of TCV. It is possible that the differences between *TIP* and *TIP** are in part responsible for the differing responses to D4N-TCV. As a preliminary analysis, the two *TIP* sequences were compared *in silica.* The eight nucleotide changes between the two TIP sequences result in only five amino acid changes in the protein (Appendix A) (Fig. 15). These amino acid differences could potentially create a significant change or changes in the protein, including protein structure and overall protein charge. There are many on-line programs available to test these possibilities. These programs are capable of converting nucleic acid sequences (DNA) to amino acid sequences (protein), calculating the predicted secondary structure of proteins from the amino acid sequences, calculating protein hydrophobicity, etc.

Amino Acid	Di-17	Di-3
237	an uncharged, polar threonine → a nonpolar isoleucine	
290	an uncharged, polar, aromatic tyrosine → a positively charged, polar histidine	
339	a basic charged lysine → a basic charged arginine	
340	a positively charged, polar arginine → an uncharged, polar threonine	
448	an uncharged, polar serine → a nonpolar glycine	

Fig. 15. **Amino acid changes in TIP.**

The amino acid changes are all found between the middle and carboxy terminal end of the protein. This is interesting since the N-terminal region of the protein has been shown to have significant homology to members of the NAC protein family. Some of these NAC proteins (i.e. *NAM, CUC2, NAP)* have been implicated in plant growth processes, including

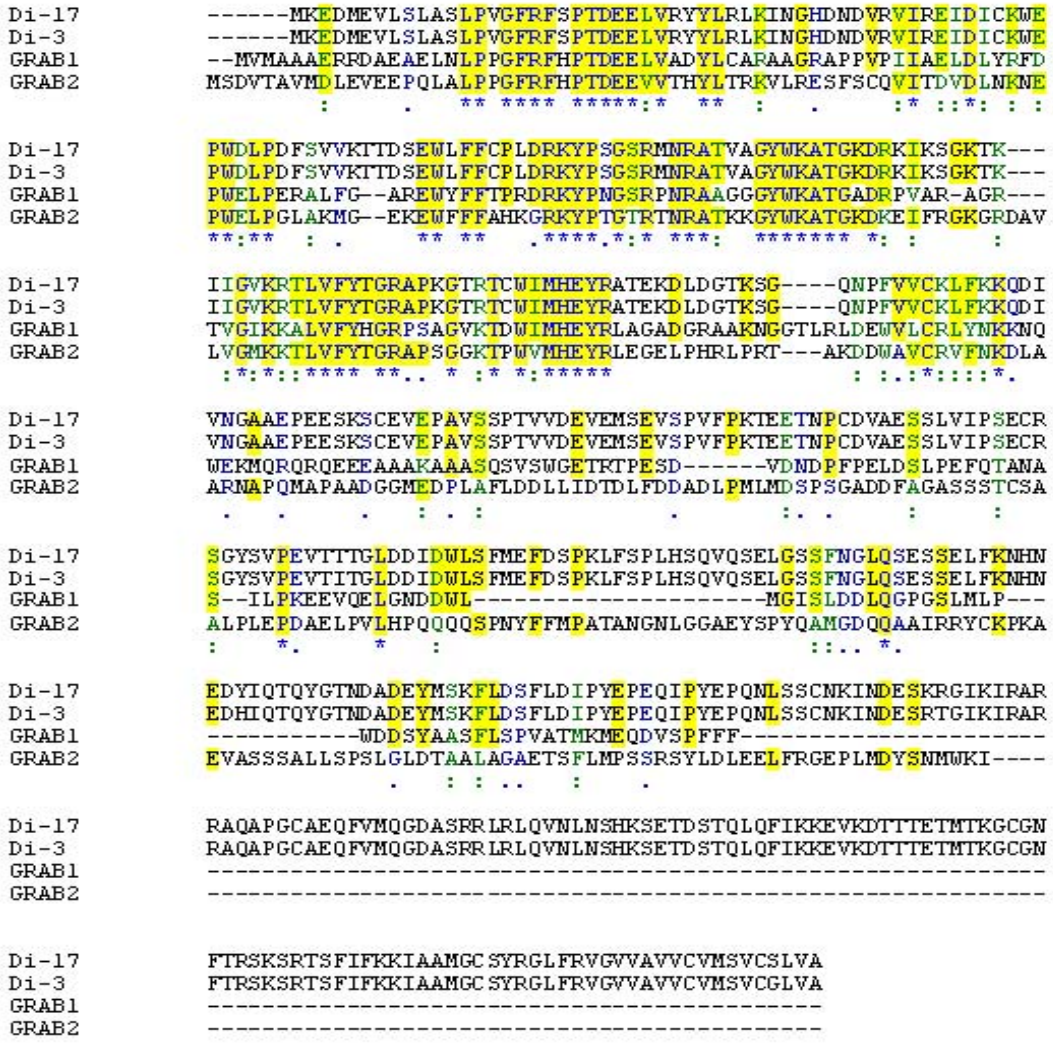
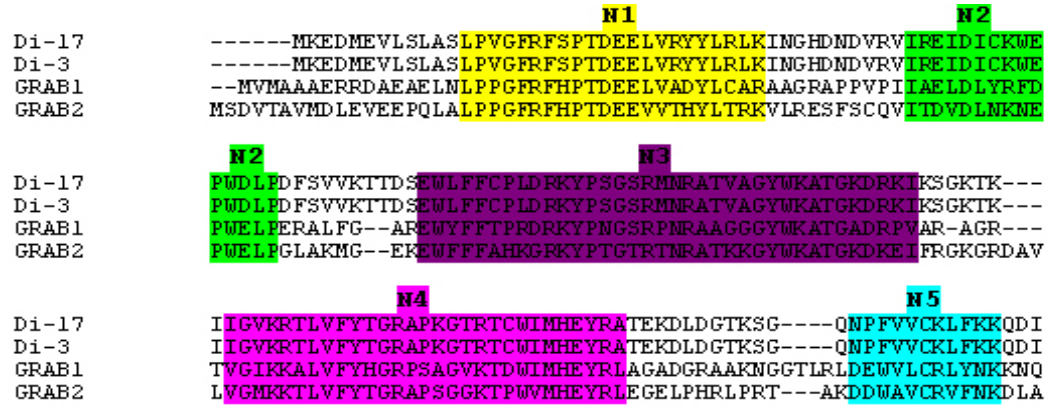


Fig. 16. **GRAB and TIP amino acid sequence alignment.** Highlighted residues (yellow) denote a consensus in the protein sequences. ***** denotes a single, fully conserved residue. **:** denotes conservation of a strong group. **.** denotes conservation of weak groups. Note the high incidence of consensus residues in the N-terminal regions of the proteins. (from CLUSTALW alignment program, Biology WorkBench v. 3.2, 2003)

separating cotyledons, generating floral organs, and cell expansion and division (Ren et al., 2000). However, it is obvious that *TIP* plays a role in plant resistance. Other NAC proteins, *GRAB1* and *GRAB2,* also demonstrate direct interaction with a plant virus.

GRAB proteins were originally isolated by their interaction with the *Wheat dwarf*

geminivirus (WDV) protein RepA. Geminiviridae are a family of plant viruses whose genome is composed of one or two 2.6-3.0 kb circular single-stranded DNA molecules. WDV has one of the smallest genomes of the geminiviridae (2750 nucleotides), and is found to replicate its DNA in the nucleus of infected plants. These Geminivirus Rep A-binding proteins, or GRAB proteins, were found to contain a NAC domain. This domain, located in the N-terminal region of GRAB is both necessary and sufficient to interact with the RepA protein of WDV (Xie et al., 1999). Fig. 16 shows a sequence alignment between GRAB1, GRAB2, and the TIP proteins. From the sequence alignments, it is possible to see that there is a great deal of amino acid homology at the N-terminal regions of the four sequences. Further alignments, provided in Appendix B, also show the amino terminal similarities between TIP and other NAC domain



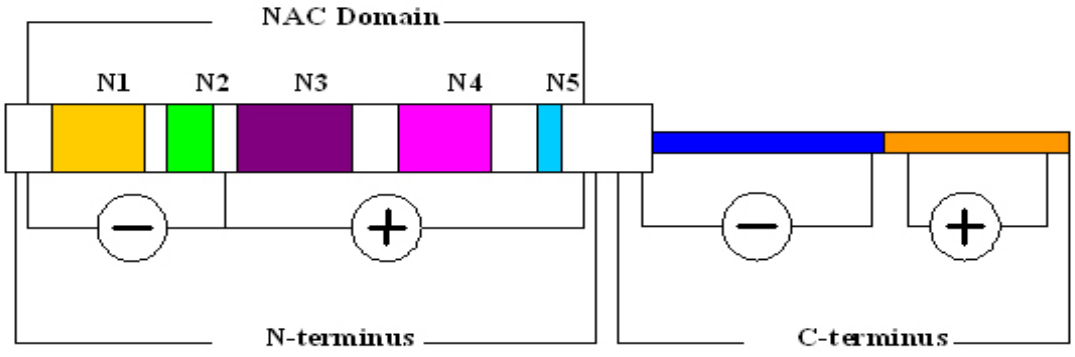


Fig. 17. **NAC consensus domain. A.** Partial GRAB and TIP sequence alignments with the five conserved blocks (N1-N5) highlighted. These motifs are conserved amongst the NAC family proteins (Xie et al., 1999). **B.** Organization of the consensus domains in NAC family proteins. The five conserved motifs and distribution of residue charge are illustrated for the TIP protein (adapted from Xie et al., 1999)

family proteins. However, each NAC protein contains a unique C-terminal domain, and the overall net charge in this region is highly negative in most NAC proteins. This is due to the fact that 15-20% of the C-terminal residues are negatively charged aspartic or glutamic acids (Fig. 17) (Xie et al., 1999). However, TIP has an extended C-terminal domain which is largely positive, although the middle section of TIP, which corresponds to the C-terminal domain of the GRAB proteins, is negatively charged. Since both the *GRAB* and *TIP* genes are implicated in plant resistance to viruses, it is highly probably that the NAC domain is significant for the protein-protein interactions. Furthermore, it is also likely that the N-terminal NAC domain is involved in signaling activity given the roles in plant development that many of the NAC family proteins play.

TIP has been implicated in having a role in transcription activation. A yeast two-hybrid

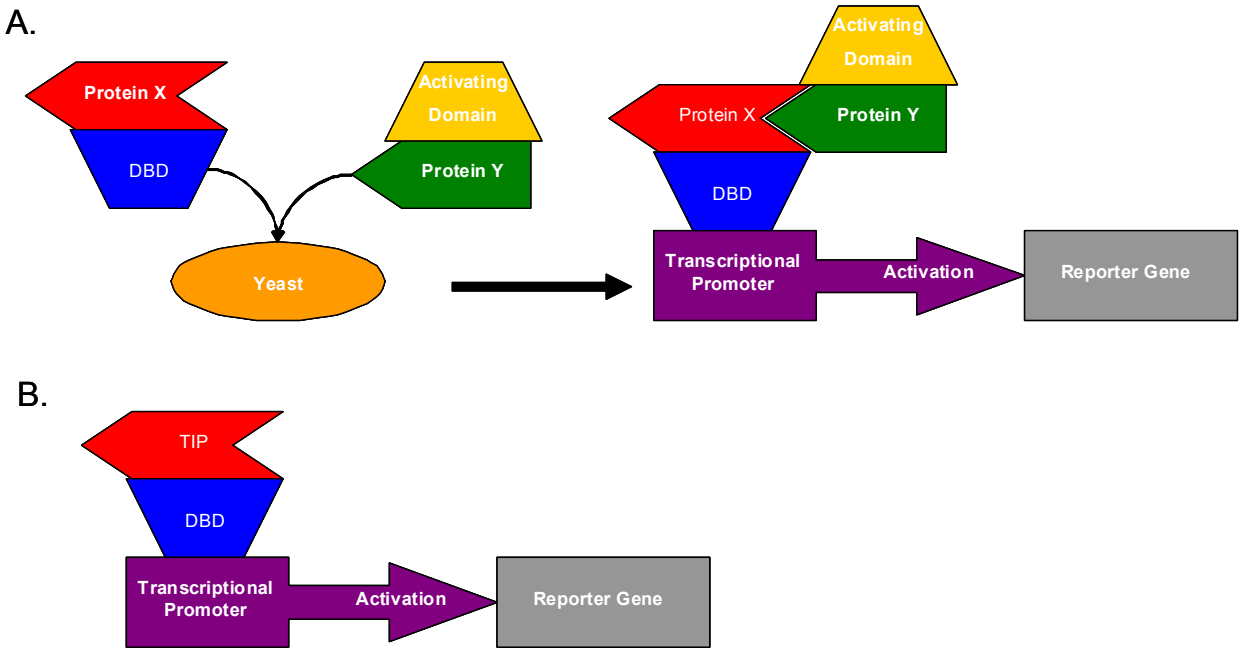


Fig. 18. **Yeast two-hybrid assay. A.** Diagram of a typical yeast two-hybrid assay. Two vectors are created, each with a different protein (X or Y) fused to either a DNA binding domain (DBD) or an activating domain. If the two proteins interact, a transcriptional promoter is activated and the reporter gene is expressed. **B.** TIP-hybrid assay. TIP fused to a DBD was capable of activating reporter gene expression without the addition of an activating domain vector, demonstrating that TIP contains a domain capable of transcriptional activation (Ren et al., 2000).

assay (Fig. 18) was performed using a DNA binding domain-TIP (DBD-TIP) fusion protein vector, in the absence of a complementing activation domain vector. The DBD-TIP protein was capable of activating the expression of the reporter gene in yeast, providing evidence that TIP

not only contains an activation domain, but that TIP is also a transcriptional activator (Ren et al., 2000). GRAB proteins have also been identified as putative transcriptional activators. Both GRAB1 and GRAB2 contain glutamine-rich C-terminal domains, implying that the proteins have a role in transcriptional regulation (Xie et al., 1999).

It is interesting to discover that the amino acid changes in the *TIP* genes appear in the C-terminal region, since the N-terminal region bears the most sequence homology to NAC proteins. Given that these proteins are likely to be involved in signaling cascades due to their roles in plant development, it would be natural to predict that the N-terminal region is involved in this signaling activity. Since the N-terminal regions of the *TIP* genes are the same, only the amino acid changes in the C-terminal area could be involved in altering resistance responses.

It is possible to predict the secondary structure of a protein using its amino acid

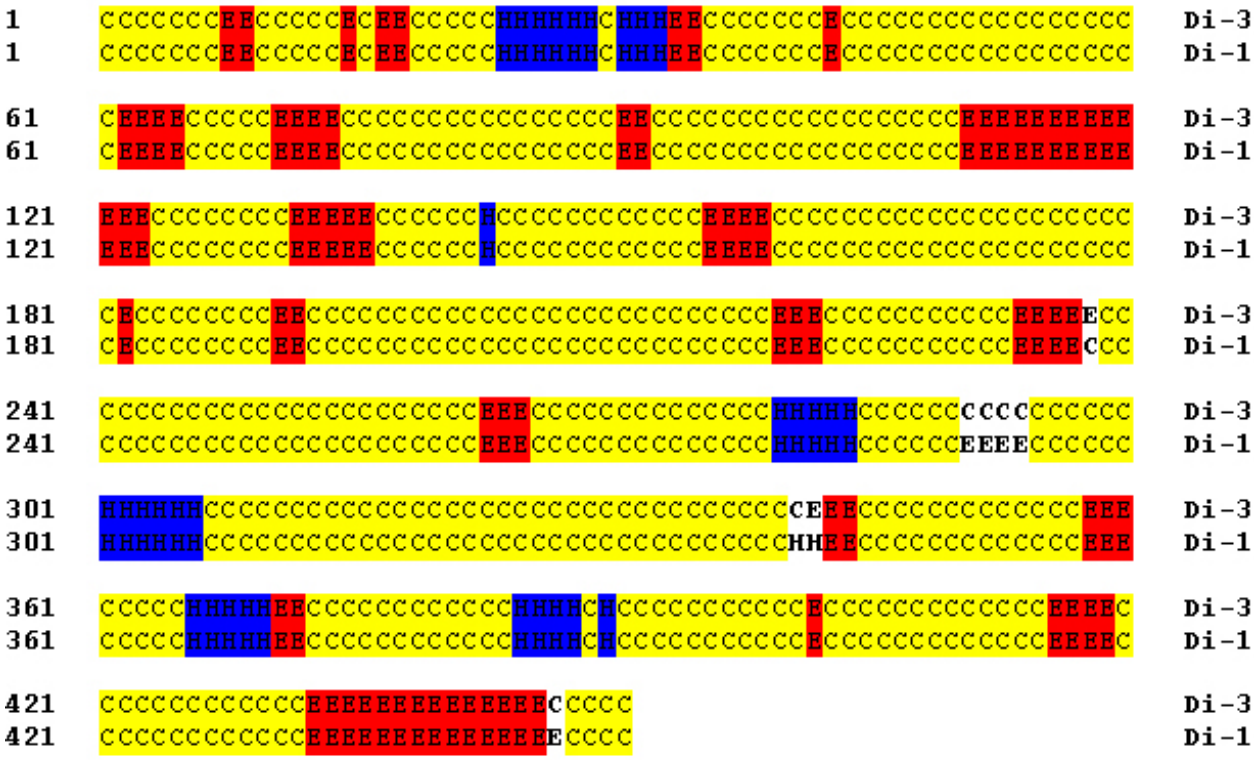


Fig. 19. **Joint prediction alignment.** The predicted secondary structure of the TIP proteins is depicted above. **C** corresponds to coils, **E** corresponds to beta-strands, and **H** corresponds to alpha-helices. Areas in **bold** (white, un-shaded areas) denote a lack of consensus between the two sequences. Four such areas were predicted by the JOI prediction program (from PELE-protein structure prediction, Biology WorkBench v. 3.2, 2003)

sequence. Both TIP sequences were subjected to various secondary structure predictions (Appendix C), one of these being a joint prediction, or a best fit prediction using a combination of seven separate predictions (Fig. 19).

There are four areas of interest in the secondary structure joint prediction depicted in Fig. 19. Three of the areas of difference in the secondary structure involve putative changes from beta-strands to coils, whereas the remaining area postulates a putative alpha-helix, a coil, and a beta-strand. Although it is possible that the areas of single amino acid changes are capable of altering the protein enough to cause differences between the resistance functions of the TIP proteins, it is more likely that the larger amino acid difference has a greater role.

It is significant that the differences between the Di-3 and Di-17 TIP proteins all fall within the last half of the protein. Ren et al. (2000) presented preliminary evidence that the 100 C-terminal amino acids of TIP are responsible for the interactions with TCV CP. Though only one amino acid is different in this region, it is distinctly possible that the more N-terminal changes result in secondary structure changes that significantly impact the structure and environment of the very C-terminal end. Thus, the differences in response to TCV-D4N could in part be explained by differences in TIP protein structures.

In order to minimize the interactions between hydrophobic amino acids and water, proteins will often fold so as to place the hydropobhic residues in the interior of the protein, and the hydrophilic residues on the surface of the protein. Therefore, it is possible to predict the amino acids that lie on the surface of a protein. Many hydrophobicity scales have been constructed by experimental techniques that involve the observation of peptides in polar and apolar solvents (Molecular Toolkit, 2003).

Hydrophobicity plots are generated by analyzing a protein using a moving "window" that scans through the amino acid sequence. The window is set to a specific number of amino acids, and proceeds to "move" down the protein sequence, calculating a value at each position that it goes through. These values are calculated using the mean value of the amino acids within the window from a hydrophobicity scale. This value is then plotted, and a graphical representation of the protein is generated. Areas of the graph that lie above the midpoint score are predicted to be hydrophobic, whereas the areas that lie below the midpoint score are predicted to be hydrophilic (Molecular Toolkit, 2003). Since there are many different hydrophobicity scales, there are many different plots that can be generated. A variety of these

plots and a corresponding table of hydrophobic values for each scale can be found in Appendix D.

Fig. 20 shows a hydrophobicity plot constructed using the Kyte-Doolittle scale. The two TIP proteins are represented by different colored lines: Di-3 in red, and Di-17 in blue. Where the proteins are identical, the line is purple. There are three areas of interest (the encircled regions). Each of these regions corresponds to amino acid variances in the protein sequence. There are distinguishable value changes between positions 200 and 250, suggesting that the hydrophobicity changes could alter protein folding or interaction. The Di-17 TIP residues in this area are found to be slightly more hydrophilic, whereas the Di-3 TIP residues are predicted to be more hydrophobic. The hydrophobicity changes, while clear, are hardly dramatic. However, given that TIP has been identified as having a role in protein-protein

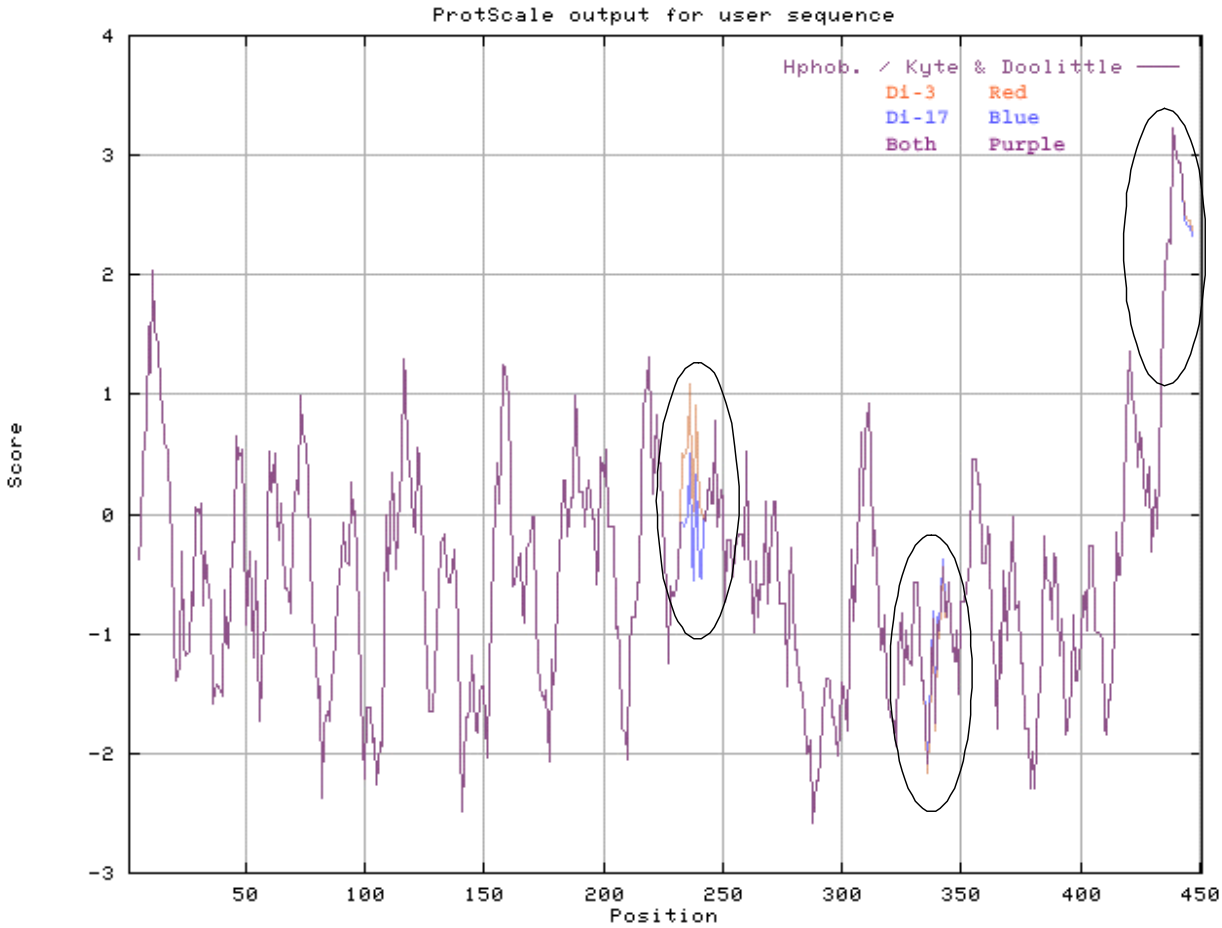


Fig. 20. **Kyte-Doolittle hydrophobicity plot.** A hydrophobicity plot created using the Kyte-Doolittle hydrophobicity scale. The three encircled regions contain areas that differ between the two TIP proteins. Values above zero indicate hydrophobic regions, whereas values below zero indicate hydrophilic regions (from ProtScale Tool, 2003)

interaction, it is possible that this change in hydrophobicity could be a factor in resistance ablility. A much smaller and less apparent difference between the plots is present between positions 300 and 350. Since there seems to be little change in this region, it is less likely that this region is involved in any phenotypic differences of the proteins. Finally, the third region of dissimilarity lies at the very end of the plot, just before position 450. Again, there is little change in the plot, suggesting a small hydrophobicity difference between the proteins. The other hydrophobicity plots, generated with different hydrophobicity programs, also show this general trend (Appendix D).

Conclusions:

The complexity of plant-pathogen interactions is both intriguing and bewildering. Although it is a small piece of the larger plant resistance pathway, determining the genetics of *A. thaliana* resistance to TCV challenge is important for understanding how many plant-virus interactions occur. The key to *A. thaliana* resistance may lie in the recognition of TCV CP by TIP.

A correlation between genotype and phenotype of *A. thaliana* plants will be essential to understanding the role of TIP and resistance genes *HRT* and *rrt*. We have laid the groundwork for future experiments by producing an optimized CAPS analysis for the *HRT* locus, by creating and analyzing an F_1 generation of Di-3 X Di-17 and Col-0 X Di-3 plants, and by predicting the outcome caused by nucleotide changes in the Di-3 *TIP* gene at the molecular level of the protein.

Future research should be focused on determining correlations between phenotype and genotype in the F_2 generations. To do this, large quantities of Di-3 X Di-17 and Col-0 X Di-3 plants should be inoculated with D4N or WT TCV, and their resistance response should be compared to their genotype. Genotypic analysis can be performed using CAPS analysis at both the *HRT* and *TIP* loci. Furthermore, it would be advantageous to identify the affect of the nucleotide differences between the two *TIP* genes, and perhaps correlate the differences to the resistance response of the plants.

Methods and Materials:

Plant Growth Conditions:

Arabidopsis plants were grown in Pro-Mix 'BX' soil in a growth chamber kept at 23 degrees Celcius, with 14 hours of daylight. The plants were fertilized at each watering with Miracle-Gro Professional Excel water-soluble fertilizer (13-2-13 +6 Ca +3 Mg Plug Special), as directed.

Crossing Experiments:

Plants were crossed by exposing the pistil in one immature bud by removing the sepals, petals, and stamens. A stamen from a flowering plant was used to pollinate the exposed pistil. After pollination, the pistil was wrapped in saran wrap until the seed pods had formed. Pods were collected as they dried. The pods were placed in Eppendorf tubes, and allowed to dry until they opened naturally. These seeds were then sealed in Eppendorf tubes.

Plant DNA Maxiprep:

Between 0.5 g and 0.75 g of leaves were ground in a mortar and pestle with liquid nitrogen until a fine, grey-green powder was obtained. Next, 15 mL EB (100 mM Tris pH 8, 50 mM EDTA pH 8, 500 mM NaCl, and 10 mM b- mercaptoethanol) was added and the mortar was placed in a 65° C water bath until the mixture had melted. The solution was then transferred to a test tube and 1 mL of 20% SDS was added. The test tube was vortexed for 1 minute, and then placed in a 65° C water bath for 10 minutes. Then, 5 mL of 5M potassium acetate was added to the solution. The test tube was vortexed and placed on ice for 20 minutes. The tubes were centrifuged at 17000 rpm for 20 minutes. The supernatent was poured through a microcloth or cheesecloth filter into a fresh tube containing 10 mL isopropanol. The solution was placed at -20° C for approximately 2 hours. The test tube was then centrifuged at 15000 rpm for 15 minutes. The supernatent was removed from the centrifuge tube, and the pellet was allowed to dry for 10 minutes. Next, 0.7 mL of 50 mM tris and 10mM EDTA were added to the centrifuge tube. The solution was transferred to an Eppendorf tube and microfuged for 1 minute at full speed. Then, 50 µL 3M NaOAc and 100 µL 1% CTAB was added. The tube was

then centrifuged for 1 minute at full speed. The supernatent was removed and the remaining pellet was washed with 70% EtOH. The pellet was redissolved in 400 µL TE. Next, the DNA was precipitated with 50 µL 3M NaOAc and 1 µL EtOH. The Eppendorf tube was microfuged for 1 minute at full speed. The supernatent was removed and 400 µL TE buffer was added to the Eppendorf tube. The pellet was resuspended and 50 µL 3M NaOAc was added and the solution was vortexed. Next, 1 µL EtOH was added and the Eppendorf tube was vortexed. The sample was placed at –20° C overnight. The next day, the sample was microfuged for 5 minutes full speed. The supernatent was removed and the Eppendorf tube was placed upside-down to dry. The pellet was resuspended in 300 µL TE and placed on ice. The sample was stored at -20° C.

Modified Baumbusch DNA Miniprep:

Two small leaf samples were ground with a pestle attached to a power drill. Then, 700 µL pre-warmed EB (200 mM Tris-HCl pH 7.5-8, 250 mM NaCl, 25 mM EDTA, and 0.5% SDS) was added. Samples were incubated at 65° C for 15 minutes, mixing at 5 minutes. Next, 220 µL 3M KoAc was added, the sample was vortexed, and placed on ice for 20 minutes. The sample was centrifuged on full for 10 minutes, and the supernatent was removed to a new Eppendorf tube. Then, 550 µL isopropanol was added. The sample was placed at room temperature for 1 hour, mixing occasionally. Next, the sample was centrifuged on full for 5 minutes. The supernatent was removed and the pellet was dried for approximately 10 minutes. The pellet was then dissolved in 100 µL TE buffer, and 2 µL RNase A was added. The sample was incubated at 37° C for 15 minutes. Then, 10 µL 3M NaOAc was added, and the sample was vortexed. Next, 200 µL 100% EtOH was added, and the sample was vortexed. The sample was placed at –20° C for 10 minutes. The samples were centrifuged on full for 10 minutes, and then washed with 80% EtOH. The samples were again centrifuged, and rinsed. The supernatent was removed, and the pellet was resuspended in 20 µL dH_20. The samples were stored at -20° C.

RNA Maxiprep from Plants:

The leaf sample (~1 g of tissue) was ground into a fine powder in a mortar and pestle in

the presence of liquid nitrogen. Tissue was resuspended in 5 mL GTC buffer (4 M guanidinium thiocyanate, 25 mM sodium citrate pH 7, 0.5% sarkosyl, and, just before use, 0.1 M b-mercaptoethanol), and vortexed. Then, 0.5 mL 2M NaOAc pH 4 was added, the sample was vortexed, and phenol/chloroform extracted (5 mL). The sample was then centrifuged at 7000 rpm for 10 minutes. The aqueous phase was removed to a new tube, and the phenol/chloroform extraction was repeated. Again, the aqueous phase was removed to a new tube, and 5 mL isopropanol was added. The sample was iced for 10 minutes to overnight. Next, the sample was centrifuged at 7000 rpm for 10 minutes. The pellet was retained, and 1.5 mL LiCl was used to resuspend the pellet. The solution was transferred to an Eppendorf tube, and centrifuged for 5 minutes at maximum speed. The supernatent was removed, and the pellet was resuspended in 750 µL resuspension buffer (0.5% SDS, 10 mM Tris pH 7.5, and 1 mM EDTA). The sample was phenol/chloroform extracted (500 µL) twice. The aqueous phase was saved, and 100 µL 1M NaOAc pH 5 and 600 µL isopropanol were added to the solution. The sample was placed at -20° C for at least 30 minutes. The pellet was then resuspended in 100-300 µL depc treated water.

Primers:

DFR Primers:

For analysis of the genotype at the *HRT* locus, DFR primers were used. The primers designed for this are as follows: AGATCCTGAGGTGAGTTTTTC and TGTTACATGGCTTCATACCA. For analysis of the genotype at the *TIP* locus, the following primers were designed: AGACCGTAAGATCAAATCAGG and TTATGCGACTAGAGTGCAGAC

PCR Programs:

CAP2:

1)	94° C	4 minutes
2)	94° C	30 seconds
3)	50° C	1 minute
4)	72° C	45 seconds
5)	Go to step 2	40 times.
6)	50° C	1 minute
7)	72° C	10 minutes
8)	4° C	24 hours

CAP3:

1)	95° C	3 minutes
2)	50° C	1 minute
3)	75° C	2 minutes
4)	94° C	1 minute
5)	Go to step 2	35 times.
6)	50° C	1 minute
7)	75° C	10 minutes
8)	4° C	24 hours

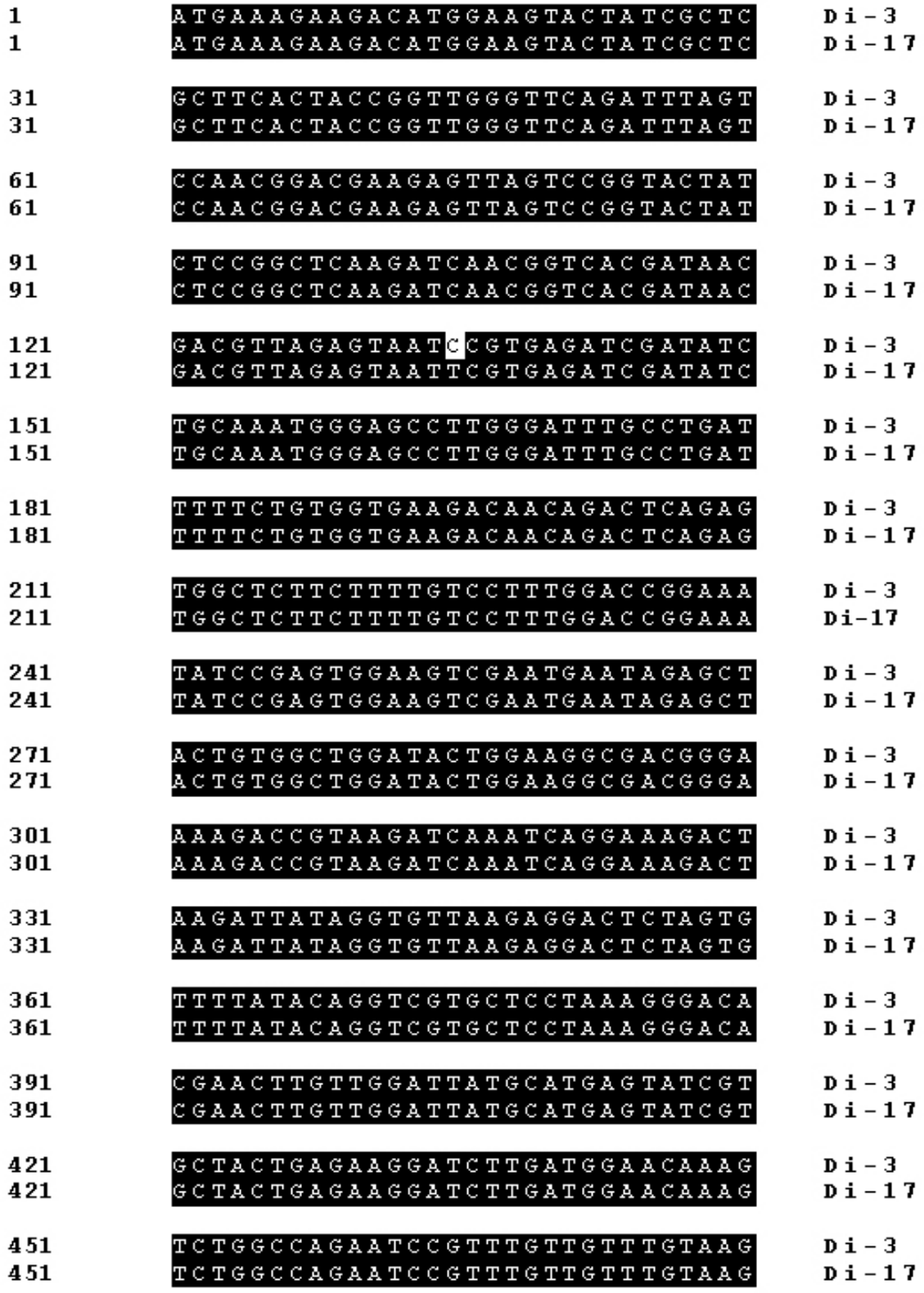
Restriction Digest:

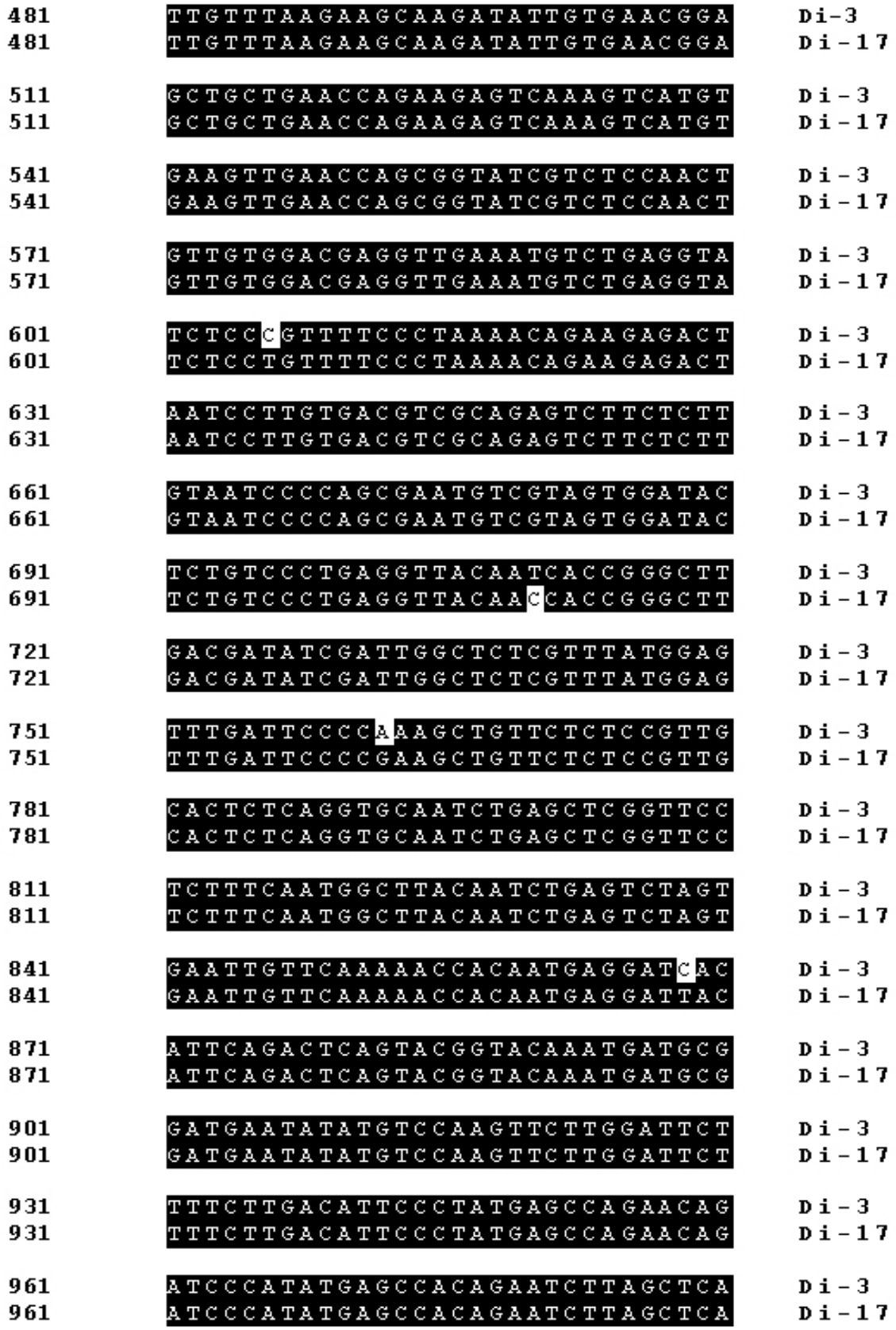
All digests were performed with appropriate DNA:enzyme concentrations at 37° C for at least 1.5 hours.

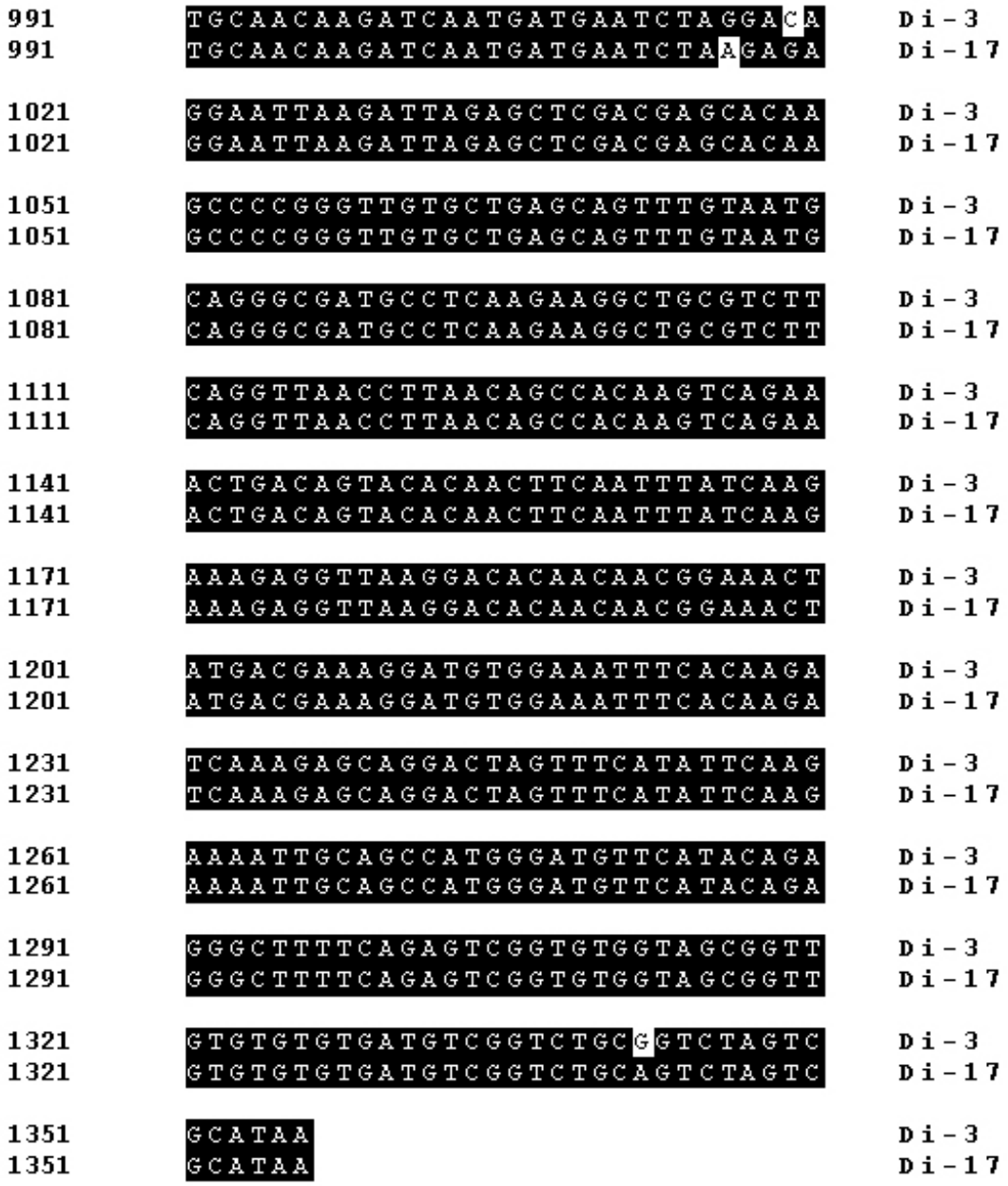
Appendix A:

A-1. Nucleotide sequence alignment
A-2. TIP protein alignment

A-1. **Nucleic acid sequence alignment**

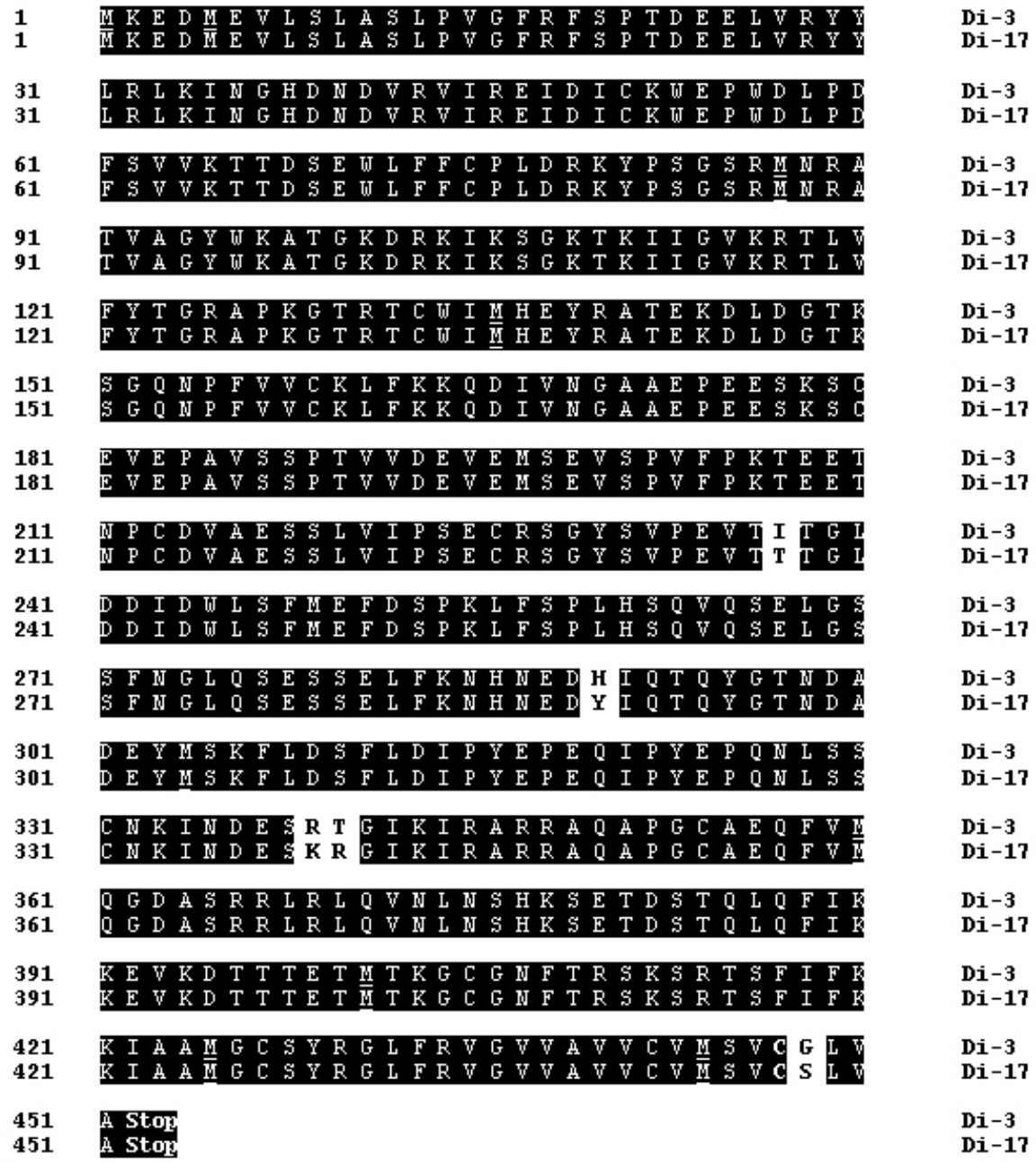



481 TTGTTTAAGAAGCAAGATATTGTGAACGGA Di-3
481 TTGTTTAAGAAGCAAGATATTGTGAACGGA Di-17
511 GCTGCTGAACCAGAAGAGTCAAAGTCATGT Di-3
511 GCTGCTGAACCAGAAGAGTCAAAGTCATGT Di-17
541 GAAGTTGAACCAGCGGTATCGTCTCCAACT Di-3
541 GAAGTTGAACCAGCGGTATCGTCTCCAACT Di-17
571 GTTGTGGACGAGGTTGAAATGTCTGAGGTA Di-3
571 GTTGTGGACGAGGTTGAAATGTCTGAGGTA Di-17
601 TCTCCCGTTTTCCCTAAAACAGAAGAGACT Di-3
601 TCTCCTGTTTTCCCTAAAACAGAAGAGACT Di-17
631 AATCCTTGTGACGTCGCAGAGTCTTCTCTT Di-3
631 AATCCTTGTGACGTCGCAGAGTCTTCTCTT Di-17
661 GTAATCCCCAGCGAATGTCGTAGTGGATAC Di-3
661 GTAATCCCCAGCGAATGTCGTAGTGGATAC Di-17
691 TCTGTCCCTGAGGTTACAATCACCGGGCTT Di-3
691 TCTGTCCCTGAGGTTACAACCACCGGGCTT Di-17
721 GACGATATCGATTGGCTCTCGTTTATGGAG Di-3
721 GACGATATCGATTGGCTCTCGTTTATGGAG Di-17
751 TTTGATTCCCCAAAGCTGTTCTCTCCGTTG Di-3
751 TTTGATTCCCCGAAGCTGTTCTCTCCGTTG Di-17
781 CACTCTCAGGTGCAATCTGAGCTCGGTTCC Di-3
781 CACTCTCAGGTGCAATCTGAGCTCGGTTCC Di-17
811 TCTTTCAATGGCTTACAATCTGAGTCTAGT Di-3
811 TCTTTCAATGGCTTACAATCTGAGTCTAGT Di-17
841 GAATTGTTCAAAAACCACAATGAGGATCAC Di-3
841 GAATTGTTCAAAAACCACAATGAGGATTAC Di-17
871 ATTCAGACTCAGTACGGTACAAATGATGCG Di-3
871 ATTCAGACTCAGTACGGTACAAATGATGCG Di-17
901 GATGAATATATGTCCAAGTTCTTGGATTCT Di-3
901 GATGAATATATGTCCAAGTTCTTGGATTCT Di-17
931 TTTCTTGACATTCCCTATGAGCCAGAACAG Di-3
931 TTTCTTGACATTCCCTATGAGCCAGAACAG Di-17
961 ATCCCATATGAGCCACAGAATCTTAGCTCA Di-3
961 ATCCCATATGAGCCACAGAATCTTAGCTCA Di-17



(Adapted from Hammond, 2001)

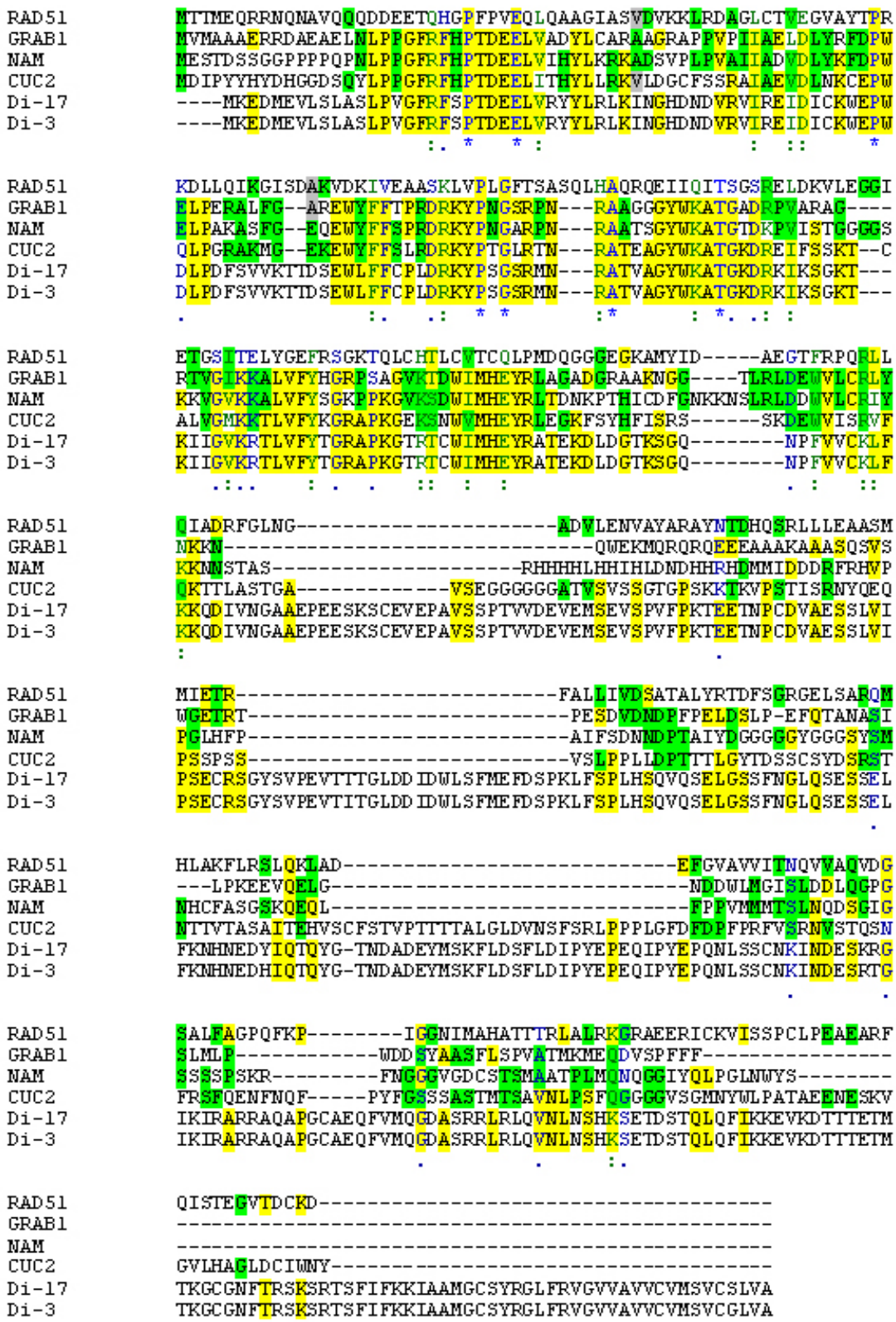
A-2. TIP Protein Alignment



Appendix B:

B-1. NAC family protein sequence alignments

B-1. NAC family protein sequence alignments



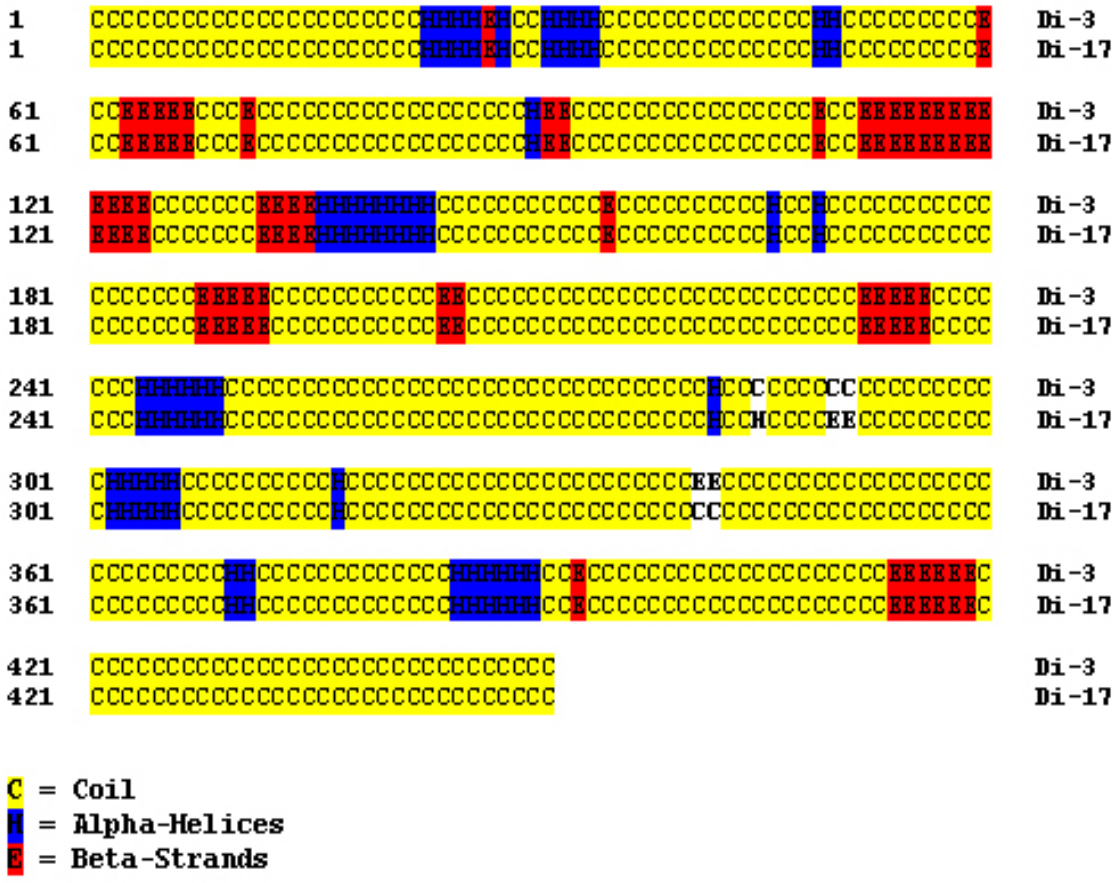
Green : NAC protein (-TIP) residue homology
Yellow : NAC protein (+TIP) residue homology
*** : Fully conserved residue**
: : Conservation of strong groups
. : Conservation of weak groups

CLUSTALW alignment program, Biology WorkBench v.3.2, accessed March 2003, http:// workbench.sdsc.edu.

Appendix C:

C-1. BPS prediction alignment
C-2. D_R prediction alignment
C-3. GGR prediction alignment
C-4. GOR prediction alignment
C-5. H_K prediction alignment
C-6. JOI prediction alignment
C-7. K_S prediction alignment

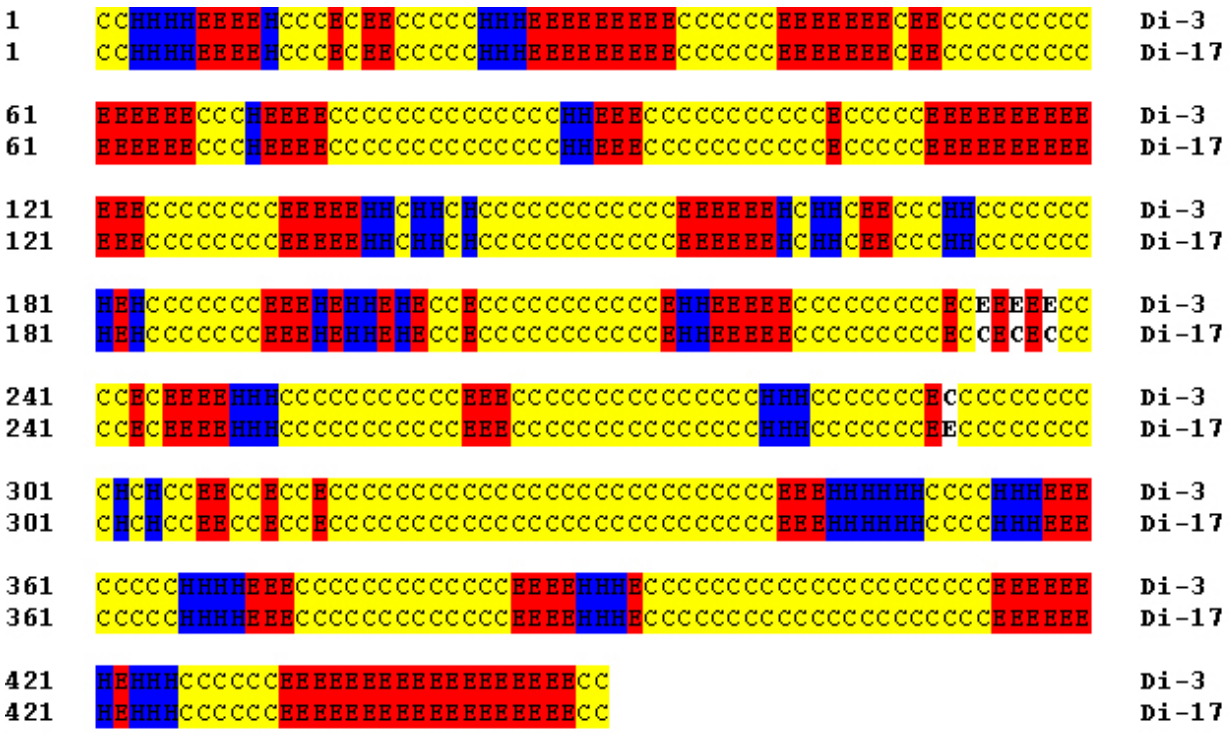
C-1. BPS prediction alignment



BPS : A. W. Burgess and P. K. Ponnuswamy and H. A. Sheraga, Analysis of conformations of amino acid residues and prediction of backbone topography in proteins, Israel J. Chem., p239-286, 1974, vol 12.

SDSC Biology Workbench v.3.2 (Accessed March 2003) PELE-protein prediction program. *http:// workbench.sdsc.edu.*

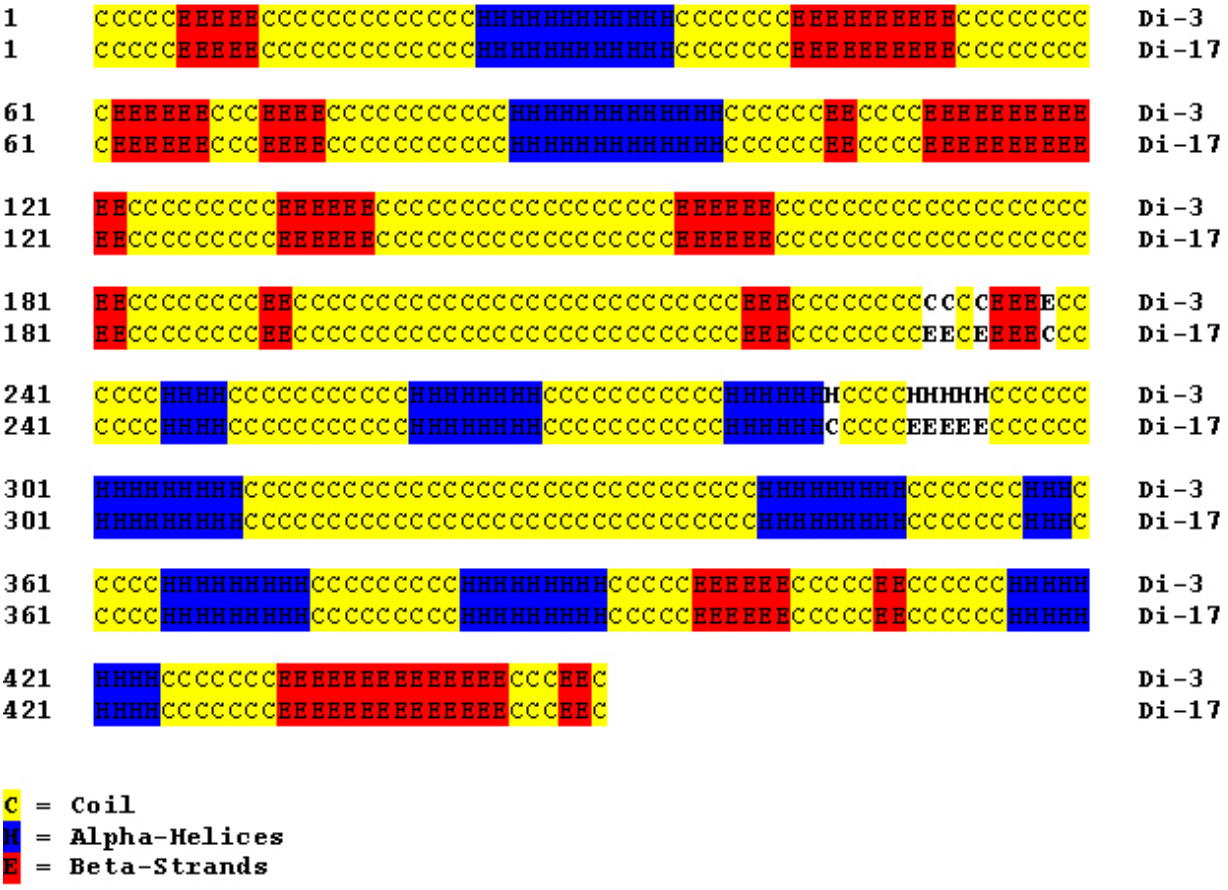
C-2. D_R prediction alignment



D_R : G. Dele`age and B. Roux, An algorithm for secondary structure prediction based on class prediction, Protein Engineering, p289-294, 1987, vol 1, num 4.

SDSC Biology Workbench v.3.2 (Accessed March 2003) PELE-protein prediction program. *http:// workbench.sdsc.edu.*

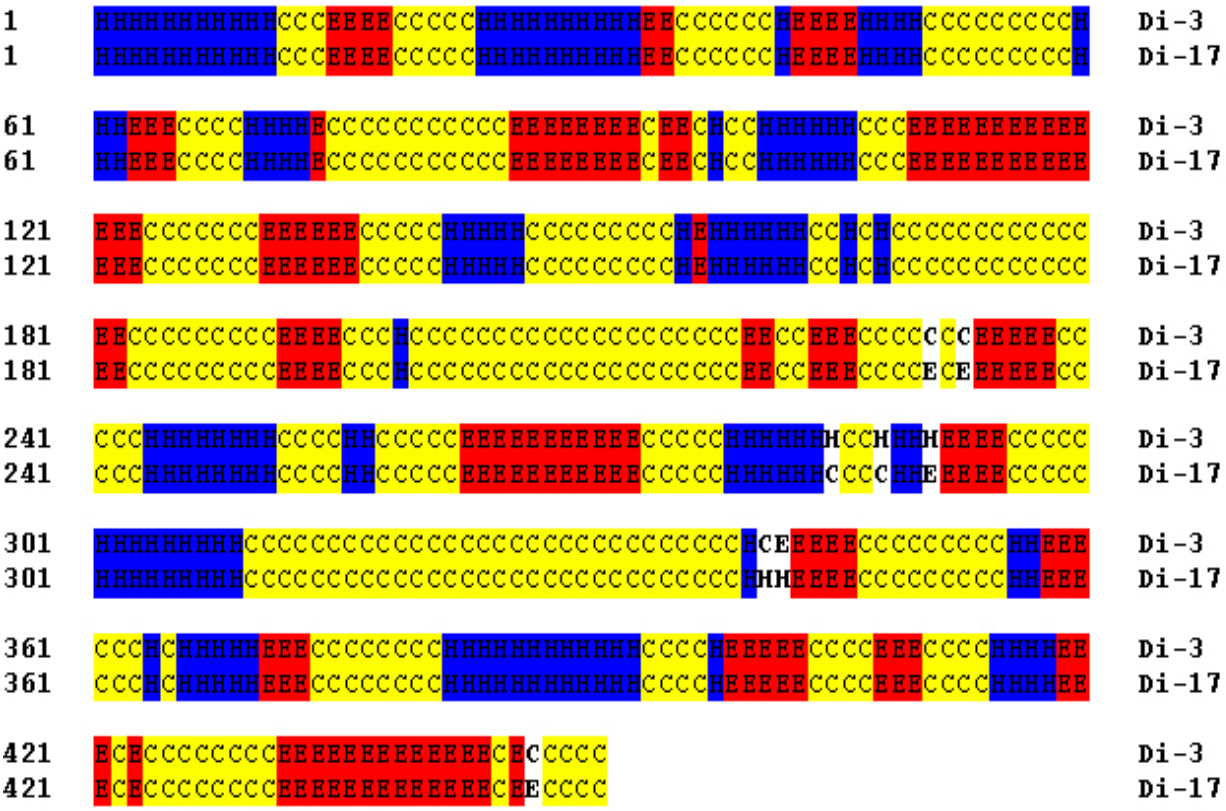
C-3. GGR prediction alignment



GGR : Garnier, Gibrat, and Robson, Meth. Enzym., R.F. Doolittle ed. 1996, 266: 97-120

SDSC Biology Workbench v.3.2 (Accessed March 2003) PELE-protein prediction program. *http:// workbench.sdsc.edu.*

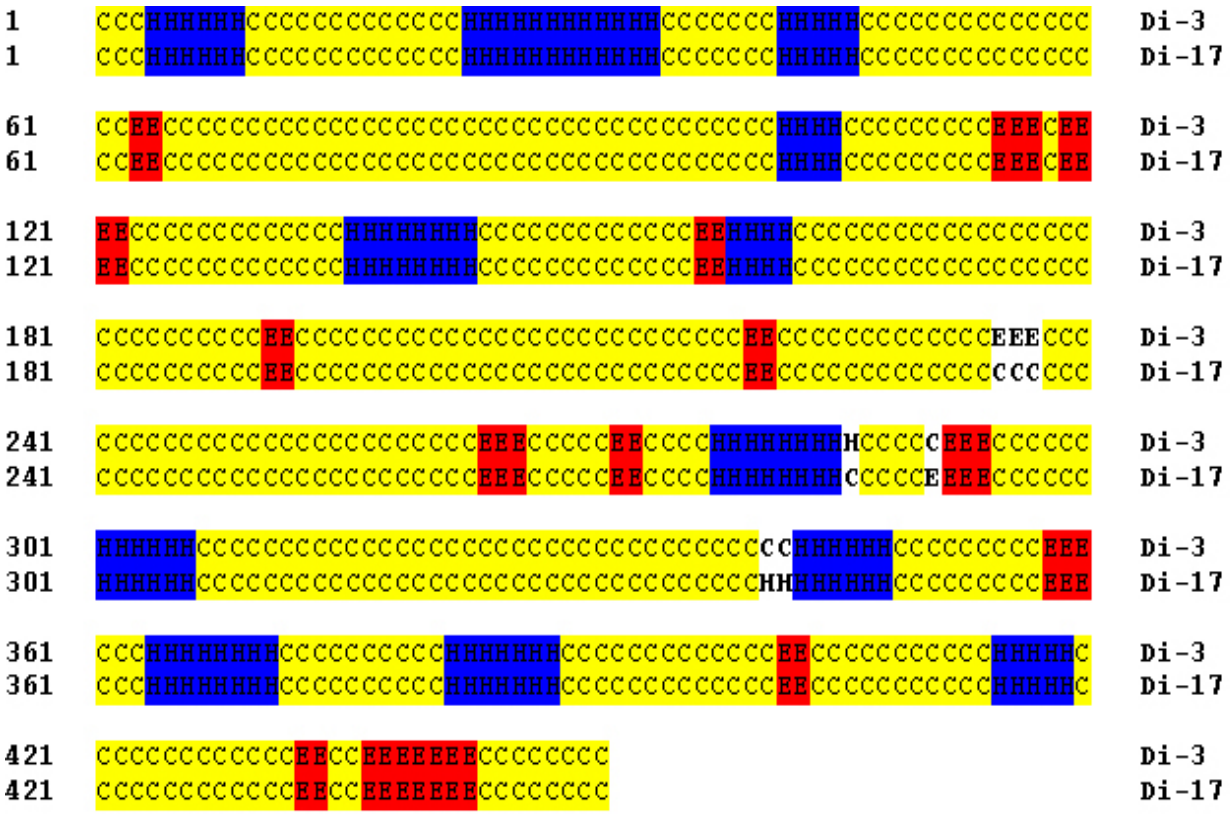
C-4. GOR prediction alignment



GOR : Jean Garnier and D. J. Osguthorpe and Barry Robson, *Analysis of the accuracy and implications of simple methods for predicting the secondary structure of proteins, J. Mol. Biol., p 97-120, 1978, vol 120.*

SDSC Biology Workbench v.3.2 (Accessed March 2003) PELE-protein prediction program. *http:// workbench.sdsc.edu.*

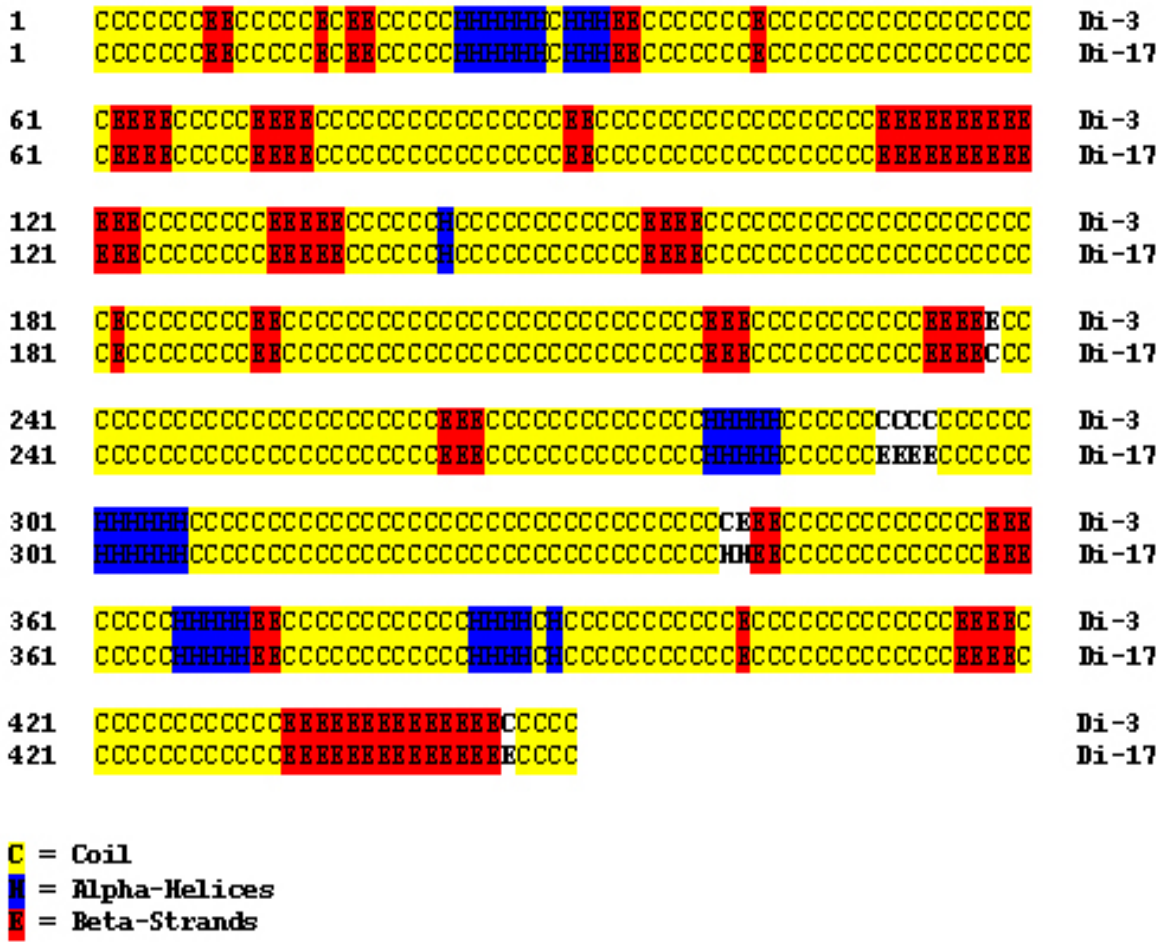
C-5. H_K prediction alignment



H_K : L. Howard Holley and Martin Karplus, Protein secondary structure prediction with a neural network, Proc. Natl. Acad. Sci. USA, p 152-156, Jan 1989, vol 86.

SDSC Biology Workbench v.3.2 (Accessed March 2003) PELE-protein prediction program. *http:// workbench.sdsc.edu.*

C-6. JOI prediction alignment



JOI Joint prediction - Prediction made by the program that assigns the structure using a "winner takes all" procedure for each amino acid prediction using the other methods.

SDSC Biology Workbench v.3.2 (Accessed March 2003) PELE-protein prediction program. *http://workbench.sdsc.edu.*

C-7. K_S prediction alignment

K_S : Ross D. King and Michael J. E. Sternberg, Machine learning approach for the prediction of protein secondary structure, J. Mol. Biol., p 441-457, 1990, vol 216.

SDSC Biology Workbench v.3.2 (Accessed March 2003) PELE-protein prediction program. *http:// workbench.sdsc.edu.*

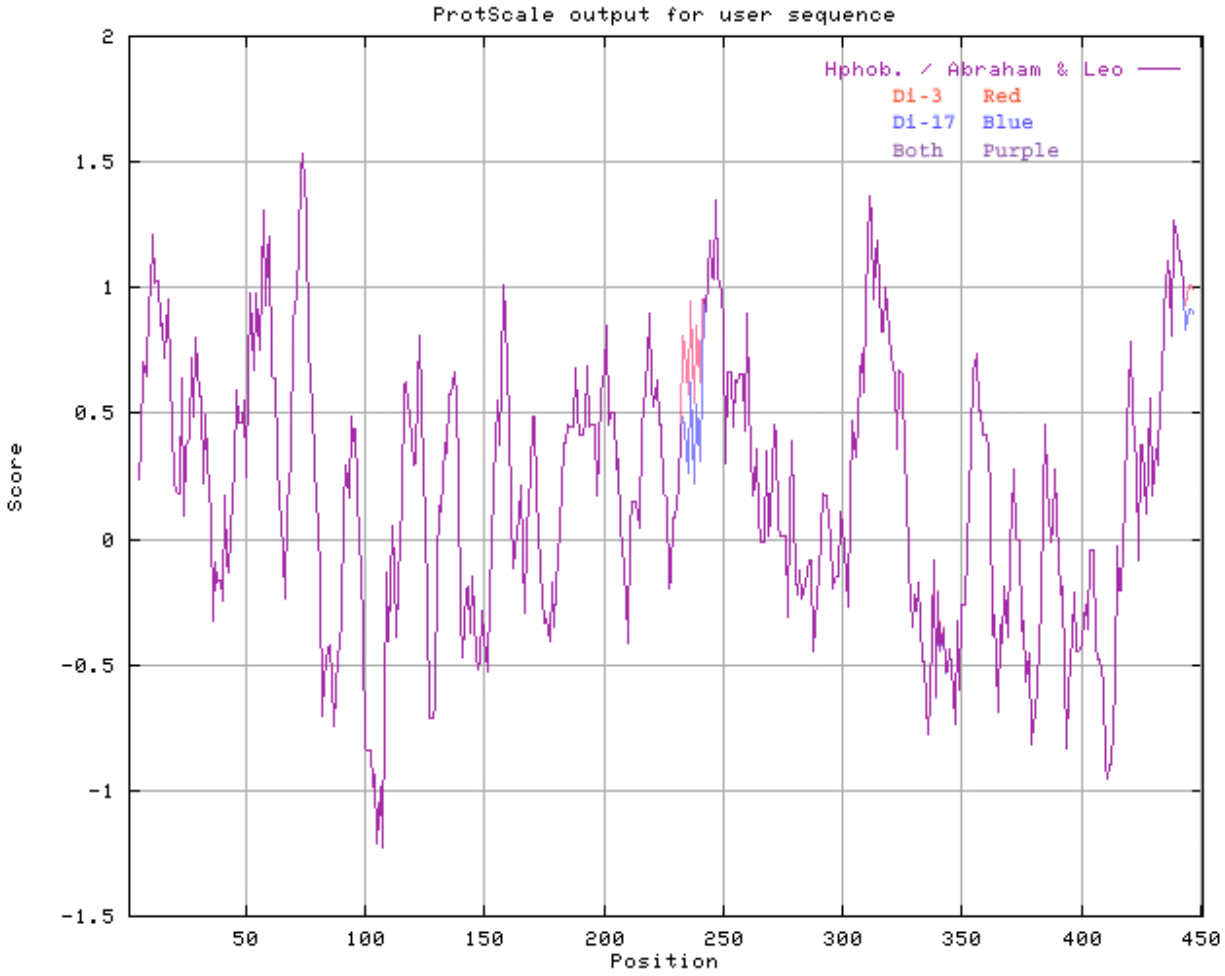
Appendix D:

D-1. Table of hydrophobicity scales
D-2. Abraham & Lee hydrophobicity overlay
D-3. Bull & Breese hydrophilicity overlay
D-4. Guy hydrophilicity overlay
D-5. Miyazawa et al. hydrophobicity overlay
D-6. Roseman hydrophobicity overlay
D-7. Sweet et al. hydrophobicity overlay
D-8. Welling et al. hydrophobicity overlay

D-1. Table of hydrophobicity scales

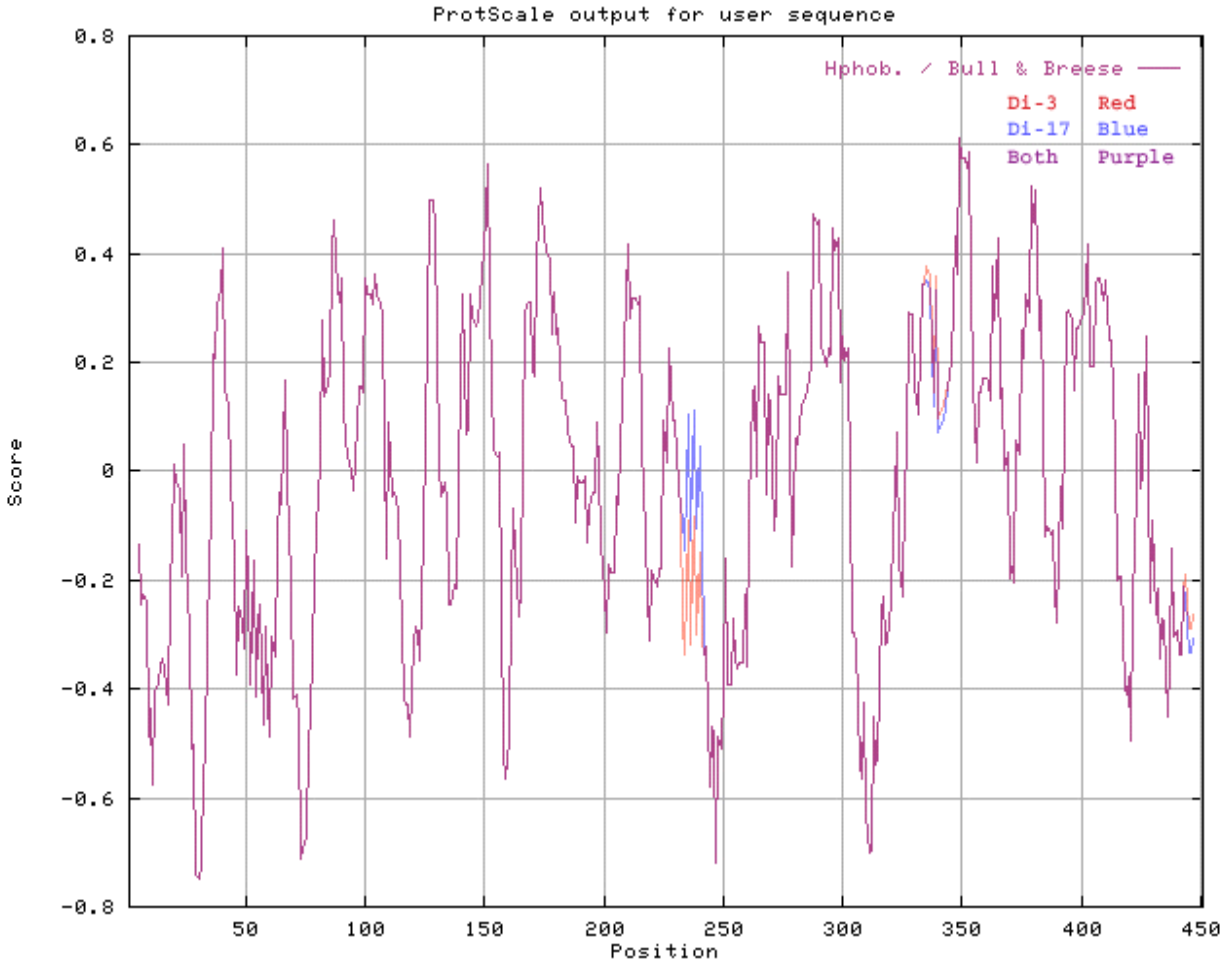
		Scale								
Amino Acid		Abraham & Leo	Bull & Breese	Chothia	Guy	Kyte & Doolittle	Miyazawa et al.	Roseman	Sweet et al.	Welling et al.
Alanine		0.440	0.610	0.380	0.100	1.800	5.330	0.390	-0.400	1.150
Arginine		-2.420	0.690	0.010	1.910	-4.500	4.180	-3.950	-0.590	0.580
Asparagine		-1.320	0.890	0.120	0.480	-3.500	3.710	-1.190	-0.920	-0.770
Aspartic Acid		-0.310	0.610	0.150	0.780	-3.500	3.590	-3.810	-1.310	0.650
Cysteine		0.580	0.360	0.500	-1.420	2.500	7.930	0.250	0.170	-1.200
Glutamine		-0.710	0.970	0.070	0.950	-3.500	3.870	-1.300	-0.910	-0.110
Glutamic Acid		-0.340	0.510	0.180	0.830	-3.500	3.650	-2.910	-1.220	-0.710
Glyceine		0.000	0.810	0.360	0.330	-0.400	4.480	0.000	-0.670	-1.840
Histidine		-0.010	0.690	0.170	-0.500	-3.200	5.100	-0.640	-0.640	3.120
Isoleucine		2.460	-1.450	0.600	-1.130	4.500	8.830	1.820	1.250	-2.920
Leucine		2.460	-1.650	0.450	-1.180	3.800	8.470	1.820	1.220	0.750
Lysine		-2.450	0.460	0.030	1.400	-3.900	2.950	-2.770	-0.670	2.060
Methionine		1.100	-0.660	0.400	-1.590	1.900	8.950	0.960	1.020	-3.850
Phenylalanine		2.540	-1.520	0.500	-2.120	2.800	9.030	2.270	1.920	-1.410
Proline		1.290	-0.170	0.180	0.730	-1.600	3.870	0.990	-0.490	-0.530
Serine		-0.840	0.420	0.220	0.520	-0.800	4.090	-1.240	-0.550	-0.260
Threonine		-0.410	0.290	0.230	0.070	-0.700	4.490	-1.000	-0.280	-0.450
Tryptophan		2.560	-1.200	0.270	-0.510	-0.900	7.660	2.130	0.500	-1.140
Tyrosine		1.630	-1.430	0.150	-0.210	-1.300	5.890	1.470	1.670	0.130
Valine		1.730	-0.750	0.540	-1.270	4.200	7.630	1.300	0.910	-0.130

D-2. Abraham & Lee hydrophobicity overlay



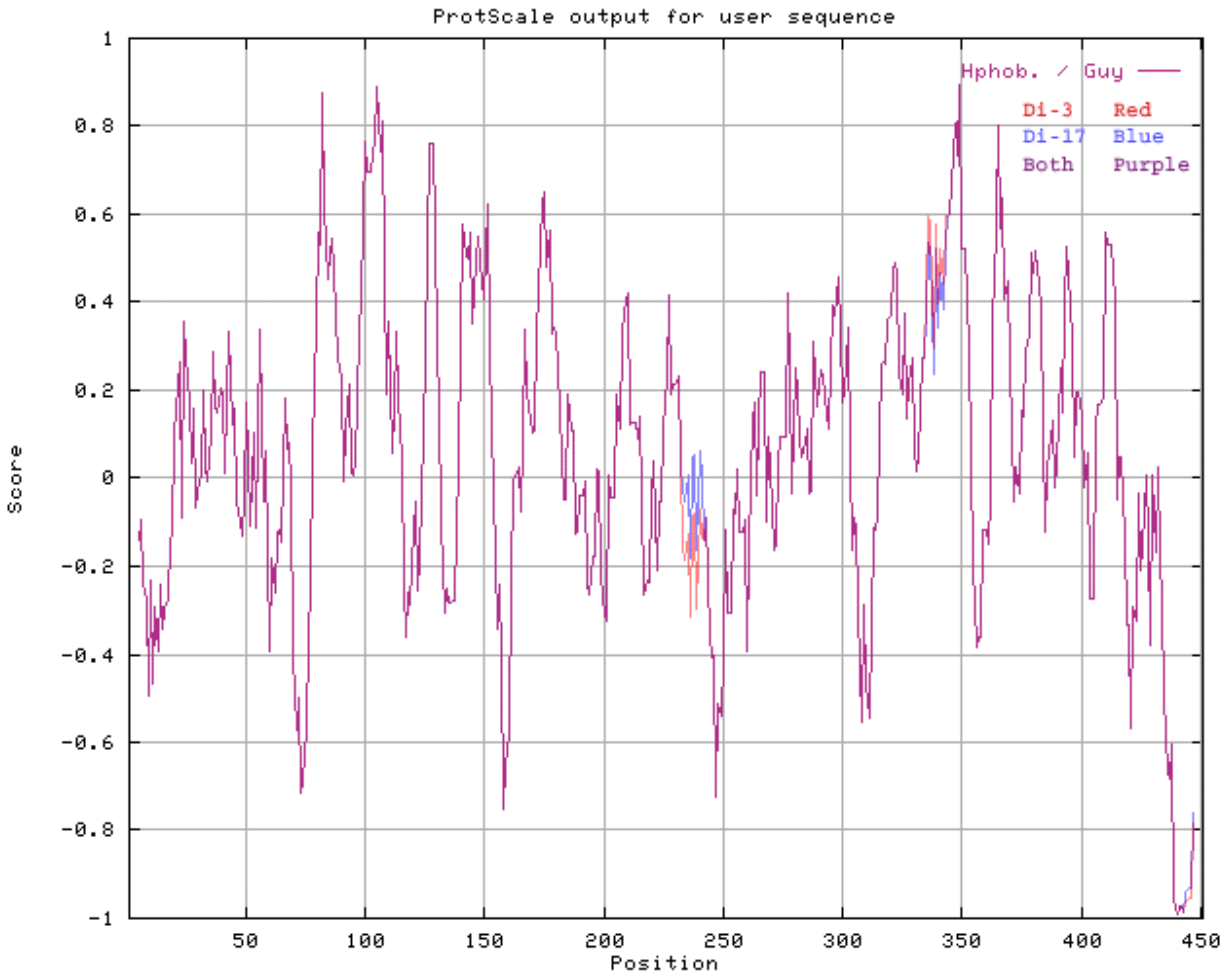
Abraham, D.J. and Leo, A.J. (1987) ***Proteins: Structure, Function and Genetics*** **2: 130-152.**

D-3. Bull & Breese hydrophilicity overlay



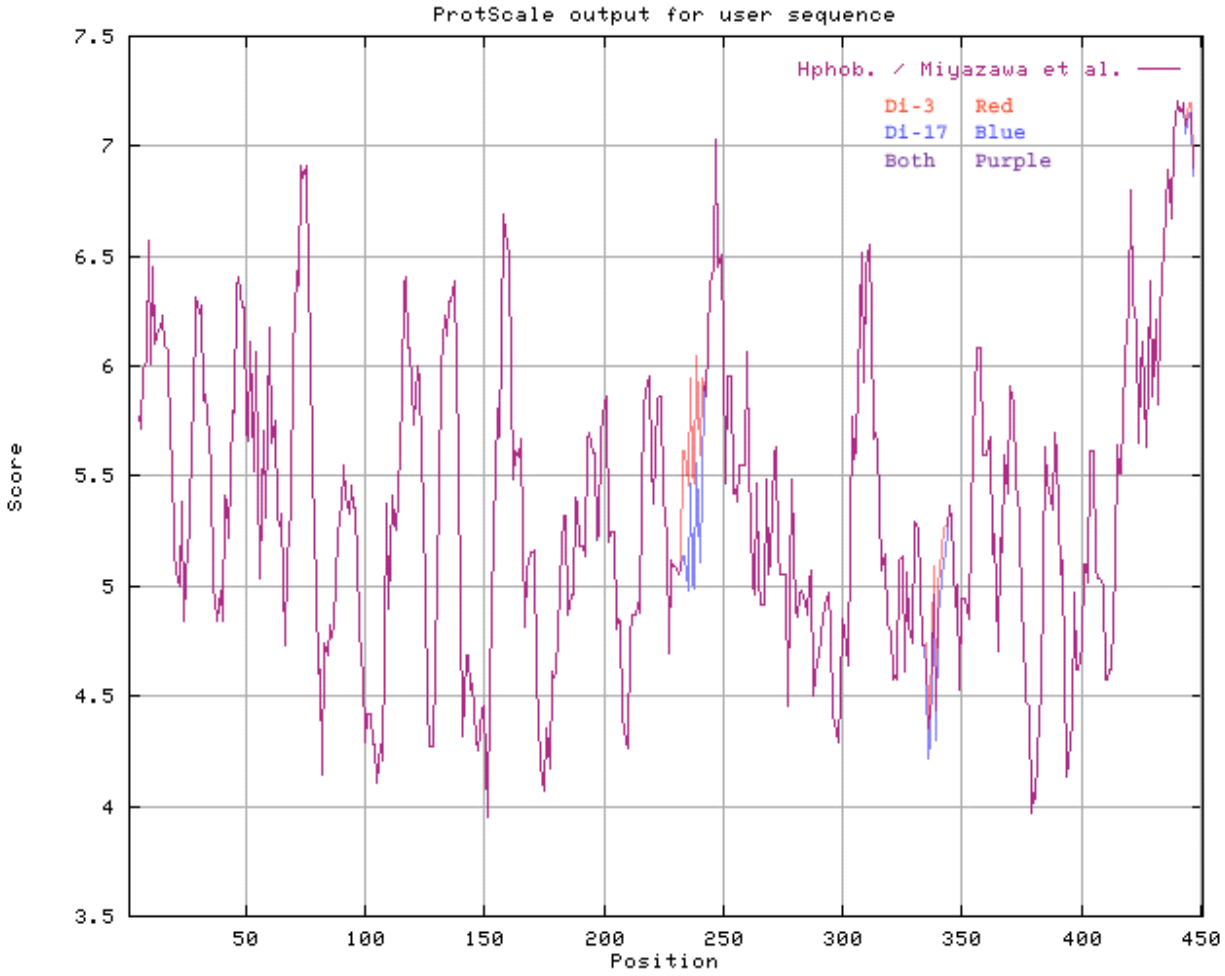
Bull, H.B. and Breese, K. (1974) ***Arch. Biochem. Biophys.*** **161: 665-670.**

D-4. Guy hydrophilicity overlay



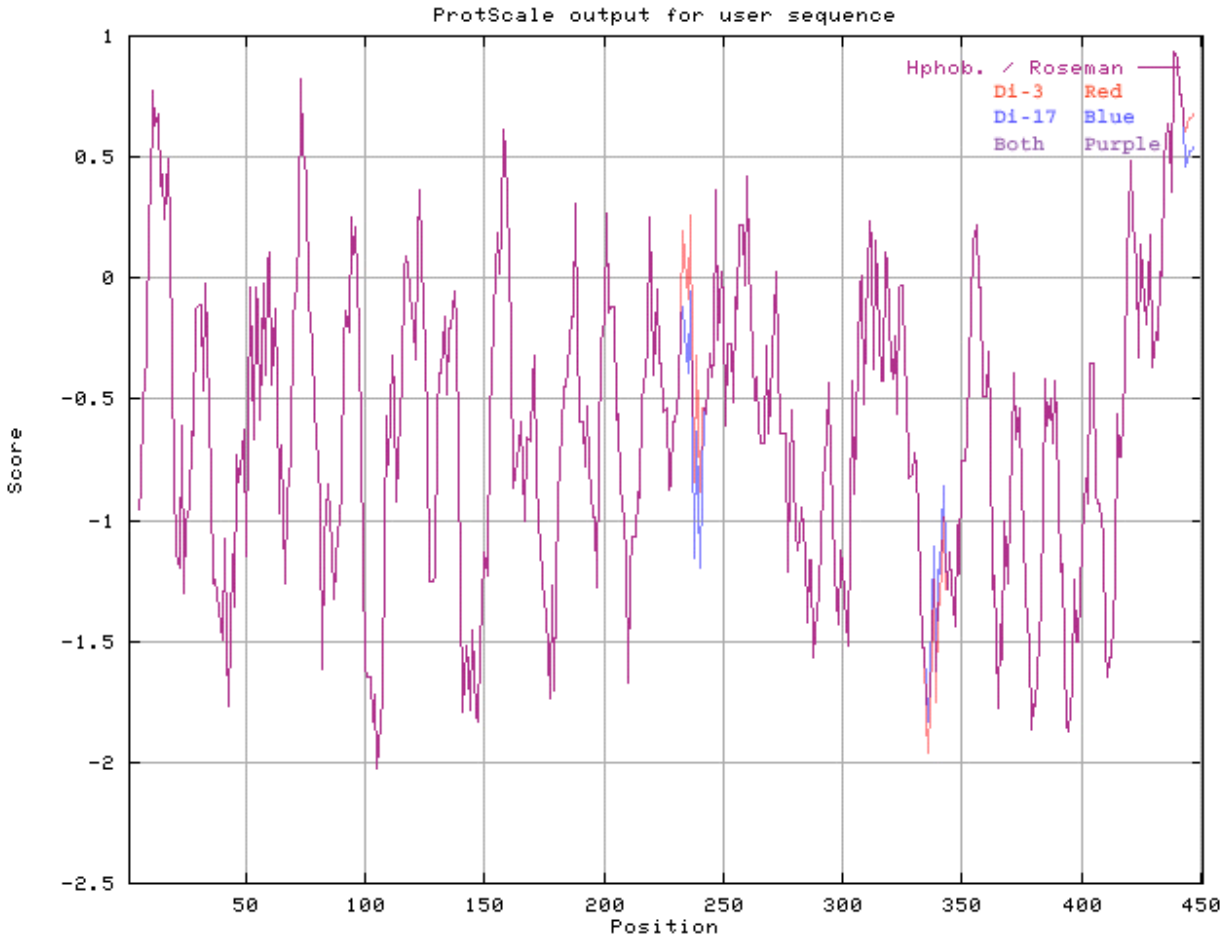
Guy, H.R. (1985) ***Biophys J.*** **47: 61-70**

D-5. Miyazawa et al. hydrophobicity overlay



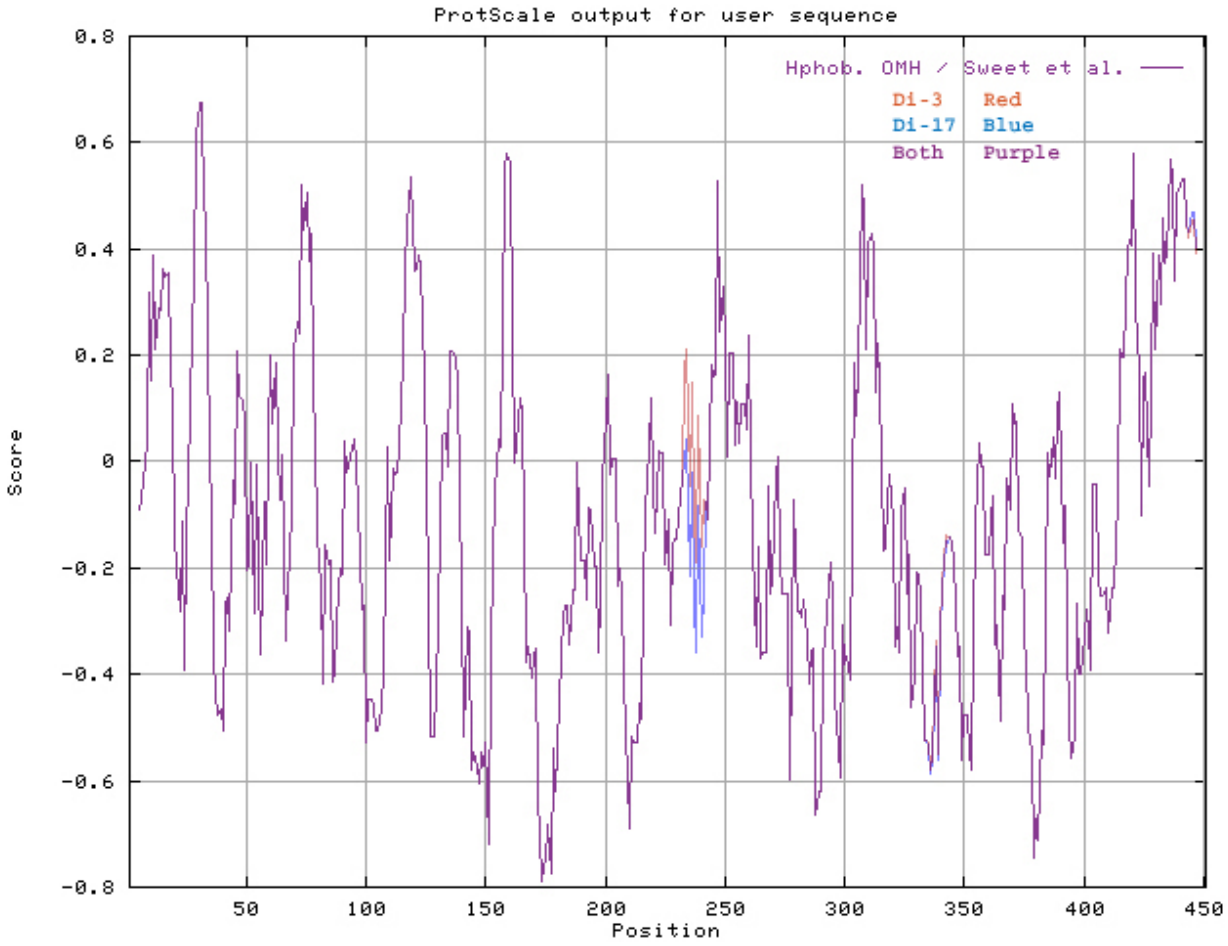
Miyazawa, S. and Jernigen. R.L. (1985) *Macromolecules* 18: 534-552.

D-6. Roseman hydrophobicity overlay



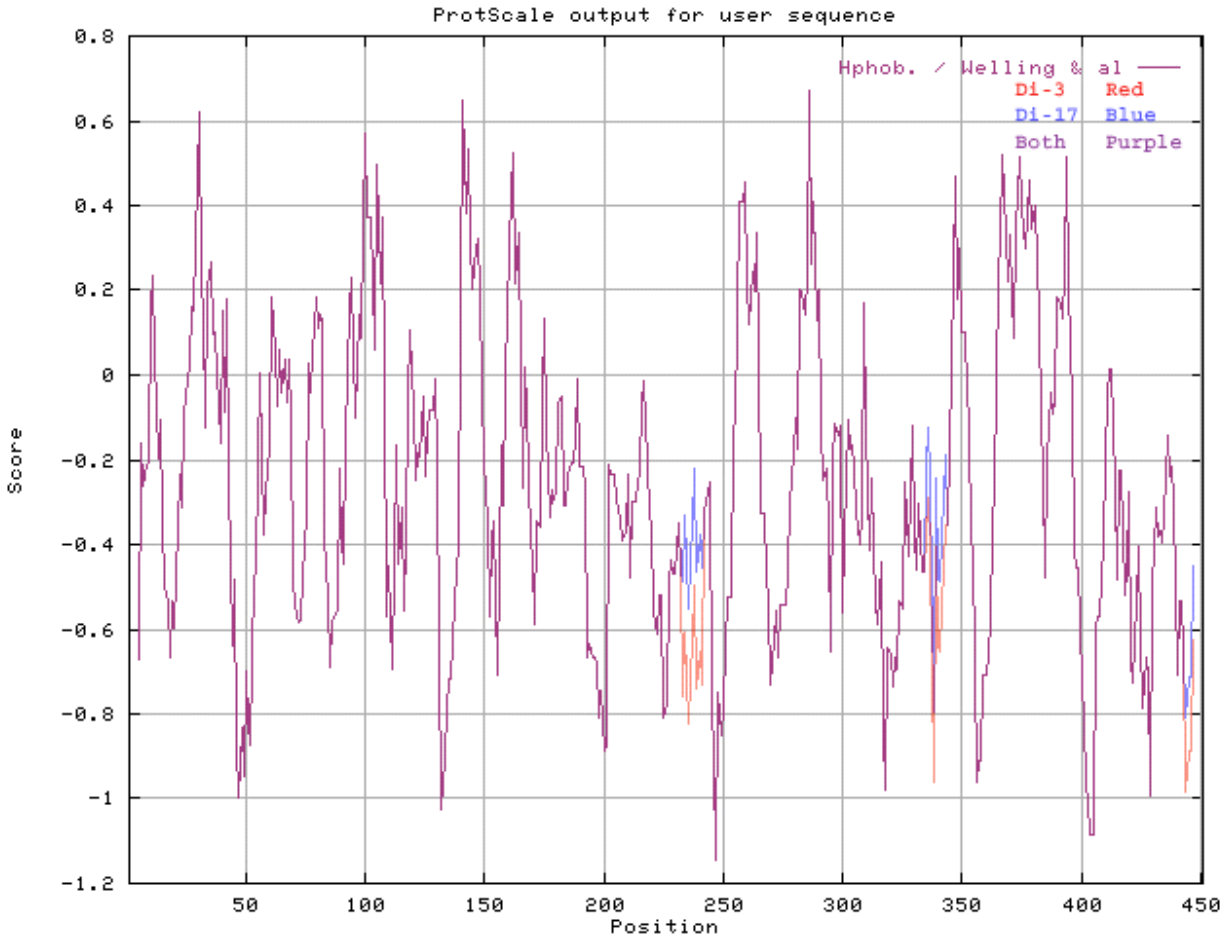
Roseman, M.A. (1988) *J. Mol. Biol.* 200: 513-522.

D-7. Sweet et al. hydrophobicity overlay



Sweet, R.M. and Eisenberg, D. (1983) *J. Mol. Biol.* 171: 479-488.

D-8. Welling et al. hydrophilicity overlay



Welling, G.W., Weijer, W.J., Van der Zee, R., and Welling-Wester, S. (1985) *FEBS Lett.* 188: 215-218.

References

Abraham, D.J., and Leo, A.J. (1987) Extension of the fragment method to calculate amino acid zwitterions and side chain partition coefficients. *Proteins* **2,** 130-152.

Baker, B., Zambryski, P., Staskawicz, B., and Dinesh-Kumar, S.P. (1997) Signaling in Plant-Microbe Interactions. *Science* **276,** 726-733.

Bull, H.B., and Breese, K. (1974) Surface tension of amino acid solutions: a hydrophobicity scale of amino acid residues. *Arch. Biochem. Biophys.* **161,** 665-670.

Burgess, A.W., Ponnuswamy, P.K., and Sheraga, H.A. (1974) Analysis of conformations of amino acid residues and prediction of backbone topography in proteins. *Israel J. Chem.* **12,** 239-286.

Carrington, J.C., Heaton, L.A., Zuidema, D., Hillman, B.I., and Morris, T.J. (1989) The genome structure of turnip crinkle virus. *Virology* **170,** 219-26.

Carrington, J.C., Morris, T.J., Stockley, P.G., and Harrison, S.C. (1987) Structure and assembly of turnip crinkle virus. IV. Analysis of the coat protein gene and implications of the subunit primary structure. *J. Mol. Biol.* **194,** 265-276.

Centre for Molecular and Biomolecular Informatics. (Accessed March 2003) "Hydrophobicity." *www.cmbi.kun.nl/gvteach/alg/infopages/hydrophobicity.shtml*.

Chothia, C. (1976) The nature of the accessible and buried surfaces in proteins. *J. Mol. Biol.* **105,** 1-14.

Cohn, J., Sessa, G., and Martin, G.B. (2001) Innate immunity in plants. *Curr Opin Immun* **13,** 55-62.

Cooley, M.B., Pathirana, S., Wu, H-J, Kachroo, P., and Klessig, D.F. (2000) Members of the Arabidopsis *HRT/RPP8* family of resistance genes confer resistance to both viral and oomycete pathogens. *Plant Cell.* **12,** 663-676.

Dangl, J.L., and Jones, J.D.G. (2001) Plant pathogens and integrated defence responses to infection. *Nature*, **411,** 826-833.

Dele'age, G., and Roux, B. (1987) An algorithm for secondary structure prediction based on class prediction. *Protein Engineering,* **1(4),** 289-294.

Dempsey, D.A., Wobbe, K.K., and Klessig, D.F. (1993) Resistance and susceptible responses of *Arabidopsis thaliana* to turnip crinkle virus. *Mol Plant Pathol* **83,** 1021-1029.

Dempsey, D.A., Pathirana, M.S., Wobbe, K.K., and Klessig, D.F. (1997) Identification of an Arabidopsis locus required for resistance to turnip crinkle virus. *Plant J.* **11,** 301-311.

Fluhr, R. (2001) Sentinels of Disease. Plant Resistance Genes. *Plant Physiology* **127,** 1367-1374.

Garnier, J., Osguthorpe, D.J., and Robson, B. (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of proteins. *J. Mol. Biol.,* **120,** 97-120.

Glazebrook, J. (1999) Genes controlling expression of defense response in *Arabidopsis. Curr Opin Plant Biol* **2,** 280-286.

Guy, H.R. (1985) Amino acid side-chain partition energies and distribution of residues in soluble proteins. *Biophys J.* **47,** 61-70.

Hammond, T.M. (2001) Turnip crinkle virus and the suppression of hypersensitive response. *M.S. Thesis,* Worcester Polytechnic Institute.

Holley, L.H., and Karplus, M. (1989) Protein secondary structure prediction with a neural network. *Proc. Natl. Acad. Sci. USA*. **86,** 152-156.

Holt, B.F., III, Hubert, D.A., and Dangl, J.L. (2003) Resistance gene signaling in plants-complex similarities to animal innate immunity. *Curr Opin Immun* **15,** 20-25.

Ji, C., Smith-Becker, J., and Keen, N.T. (1998) Genetics of plant-pathogen interactions. *Curr Opin Biotechnol* **9,** 202-207.

Jia, Y., McAdams, S.A., Bryan, G.T., Hershey H.P., and Valent, B. (2000) Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *EMBO J* **19,** 4004-4014.

Jones, J. Benton, Jr. (1998) *Plant Nutrition Manual*. CRC Press: USA.

Kachroo, P., Yoshioka, K., Shah, J., Dooner, H.K., and Klessig, D.F. (2000) Resistance to turnip crinkle virus in Arabidopsis is regulated by two host genes and is salicylic acid dependent but *NPR1*, ethylene, and jasmonate independent. *Plant Cell.* **12,** 677-690.

King, R.D., and Sternberg, M.J.E. (1990) Machine learning approach for the prediction of protein secondary structure. *J. Mol. Biol.* **216,** 441-457.

Kyte, J., and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157,** 105-132.

Lam, E., Kato, N., and Lawton, M. (2001) Programmed cell death, mitochondria and

the plant hypersensitive response. *Nature* **411,** 848-853.

Leutwiler, L.S., Hough-Evans, B.R., and Meyerowitz, E.M. (1984) The DNA of *Arabidopsis thaliana. Mol Gen Genet* **194,** 15-23.

Luck, J.E., Lawrence, G.J., Dodds, P.N., Shepherd, K.W., and Ellis, J.G. (2000) Regions outside of the leucine-rich repeats of flax rust. *Plant Cell* **12,** 1367-1377.

Martin, G.B. (1999) Functional analysis of plant disease resistance genes and their downstream effectors. *Curr Opin Plant Biol* **2,** 273-279.

Miyazawa, S., and Jernigan, R.L. (1985) Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation. *Macromolecules* **18,** 534-552.

Montesinos, E., Bonaterra, A., Badosa, E., Francés, J., Alemany, Llorente, I., and Moragrega, C. (2002) Plant-microbe interactions and the new biotechnological methods of plant disease control. *Int Microbiol* **5,** 169-175.

Morris, T.J., and Hacker, D.L. Edited by R.G. Webster and A. Granoff (1994) *Encyclopedia of Virology.* San Diego: Academic Press Ltd.

Nimchuk, Z., Rohmer, L., Chang, J.H., and Dangl, J.L. (2001) Knowing the dancer from the dance: *R*-gene products and their interactions with other proteins from host and pathogen. *Curr Opin Plant Biol* **4,** 288-294.

Pan, Q., Liu, Y.-S., Budai-Hadrian, O., Sela, M., Carmel-Goren, L., Zamir, D., and Fluhr, R. (2000) Comparative genetics of nucleotide binding site-leucine rich repeat resistance gene homologs in the genomes of two dicotyledons: Tomato and Arabidopsis. *Genetics*, **155,** 309-322.

ProtScale Tool. (Accessed March 2003) *http://us.expasy.org/tools/pscale*.

Ren, T., Qu, F., and Morris, T.J. (2000) *HRT* gene function requires interaction between NAC protein and viral capsid protein to confer resistance to turnip crinkle virus. *Plant Cell.* **12,** 1917-1925.

Roseman, M.A. (1988) Hydrophilicity of polar amino acid side-chains is markedly reduced by flanking peptide bonds. *J. Mol. Biol.* **200,** 513-522.

Sander, D.M. (Accessed January 2003) Virus structure. *www.tulane.edu/~dmsander/WWW/335/335Structure.html.*

Scofield, S.R., Tobias, C.M., Rathjen, J.P., Chang, J.H., Lavelle, D.T., Michelmore, R.W., and Staskawicz, B.J. (1996) Molecular basis of gene-for-gene specificity in bacterial speck disease of tomato. *Science* **274,** 2063-2065.

SDSC Biology Workbench v.3.2 (Accessed March 2003) CLUSTALW alignment program and PELE-protein prediction program. *http://workbench.sdsc.edu.*

Simon, A.E., Li, X.H., Lew, J.E., Stange, R., Zhang, C., Polacco, M., and Carpenter, C.D. (1992) Susceptibility and resistance of *Arabidopsis thaliana* to turnip crinkle virus. *Mol Plant-Microbe Interact* **5,** 496-503.

Staskawicz, B.J. (2001) Genetics of plant-pathogen interactions specifying plant disease resistance. *Plant Physiology* **125,** 73-76.

Staskawicz, B.J., Ausubel, F.M., Baker, B.J., Ellis, J.G., and Jones, J.D.G. (1995) Molecular genetics of plant disease resistance. *Science*, **268,** 661-667.

Sweet, R.M., and Eisenberg, D. (1983) Correlation of sequence hydrophobicities measures similarity in three-dimensional protein structure. *J. Mol. Biol.* **171,** 479-488.

Translate Tool (Accessed March 2003) *http://us.expasy.org/tools/dna.html.*

Welling, G.W., Weijer, W.J., Van der Zee, R., and Welling-Wester, S. (1985) Prediction of sequential antigenic regions in proteins. *FEBS Lett.* **188,** 215-218.

Xie, Q., Sanz-Burgos, A.P., Guo, H., Garcia, J.A., and Gutierrez, C. (1999) GRAB proteins, novel members of the NAC domain family, isolated by their interaction with a geminivirus protein. *Plant Mol. Biol.* **39,** 647-656.

Yu, I-C., Parker, J., and Bent, A.F. (1998) Gene-for-gene disease resistance without the hypersensitive response in *Arabidopsis dnd1* mutant. *Plant Biology* **95,** 7819-7824.

Zhao, Y., DelGrosso, L., Yigit, E., Dempsey, D.A., Klessig, D.F., and Wobbe, K.K. (2000) The amino terminus of the coat protein of turnip crinkle virus is the AVR factor recognized by resistant Arabidopsis. *Mol. Plant Microbe. Interact.* **9,** 1015-1018.