MOLECULAR AND CELLULAR BIOLOGY, Feb. 2001, p. 1045–1057 Vol. 21, No. 4
0270-7306/01/$04.00+0 DOI: 10.1128/MCB.21.4.1045–1057.2001

Identification of Specificity Determinants and Generation of Alleles with Novel Specificity at the *het-c* Heterokaryon Incompatibility Locus of *Neurospora crassa*

JENNIFER WU[1]† AND N. LOUISE GLASS[1,2]*

The Biotechnology Laboratory and The Department of Botany, The University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada,[1] and The Plant and Microbial Biology Department, The University of California, Berkeley, California 94720[2]

Received 2 August 2000/Returned for modification 10 October 2000/Accepted 27 October 2000

The capacity for nonself recognition is a ubiquitous and essential aspect of biology. In filamentous fungi, nonself recognition during vegetative growth is believed to be mediated by genetic differences at heterokaryon incompatibility (*het*) loci. Filamentous fungi are capable of undergoing hyphal fusion to form mycelial networks and with other individuals to form vegetative heterokaryons, in which genetically distinct nuclei occupy a common cytoplasm. In *Neurospora crassa*, 11 *het* loci have been identified that affect the viability of such vegetative heterokaryons. The *het-c* locus has at least three mutually incompatible alleles, termed *het-c*^{OR}, *het-c*^{PA}, and *het-c*^{GR}. Hyphal fusion between strains that are of alternative *het-c* specificity results in vegetative heterokaryons that are aconidial and which show growth inhibition and hyphal compartmentation and death. A 34- to 48-amino-acid variable domain, which is dissimilar in HET-C^{OR}, HET-C^{PA}, and HET-C^{GR}, confers allelic specificity. To assess requirements for allelic specificity, we constructed chimeras between the *het-c* variable domain from 24 different isolates that displayed amino acid and insertion or deletion variations and determined their *het-c* specificity by introduction into *N. crassa*. We also constructed a number of artificial alleles that contained novel *het-c* specificity domains. By this method, we identified four additional and novel *het-c* specificities. Our results indicate that amino acid and length variations within the insertion or deletion motif are the primary determinants for conferring *het-c* allelic specificity. These results provide a molecular model for nonself recognition in multicellular eucaryotes.

The ability to distinguish self from nonself is a critical feature in most multicellular eucaryotic organisms for maintenance of integrity and individuality. In filamentous fungi, nonself recognition during vegetative growth is mediated through a system known as vegetative or heterokaryon incompatibility (22, 29, 37). A filamentous fungal individual grows as an interconnected network of multinuclear hyphal filaments that are formed via hyphal self-fusion. Filamentous fungi also possess the remarkable attribute of being able to undergo hyphal fusion between different individuals to form vegetative heterokaryons (genetically different nuclei in a common cytoplasm). However, if fungal individuals undergo hyphal fusion but differ in allelic specificity at any one of a number of heterokaryon incompatibility loci (*het*; sometimes referred to as *vic* for vegetative incompatibility), the hyphal fusion cell is compartmentalized and dies (Fig. 1). Heterokaryon incompatibility is believed to play a role in filamentous fungi to prevent the spread of mycoviruses and debilitated organelles throughout fungal populations and to restrict resource plundering between individuals (7, 11, 12).

In *Neurospora crassa*, 11 *het* loci have been genetically characterized (33, 35); the *het-c* locus was one of the first *het* loci identified (17). Vegetative heterokaryons forced by auxotrophic markers between strains that differ in *het-c* specificity or partial diploids that are heterozygous at *het-c* are aconidial, show greatly reduced growth rates, and exhibit hyphal compartmentation and death (17, 18, 25, 31, 34). Wild-type isolates fall into three *het-c* specificity groups (*het-c*^{OR} compatible, *het-c*^{PA} compatible, or *het-c*^{GR} compatible) based on results from crosses using translocation strains (32) that generate *het-c* partial diploid progeny (24, 39). Genetic differences at *het* loci in *N. crassa* do not interfere with sexual fertility (34). Representatives from the three distinct and mutually incompatible allele types (*het-c*^{OR}, *het-c*^{PA}, and *het-c*^{GR}) have been molecularly characterized (39, 40). The *het-c* locus encodes a polypeptide containing a consensus signal peptide sequence and a glycine-rich carboxyl-terminal domain (40). By chimeric construction between the three *het-c* allele types, it was determined that a 34- to 48-amino-acid (aa) variable domain (which is dissimilar in HET-C^{OR}, HET-C^{PA}, and HET-C^{GR}) confers allelic specificity (39). The *het-c* specificity domain of *het-c*^{OR}, *het-c*^{PA}, and *het-c*^{GR} differs in both predicted amino acid sequence and in the pattern of insertion and deletion (indel) (Fig. 2).

A molecular survey of the *het-c* variable domain in 15 *N. crassa* isolates and 25 isolates from related species and genera showed that all could be placed into one of the three previously identified *het-c* specificities based on indel motif (39, 48). Phylogenetic analyses based on DNA sequences of the *het-c* variable domain from *N. crassa* and related species showed trans species polymorphisms, i.e., isolates did not group according to genus or species but rather grouped according to DNA sequence type of the *het-c* specificity domain (48). Additional

* Corresponding author. Mailing address: Department of Plant and Microbial Biology, 111 Koshland Hall, The University of California, Berkeley, CA 94720-3102. Phone: (510) 643-2399. Fax: (510) 642-4995. E-mail: Lglass@uclink.berkeley.edu.

† Present address: Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109.

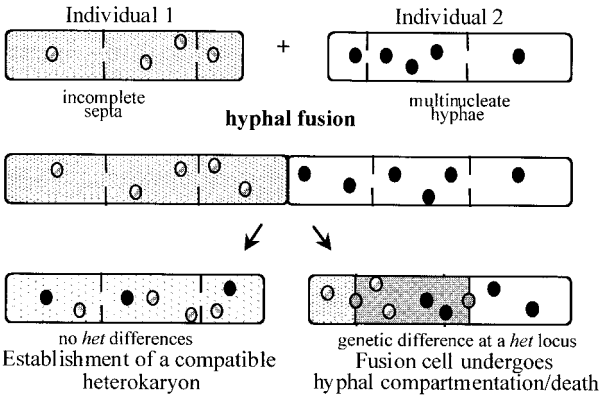

FIG. 1. Diagrammatic representation of the consequences of hyphal fusion between fungal individuals that do, or do not, differ in specificity at heterokaryon incompatibility (*het*) loci, such as *het-c* in *N. crassa*.

clades within each *het-c* specificity group were observed, suggesting that diversity within each *het-c* indel type could represent additional *het-c* specificities. Analysis of other loci that mediate nonself recognition, such as loci in the major histocompatibility complex (MHC) and the *S* locus in plants, also show multiple allelic polymorphisms that display transspecies polymorphisms (reviewed in reference 27). These data indicate that, like the MHC and the *S* locus, the *het-c* locus is subject to balancing selection and suggests that its function in mediating heterokaryon incompatibility is biologically significant as a nonself recognition system in this group of fungi.

Common molecular features are apparent among loci that mediate self and nonself recognition. Alleles that confer specificity are polymorphic and recognition is generally mediated by protein-protein interactions. The mechanism of allelic specificity has been examined in several fungal nonself recognition systems by the construction of chimeric (or hybrid) alleles. In *Podospora anserina*, a single amino acid difference in the alternative proteins encoded by the vegetative incompatibility locus, *het-s*, was sufficient to confer allelic specificity (14). In *Ustilago maydis*, a region composed of 30 to 48 amino acid residues was identified that regulates specificity at the *b* mating locus (49); artificial hybrid *b* alleles with novel specificity were generated by chimeric construction within this border region (50). In *Coprinus cinereus*, specificity of the homeodomain mating proteins HD1 and HD2 was determined by the N-terminal 160 to 170 aa (1). In *N. crassa*, amino acid sequence differences within the *het-c* variable domain could be the critical determinant for *het-c* specificity, or spatial differences within the indel motif could be the most important factor. To differentiate these two possibilities, we took advantage of the natural amino acid variation and indel motifs observed in the *het-c* variable domain among natural isolates (48) to construct chimeric alleles. We first asked whether the *het-c* specificity of these alleles, as assayed in *N. crassa*, was consistent with their grouping based on phylogenetic analyses. Second, we asked whether *het-c* specificity could be affected either by amino acid sequence and/or indel pattern within the *het-c* specificity domain. Third, we constructed a number of artificial *het-c* alleles that contained combinations of amino acid sequences and indel motifs that were not observed in our survey of naturally occurring *het-c* alleles. By this method, we were able to identify four additional, novel *het-c* specificities. Our results indicate that amino acid and spatial characteristics of the indel motif are the primary determinant for conferring *het-c* allelic specificity. These findings provide insight into understanding the molecular mechanism of nonself recognition during heterokaryon incompatibility in filamentous fungi and provide a molecular model for allelic specificity and nonself recognition mechanisms in multicellular eucaryotes.

MATERIALS AND METHODS

Strains and media. *Escherichia coli* strain DH5α (F^- *endA1 hsdR17 supE44 lacZM15*) (Bethesda Research Laboratories, Gaithersburg, Md.) was used for routine DNA manipulation work. Strains used for chimeric construction and transformation recipients are listed in Table 1, along with their origin and *het-c* specificity. All strains were grown on Vogel's vegetative growth media (VM) (46). A cross to construct *het-c* deletion strain CJ44 was performed between X22-2 and Xa-3 (Table 1; Q. Xiang and N. L. Glass, unpublished results) by using Westergaard's media and standard crossing conditions (10) and selecting for progeny that formed heterokaryons compatible with both *het-c*^{OR} and *het-c*^{PA} strains.

Construction of chimeric alleles. Construction and identification of *het-c* chimeric alleles were as previously described (39). Previous results indicated that the *het-c* variable domain was necessary and sufficient to confer allelic specificity; amino acid differences outside of this region between HET-C^{OR}, HET-C^{PA}, and HET-C^{GR} did not affect allelic specificity (underlined region in Fig. 2) (39). Plasmids carrying *N. crassa het-c*^{OR} and *het-c*^{PA} alleles were used for chimeric allele construction (39), depending on availability of restriction sites. A unique DNA 220-bp *Stu*I-*Sal*I or 650-bp *Eco*RV-*Sal*I fragment encompassing the *het-c* variable domain from naturally occurring and artificial alleles was exchanged in frame with an otherwise *het-c*^{OR} or *het-c*^{PA} allele. Chimeric alleles were identified by restriction digests using conserved *Xho*I or *Apa*I restriction site differences located within the exchanged fragment. The chimeric alleles were cloned into pCB1004 (6) or pOKE103 vector (gift of R. L. Metzenberg) to test for *het-c* specificity in *N. crassa* by transformation assays.

Generation of artificial *het-c* specificity regions by PCR mutagenesis. The specificity domains for artificial alleles were generated by a recombinant PCR technique (23, 45) that requires two sets of primers and two rounds of PCR. The oligonucleotides used in the construction of artificial *het-c* specificity domains are listed in Table 2. Plasmids containing *het-c* were used as first-round PCR templates. The first round of PCR was performed by standard protocol: initial 5-min denaturation at 94°C, followed by 30 step cycles of 1 min at 94°C, 1 min at 55°C, and 1 min at 72°C, and a final extension at 72°C for 10 min. Reactions were carried out in volumes of 50 μl containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2.0 mM $MgCl_2$, 100 mM (each) dATP, dCTP, dGTP, and dTTP, 0.2 mM concentrations of each primer, 1.25 U of *Taq* polymerase, and 50 ng of plasmid DNA as template. The first-round PCR products were purified from agarose gel using a QIAEXII Gel Purifying kit (Qiagen Inc., Mississauga, Ontario, Canada). Two corresponding first-round PCR products were combined and used as templates for the second round of PCR to construct recombinant products. The PCR conditions resembled those of the first round except that annealing was done at 51°C. PCR products from the second round were cloned into pCRII vector (Invitrogen, San Diego, Calif.) for further manipulations. The predicted DNA sequence of the artificial constructs was confirmed by DNA sequence analysis (Nucleic Acid and Protein Synthesis Unit; The University of British Columbia).

Secondary-structure predictions. Secondary-structure predictions of the HET-C variable domain were examined using Gibrat (21), Levin (30), DPM (13), and SOPMA (19, 20) prediction programs (http://www.bcp.fr) and the proteomics tools at the ExPASy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (SIB) (http://www.expasy.ch/).

Transformation assays. *N. crassa* spheroplasts were prepared as described by Schweizer et al. (41). Strains C9-2, C2-2-9, and FGSC2193 (Table 1) were used as recipients for transformation assays with the pCB1004 (*hyg*^{R}) (6) vector constructs. Strain CJ44 (Table 1) was used for cotransformation with pCB1004 and pOKE103 (*pan-2*$^+$; gift of R. L. Metzenberg) vector constructs. For transformation experiments, a modified procedure from *P. anserina* transformation (4) was applied to *N. crassa*. Fifty microliters of spheroplasts was thawed on ice and subsequently heat shocked at 48°C for 5 min, followed by a 30-s incubation on ice. The spheroplasts were then placed at room temperature for 10 min before

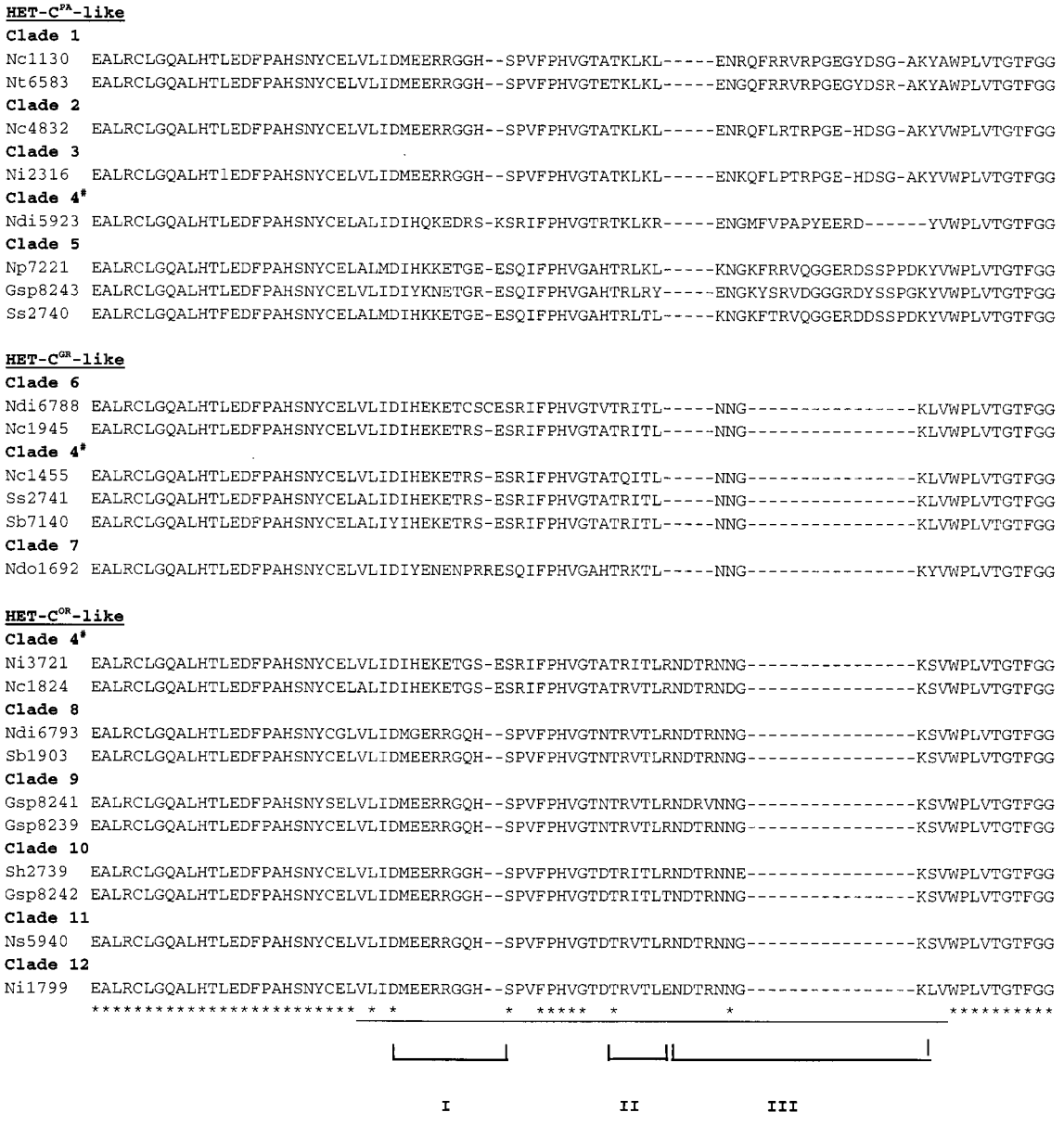

FIG. 2. Predicted peptide sequences of the variable domain in naturally occurring *het-c* alleles. The underlined area is the *het-c* variable domain that is necessary and sufficient to confer *het-c* allelic specificity in the *N. crassa het-c*^{OR} (FGSC 2489), *het-c*^{PA} (FGSC 1130), and *het-c*^{GR} (FGSC1945) alleles (39, 40). *, conserved sites. Dashes represent deletions. Regions I and II are polymorphic blocks; region III encompasses the indel motif. Each clade is supported by over 94% bootstrap values (48). HET-C^{PA}-like, HET-C^{OR}-like, and HET-C^{GR}-like refers to the indel pattern observed in the genetically characterized *N. crassa het-c*^{PA}, *het-c*^{OR} and *het-c*^{GR} alleles, respectively (39). Clade 4 has representatives in all three HET-C-like classes based on the indel motif; separation of these three *het-c* types is supported by bootstrap values of less than 94%. GenBank accession numbers for nucleotide sequences of these alleles were previously cited (48); accession numbers are AF092695 to AF092732 and AF093679.

DNA was added. One microgram of DNA construct was used for each 50 μl of spheroplasts. The mixture of spheroplasts and DNA was incubated at room temperature for an additional 10 min and then added to 1 ml of 40% polyethylene glycol (3350)–10 mM MOPS (morpholine propanesulfonic acid)–50 mM $CaCl_2$. After 10 min, 7 ml of prewarmed top agar was added into the mixture and poured onto agar plates containing 250 μg of hygromycin (Calbiochem, San Diego, Calif.)/ml or media lacking pantothenic acid. For each transformation experiment, 15 individual transformants were transferred to separate VM plates containing 200 μg of hygromycin/ml or lacking pantothenic acid and were incubated at 30°C. Transformants were inspected for up to a week for growth inhibition or altered morphology that is characteristic of *het-c* incompatibility

TABLE 1. List of *N. crassa* strains

Strain	Genotype	Source or reference
C2-2-9	*het-c*^{OR} *thr-2 A*	43
C9-2	*het-c*^{PA} *thr-2 a*	43
FGSC2193	*NM149 T(II;V) het-c*^{GR} *A*	32; D. D. Perkins
C15-1	*het-c*^{OR} *arg-5; pan-2; inl a*	R. Todd and L. Glass
Xa-3	*het-c*^{PA} *arg-5; pan-2 A*	Q. Xiang and L. Glass
X22-2	Δ*het-c thr-2 a*	Q. Xiang and L. Glass
CJ44	Δ*het-c arg-5; pan-2 A*	This study

TABLE 2. Oligonucleotide primers used for the construction of artificial *het-c* alleles

Primer name	Sequence[a]
Red	GGA GAC ATG GCG ATA TCG
SP1.3	GCT CAT GCC AGG AAC AAC
del3.3	AGC GTA CTT GAG TTT GAG TTT AGT CGC AGT GC
del3.5	ACT CAA ACT CAA GTA CGC TTG GCC CCT GGT
po1.3	CTG AAC TGT CCA TTG TTA CGT GTG TCA
po1.5	CAC GTA ACA ATG GAC AGT TCA GGA GCG TAA GA
po2.3	GTC TAT TCT CAC GTG TGT CAT TCT TGA G
po2.5	GAC ACA CGT GAG AAT AGA CAG TTC AGG CGC GTA AGA
po3.3	CTG TCT ATT CTC GAG TGT GAT TCG CGT ATC
po3.5	CAC ACT CGA GAA TAG ACA GTT CAG GCG CGT AAG A
po4.3	GTG TCA TTC TTG AGT TTG AGT TTA GTC GC
po4.5	CTC AAA CTC AAG AAT GAC ACA CGT AAC
p26m-3	CGC CCC AGA ATC ATC TCT TTC CTC ATA TGG
p26m-5	GAA AGA GAT GAT TCT GGG GCG AAG TAC ACG
pd1.3	AGC GTA CTT TCC TTC TCC GGG TCT TAC
pd1.5	GGA GAA GGA AAG TAC GCT TGG CCC CTG
pd2.3	AGC ATA CTT TAC GCG CCT GAA CTG TCT
pd2.5	AGG CGC GTA AAG TAT GCT TGG CCC CTG
pd3.3	AAG CGT ACT TCG GTC TTA CGC GCC TGA A
pd3.5	GTA AGA CCG AAG TAC GCT TGG CCC CTG
pd4.3	AGC GTA CTT TCT ATT CTC GAG TTT GAG
pd4.5	CTC GAG AAT AGA AAG TAC GCT TGG CCC CTG
od-3	TCC ATT GTT GAG TGT GAT TCG AGT ATC TG
od-5	ATC ACA CTC AAC AAT GGA AAG TCG GTT
go-3	CTT ACG TGT GTC ATT CTT GAG TGT GAT TCG AGT AGC
go-5	AAG AAT GAC ACA CGT AAG TTG GTT TGG CCC TTG G

[a] The primer sequences are shown in the 5′-to-3′ orientation.

(18). For cotransformation experiments, 0.5 μg of each plasmid was used per 50 μl of spheroplasts.

Growth rate determinations. The linear growth rate (LGR) of transformants was measured in race tubes as previously described (10). Individual colonies cut from a transformation plate were placed at one end of a 40-ml glass tube containing 25 ml of VM, plus supplements and/or hygromycin. All tubes were incubated at 25°C. The starting point was marked as the leading edge of the colony after overnight growth; subsequent growth was recorded as the distance to the leading edge of the colony at 24-h intervals. Each experiment contained at least three replicate transformants for each construct, plus controls.

Light microscopy. The stain Evan's Blue (Direct Blue 53, CI23860; Aldrich Chemical Co., Milwaukee, Wis.) is excluded by cells with intact plasma membranes (16) and was used to identify dead hyphal compartments (25). Transformants were inoculated on cellophane membrane layered on top of petri dishes containing selective media and incubated at 30°C. Cellophane membranes with adherent mycelium were removed from the medium, placed on glass slides, and flooded with 1% (wt/vol) Evans blue in water. After 5 min, mycelia were rinsed in distilled water, floated off cellophane onto the slide, and mounted in glycerol-phosphate buffer under a coverslip. Samples were examined under bright-field illumination on an Olympus microscope.

RESULTS

Phylogenetic and structural features of the *het-c* variable domain. Allelic specificity of *het-c* is dependent upon a 34- to 48-aa region (Fig. 2), which is variable between HET-C^{OR}, HET-C^{PA}, and HET-C^{GR}; exchange of this region by chimeric construction is sufficient to switch allelic specificity (39). The variable domain differs in both predicted amino acid sequence and in indel motif between *het-c*^{OR}, *het-c*^{PA}, and *het-c*^{GR} alleles. Phylogenetic analysis of the *het-c* variable domain among *N. crassa* isolates and related species and genera showed 12 major clades, which were supported by bootstrap levels of over 94% (48) (Fig. 2). Isolates within each clade group by *het-c* indel motif, with the exception of clade 4, which includes *het-c*^{OR}-like, *het-c*^{GR}-like, and *het-c*^{PA}-like isolates.

Two regions that show high amino acid diversity are designated "I" and "II" in Fig. 2. The 24 isolates fall into two groups based on amino acid variations within region I. The first group contains a consensus sequence of M(G)EERRGG(Q)H and includes isolates that have either a HET-c^{PA}-like or HET-c^{OR}-like indel motif. The second group of alleles is much more variable in amino acid sequence in region I but has a consensus of IH(Y)E(Q/K)K(N)ET(N/D)G(R/P/C)S(E/R). All of the HET-c^{GR}-like peptides have this consensus sequence. Secondary-structure predictions indicate that both amino acid variations found in region I would form an α-helix. The second variable region (II; Fig. 2) has a consensus sequence of T(A)XTR(Q)L(I/V/K)T(R/K)L(Y/R); only the second T residue is conserved among all 24 alleles. Although this region is variable, it is predicted to form or be part of a β-sheet (antiparallel) structure.

Two regions that are highly conserved among all 24 predicted peptides immediately bracket the indel region, I(V)FPHVG and WPLVTGTF (Fig. 2). The I(V)FPHVG region is predicted to be part of a β-sheet with adjacent sequences that are part of region II. The second highly conserved region flanking the indel region, WPLVTGTF, is also predicted to form a β-sheet (antiparallel) structure. The indel region (III; Fig. 2) is predicted to form a loop or coil. These secondary-structure predictions indicate that the variable indel loop region is flanked by the two conserved regions predicted to form antiparallel β-strands.

Introduction of different *het-c* constructs yields three phenotypic classes of transformants. To determine whether the 12 *het-c* phylogenetic branches each confer a different *het-c* allelic specificity (Fig. 2) or whether they fall into one of the three *het-c* specificities previously identified (*het-c*^{OR}, *het-c*^{PA}, or *het-*

TABLE 3. Phenotypes of transformants containing constructs of naturally occurring *het-c* alleles[a]

Allele	C9-2 (PA)	C2-2-9 (OR)	FGSC2193 (GR)
het-c^{PA}-like			
Nc1130	C	I^2	C[b]
Nc6583	C	I^2	C
Ni4832	C	I^2	C
Nt2316	C	I^2	C
Ndi5923	I^2	I^2	I^2
Np7221	C	I^2	C
Gsp8243	C	I^2	C
Ss2740	C	I^2	C
het-c^{GR}-like			
Ndi6788	I^2	I^2	C
Nc1945	I^2	I^2	C
Nc1455	I^2	I^2	C
Ss2741	I^2	I^2	C
Sh7140	I^2	I^2	C
Ndo1692	I^2	I^2	C
het-c^{OR}-like			
Ni3721	I^3	C	I^3
Nc1824	I^3	C	I^3
Ndi6793	I^3	C	I^3
Sb1903	I^3	C	I^3
Gsp8241	I^3	C	I^3
Gsp8239	I^3	C	I^3
Sh2739	I^3	C	I^3
Gsp8242	I^3	C	I^3
Ns5940	I^3	C	I^3
Ni1799	I^3	C	I^3

[a] I^2, incompatible transformants with an LGR of 1 to 2.5 cm/day showing ~20 to 30% HCD; I^3, severely incompatible transformants with an LGR of <1 cm/day showing 20 to 30% HCD; C, fully compatible transformants with an LGR of 4 to 6.5 cm/day with <1% HCD.

[b] Transformation results (39) and analysis of partial diploid progeny (24) indicate that strain FGSC2193 has a recessive allele-specific suppressor for *het-c* incompatibility. FGSC2193 and C9-2 are not isogenic strains and probably differ not only in *het* genotype but also at a large number of other loci that may affect the morphological phenotype associated with *het-c* vegetative incompatibility.

c^{GR} compatible), we assessed the *het-c* specificity of chimeric alleles by transformation assays with *N. crassa*. Chimeric constructs between the *het-c* variable domain from the 24 naturally occurring alleles (Fig. 2) and an otherwise *het-c*^{OR} and/or *het-c*^{PA} allele were constructed (see Materials and Methods). Each chimeric construct was introduced into strains that differed in *het-c* specificity, C2-2-9 (*het-c*^{OR}), C9-2 (*het-c*^{PA}), and FGSC2193 (*het-c*^{GR}) (Table 1). Three phenotypic classes could be distinguished by the morphology of colonies, LGR, and occurrence of hyphal compartmentation and death (HCD) in the transformants. Compatible (class 1) transformants [for example, introduction of a *het-c*^{PA} allele into C9-2 (PA); Table 3] displayed vigorous growth and conidiation with an LGR of 4.0 to 6.5 cm/day (Fig. 3). Less than 1% HCD was observed in compatible transformants. Class 2 (intermediate incompatible) transformants [for example, introduction of a *het-c*^{PA} allele into a C2-2-9 (OR); Table 3] had an LGR of 1.0 to 2.5 cm/day, were aconidial, and exhibited a swollen hyphal morphology (Fig. 3). HCD (approximately 20%) was observed after 2 days of growth. Class 3 transformants [for example, the introduction of a *het-c*^{OR} allele into C9-2 (PA); Table 3] displayed severely inhibited growth with an LGR of <1.0 cm/day. These transformants showed a flat, curling, aconidial morphology and approximately 20 to 30% dead hyphal compartments (Fig. 3).

Chimeric alleles containing *het-c*^{OR}-like variable domains conferred *het-c*^{OR} allelic specificity. The *het-c*^{OR}-like alleles fell into six distinct clades by phylogenetic analysis (48) (Fig. 2). All of the *het-c* chimeric constructs containing a *het-c*^{OR}-like indel motif produced class 3, severely incompatible transformants when introduced into C9-2 (PA) and FGSC2193 (GR) (Table 3). Fully compatible transformants were obtained when these constructs were introduced into C2-2-9 (OR). The chimeric *het-c*^{OR}-like alleles thus displayed an identical *het-c* specificity to a *het-c*^{OR} allele, even though amino acid variability occurred in both regions I and II (Fig. 2). In particular, *het-c*^{OR}-like chimeric alleles containing either of the two motifs in region I displayed identical *het-c* specificity, for example, *Ni3721* and *Sb1903* (Table 3).

Chimeric *het-c*^{GR}-like alleles conferred a *het-c* specificity identical to that of *N. crassa het-c*^{GR}. The *het-c*^{GR}-like alleles group together based on indel motif within the *het-c* variable domain but show three distinct clades by phylogenetic analyses (48). The *het-c*^{GR}-like alleles lack the *het-c*^{PA}-like and *het-c*^{OR}-like insertions and have a similar motif in regions I and II, although some amino acid variability is present (Fig. 2). The introduction of all of the *het-c*^{GR}-like chimeric constructs into C9-2 (PA), C2-2-9 (OR), and FGSC2193 (GR) yielded identical transformation results as a canonical *het-c*^{GR} allele (Table 3). Thus, the observed variations in the predicted amino acid sequence among the *het-c*^{GR}-like alleles did not affect *het-c* allelic specificity. In particular, although phylogenetic analysis showed a relationship between *het-c*^{GR}-like alleles *Nc1455*, *Ss2741*, and *Sb7140* and *het-c*^{OR}-like alleles *Ni3721* and *Nc1824* (clade 4, Fig. 2), *Nc1455*, *Ss2741*, and *Sb7140* conferred *het-c*^{GR} specificity while *Ni3721* and *Nc1824* conferred *het-c*^{OR} specificity.

Chimeric *het-c*^{PA}-like alleles conferred an identical *het-c* specificity as a *het-c*^{PA} allele with the exception of *Ndi5923*. The *het-c*^{PA}-like alleles contain a 42- to 48-bp insertion (14 to 16 aa) compared to *het-c*^{GR}-like and *het-c*^{OR}-like alleles. Phylogenetic analysis showed that these *het-c*^{PA}-like alleles fell into five distinct clades (Fig. 2). A high degree of nonsynonymous substitutions among these alleles results in amino acid variation in regions I and II and especially in the *het-c*^{PA}-specific insertion (III; Fig. 2). Secondary-structure predictions of the indel motif in the HET-C^{PA}-like peptides showed an additional short β-sheet, plus a variable loop between the conserved antiparallel β-strands. When the *het-c*^{PA}-like chimeric constructs were introduced into C2-2-9 (OR), C9-2 (PA), and FGSC2193 (GR), all of the *het-c*^{PA}-like chimeric constructs displayed a *het-c* specificity pattern identical to those of transformants containing the *N. crassa* canonical *het-c*^{PA} allele (Table 3), with the single exception of the *Ndi5923* construct.

Unlike the other *het-c*^{PA}-like chimeric constructs, the *Ndi5923* construct yielded incompatible transformants when introduced into C2-2-9 (OR), C9-2 (PA), and FGSC2193 (GR) strains (Table 3 and Fig. 4). The class 2-incompatible transformants in all three recipient strains displayed a similar phenotype, characterized by an LGR of approximately 1.3 to 1.8 cm per day, abnormal swollen hyphal morphology, and dead hyphal compartments (Fig. 4). The *Ndi5923* allele has a *het-c*^{PA}-like insertion that is 12 to 18 bp (four to six amino acids)

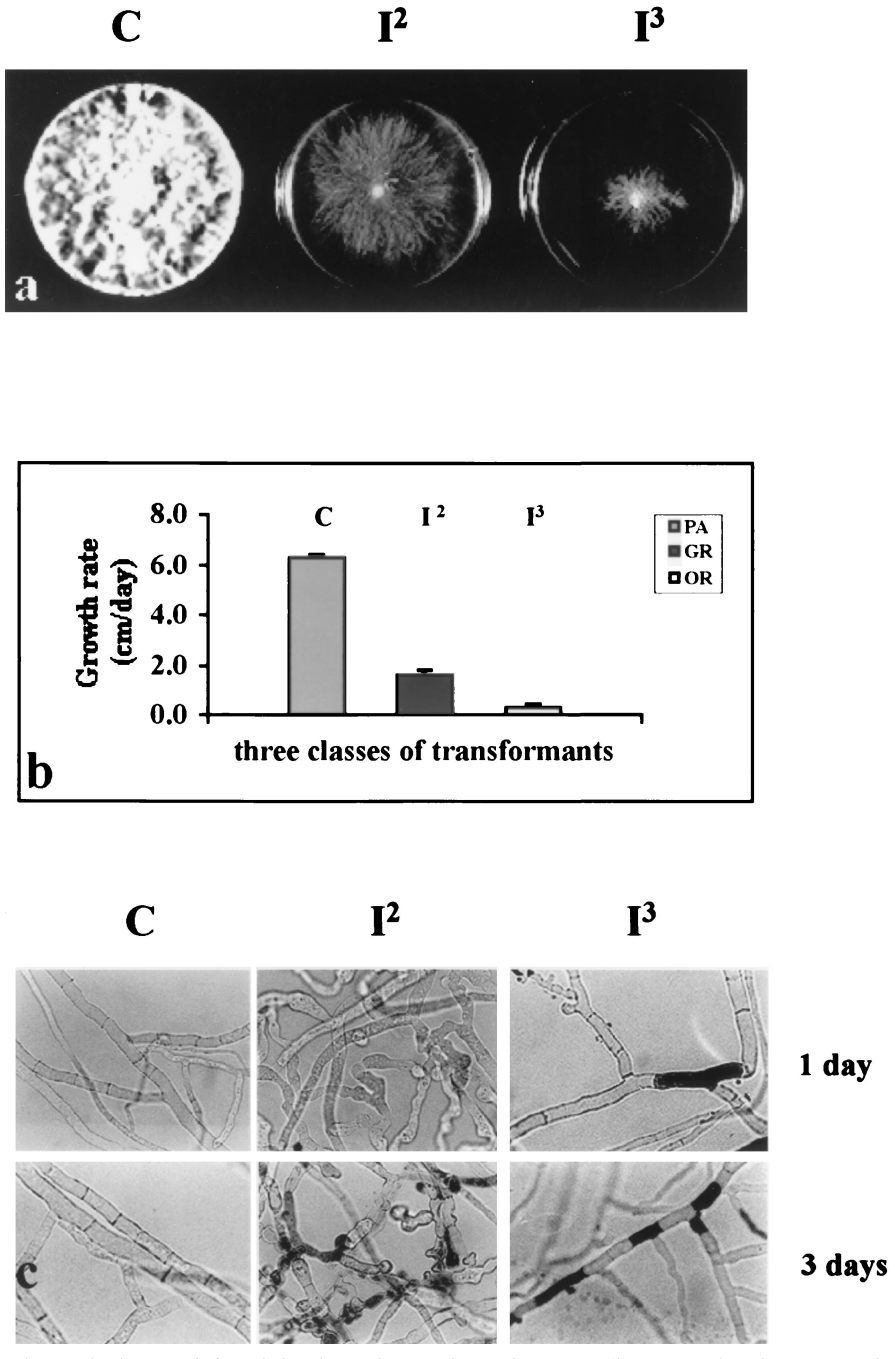

FIG. 3. Phenotypic and growth characteristics of the three classes of transformants. Shown are the phenotypes of compatible (class 1 [C]) transformants, such as the introduction of a *het-c*^{PA} allele into C9-2 (PA) spheroplasts, incompatible class 2 (I^2) transformants, such as introduction of a *het-c*^{GR} allele into C9-2 (PA) spheroplasts, and class 3 (I^3) transformants, such as introduction of a *het-c*^{OR} allele into C9-2 (PA) spheroplasts. The genotypes of strains are given in Table 1. (a) Representative colony growth of the three classes of transformants. All are shown after 3 days of incubation at 30°C on solid VM (see Materials and Methods). (b) Average growth rate (centimeters/day) of three classes of transformants represented by the introduction of *het-c*^{PA} (C), *het-c*^{GR} (I^2), and *het-c*^{OR} (I^3) alleles into a C9-2 (PA) strain (see above). (c) HCD in the three classes of transformants represented by the introduction of *het-c*^{PA} (C), *het-c*^{GR} (I^2), and *het-c*^{OR} (I^3) alleles into a C9-2 (PA) strain (see above) after 1 and 3 days of vegetative growth. Evan's blue (16) was used to stain dead hyphal compartments as described in Materials and Methods. Magnification, ×61.

shorter than other *het-c*^{PA}-like alleles (depending on the reference allele), in addition to amino acid differences.

The transformation results using the chimeric alleles showed that grouping of alleles by indel motif was a good predictor of *het-c* specificity. We did not observe *het-c* specificity differences among the 12 different clades within each *het-c* specificity group that were observed by phylogenetic analysis (48). However, we did identify a chimeric allele, *Ndi5923*, which showed

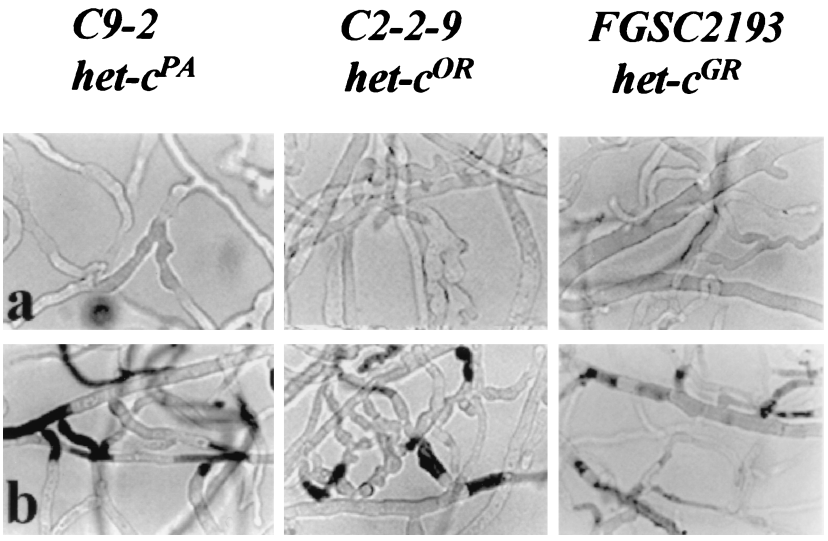

FIG. 4. Phenotypes of the class-2-incompatible transformants caused by the introduction of chimeric allele *Ndi5923* into C9-2 (PA), C2-2-9 (OR), and FGSC2193 (GR) after 1 day of growth (a) and 3 days of growth (b). Hyphae were treated with the vital dye Evan's blue (16), which stains dead hyphal compartments (25). Magnification, ×68.

length and amino acid differences within the *het-c*^{PA}-like insertion and conferred a novel *het-c* specificity.

Artificially constructed amino acid variations in regions I and II do not affect *het-c* specificity. The transformation results using the chimeric alleles showed that naturally occurring amino acid variations in regions I and II (Fig. 2) did not affect *het-c* allelic specificity. However, some variations in amino acid composition in region II were specific for a *het-c* type, and therefore, the role of amino acid variability on *het-c* specificity could not be completely addressed by our chimeric constructs using naturally occurring *het-c* alleles. We therefore constructed a number of artificial *het-c* alleles with novel combinations of region II and indel motifs (Fig. 2, region III).

The first type of allele (Fig. 5, *po3* and *po4*) had a mosaic combination of predicted amino acid sequence and indel motif between a *het-c*^{PA} and a *het-c*^{OR} allele. The *po3* allele has a *het-c*^{PA}-specific insertion in an otherwise *het-c*^{OR} allele. The introduction of *po3* into C9-2 (PA), C2-2-9 (OR), and FGSC2193 (GR) yielded an incompatibility-compatibility spectrum identical to that of a canonical *het-c*^{PA} allele (Table 4). The *po4* allele has a *het-c*^{OR} indel motif but is otherwise identical to a *het-c*^{PA} allele in predicted amino acid sequence (Fig. 5). When introduced into C9-2 (PA), C2-2-9 (OR), and FGSC2193 (GR), *po4* displayed a *het-c* specificity identical to that of a canonical *het-c*^{OR} allele (Table 4).

The second type of construct (*od1* and *go1*) contains mosaic combinations between region II and the indel motif between a *het-c*^{GR} and a *het-c*^{OR} allele (Fig. 5). Construct *od1* was generated from *het-c*^{OR} by a 15-bp deletion of the *het-c*^{OR}-specific insertion sequence (encoding RNDTR), thus creating an allele that is HET-C^{OR} for amino acid sequences but *het-c*^{GR}-like for the indel motif. The *go1* construct was generated from a *het-c*^{GR} allele by the addition of 15 bp encoding the *het-c*^{OR}-specific insertion sequence (RNDTR) and is thus *het-c*^{GR} for predicted amino acid sequence within the *het-c* variable region, with the exception of the *het-c*^{OR}-like insertion. The *od1* and

```
HET-C^PA   EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGH-SPVFPHVGTATKLKL-----ENRQFRRVRPGEGYDSGAKYAWPLVTGTFGG
PO3        EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGH-SPVFPHVGTDTRVTL-----ENRQFRRVRPGEGYDSGAKYAWPLVTGTFGG
Ndi5923    EALRCLGQALHTLEDFPAHSNYCELALIDIHQKEDRSKSRIFPHVGTRTKLKR-----ENGMFVPAPYEERD-----YVWPLVTGTFGG
Ndi5923m   EALRCLGQALHTLEDFPAHSNYCELALIDIHQKEDRSKSRIFPHVGTRTKLKR-----ENGMFVPAPYEERDDSGAKYVWPLVTGTFGG
PD1        EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGH-SPVFPHVGTATKLKL-----ENRQFRRVRPGEG-----KYAWPLVTGTFGG
PD2        EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGH-SPVFPHVGTATKLKL-----ENRQFRRVRPG-------KYAWPLVTGTFGG
PD3        EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGH-SPVFPHVGTATKLKL-----ENRQFRRV----------KYAWPLVTGTFGG
PD4        EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGH-SPVFPHVGTATKLKL-----ENR---------------KYAWPLVTGTFGG

HET-C^GR   EALRCLGQALHTLEDFPAHSNYCELVLIDIHEKETRSESRIFPHVGTATRITL-----NNG---------------KLVWPLVTGTFGG
OD1        EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGH-SPVFPHVGTDTRVTL-----NNG---------------KSVWPLVTGTFGG
DEL3       EALRCLGQALHTLEDFPAHSNYCELVLIDIHEKETRSESRIFPHVGTATRITL----------------------KLVWPLVTGTFGG

HET-C^OR   EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGH-SPVFPHVGTDTRVTLRNDTRNNG---------------KSVWPLVTGTFGG
PO4        EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGH-SPVFPHVGTATKLKLRNDTRNNG---------------KSVWPLVTGTFGG
GO1        EALRCLGQALHTLEDFPAHSNYCELVLIDIHEKETRSESRIFPHVGTATRITLRNDTRNNG---------------KLVWPLVTGTFGG
PO1        EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGH-SPVFPHVGTDTRVTLRNDTRNNGQFRRVRPGEGYDSGAKYAWPLVTGTFGG
PO2        EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGH-SPVFPHVGTDTRVTLRNDTRENRQFRRVRPGEGYDSGAKYAWPLVTGTFGG
```

FIG. 5. Predicted amino acid sequences in the variable domain of artificially constructed *het-c* alleles (for details on construction, see Materials and Methods).

TABLE 4. Phenotypes of transformants containing artificially constructed *het-c* alleles in recipient strains that differ in *het-c* specificity[a]

Allele	C9-2 (PA)	C2-2-9 (OR)	FGSC2193 (GR)
het-c^{PA}	C	I^2	C[b]
het-c^{GR}	I^2	I^2	C
het-c^{OR}	I^3	C	I^3
po3	C	I^2	C
po4	I^3	C	I^3
od1	I^2	I^2	C
go1	I^3	C	I^3
del3	C	C	C
Ndi5923	I^2	I^2	I
Ndi5923m	C	I^2	C
pd1	I^2	I^3	I^2
pd2	I^2	I^2	I^2
pd3	I^2	I^3	I^2
pd4	I^2	I^3	C
po1	I^3	I^2	I^3
po2	I^3	I^3	I^3

[a] C, fully compatible transformants with an LGR of 4 to 6.5 cm/day with <1% HCD; I^2, incompatible transformants with an LGR of 1 to 2.5 cm/day which show ~20 to 30% HCD; I^3, severely incompatible transformants that have an LGR of <1 cm/day and show 20 to 30% HCD.

[b] Transformation results (39) and analysis of partial diploid progeny (24) indicate that strain FGSC2193 has a recessive allele-specific suppressor for *het-c* incompatibility.

go1 chimeric constructs displayed a *het-c* specificity that was indistinguishable from the specificities of *het-c*^{GR} and *het-c*^{OR} alleles, respectively (Table 4). Thus, these data support what was observed in transformation experiments with the naturally occurring *het-c* chimeric constructs: allelic specificity is completely dependent upon the indel motif and amino acid variability outside of the indel motif does not contribute to *het-c* allelic specificity.

The indel motif of three amino acids that define a *het-c*^{GR} allele is essential for vegetative incompatibility. A *het-c*^{GR} allele has the smallest variable domain, which is characterized by a 9-bp indel motif (encoding NNG), which is also conserved among almost all *het-c*^{OR} alleles. In *het-c*^{PA}-like alleles, the predicted amino acid sequence of this 9-bp indel is variable but includes the predicted asparagine residue (N) (Fig. 2). The indel motif of HET-C^{GR} is predicted to form a short loop between the two conserved antiparallel β-strands. Removal of the codons for the NNG residues results in the predicted formation of a single β-sheet structure that includes both the conserved PHVGTRITL and WPLVTGTF regions (Fig. 2).

A *het-c*^{GR} allele that contained a 9-bp deletion that removed the codons for NNG was generated (Fig. 5, *del3*). Only compatible transformants were obtained when *del3* was introduced into C2-2-9 (OR), C9-2 (PA), and FGSC2193 (GR) strains by transformation. The *del3* transformants displayed growth rates comparable to those of other compatible transformants and did not display HCD. These data suggest that the loop structure between the antiparallel β-strands is important in the structural maintenance of the *het-c* specificity domain and that loss of this region produces a nonfunctional allele.

The novel specificity of *Ndi5923* could be converted to *het-c*^{PA} specificity by increasing the length of the *het-c*^{PA}-type insertion. The results obtained with the artificial and naturally occurring chimeric constructs indicated that amino acid variations in regions I and II do not materially affect *het-c* specificity; *het-c* specificity is dependent upon the indel motif in region III (Fig. 2). The *Ndi5923* allele displayed a novel *het-c* specificity and produced incompatible transformants in C9-2 (PA), C2-2-9 (OR), and FGSC2193 (GR) strains (Table 3). The *het-c*^{PA}-like insertion in *Ndi5923* shows both amino acid composition and length variations compared to those of other *het-c*^{PA}-like alleles.

To determine whether the predicted amino acid sequence of the *het-c*^{PA}-like insertion from *Ndi5923* was the most important factor in *het-c* specificity or whether the length of the insertion mattered, we constructed an artificial allele, *Ndi5923m*. Fifteen base pairs were inserted into the *Ndi5923* construct, thus making an allele with a *het-c*^{PA} indel insertion size identical to that of the *het-c*^{PA} allele from C9-2 (Fig. 5). The 15-bp addition encoded 5 aa that are observed in the C9-2 *het-c*^{PA} allele. Of these five codons, only the serine (S) and lysine (K) residues are conserved among HET-C^{PA}-like peptides (Fig. 2). When *Ndi5923m* was introduced into C2-2-9 (OR), C9-2 (PA), and FGSC2193 (GR) strains, it behaved identically to *het-c*^{PA} and yielded wild-type transformants in C9-2 and FGSC2193 and class-2-incompatible transformants in C2-2-9 (Table 4). These data showed that the novel *het-c* specificity displayed by *Ndi5923* could be converted to *het-c*^{PA} specificity by altering the size and amino acid composition of the *het-c*^{PA}-like insertion.

Variations in amino acid sequence in the indel motif affect *het-c* specificity and severity of incompatibility response. A set of artificial *het-c* constructs, *pd1*, *pd2*, *pd3*, and *pd4*, which contained variations in length of the *het-c*^{PA} insertion, was generated from the C9-2 *het-c*^{PA} allele (Fig. 5). The *pd1* construct has a deletion of 15 bp (5 aa) and thus has a *het-c*^{PA}-like insertion of identical size but with a predicted amino acid composition different from that of *Ndi5923*. The lysine (K) residue that is conserved among all naturally occurring HET-C polypeptides, with the exception of *Ndi5923* HET-C, is included in the *pd1* construct (Fig. 2 and 5). The *pd2* construct has a deletion of 21 bp (removing 7 aa) within the *het-c*^{PA}-like insertion. The *pd3* construct has a deletion of 30 bp (removing 10 aa) within the *het-c*^{PA}-specific insertion and thus has a size identical to that of a *het-c*^{OR} allele, although the 5-aa insertion is placed differently in respect to the NNG motif (Fig. 5). The *pd4* construct is missing the entire *het-c*^{PA}-specific insertion and thus has an indel motif that is identical in size to that of a *het-c*^{GR} allele, although it has the predicted amino acid composition of a *het-c*^{PA} allele (Fig. 5). In particular, instead of having the 9-bp indel motif (NNG) from *het-c*^{GR}, the *pd4* HET-C has ENR from *het-c*^{PA}.

The *pd4* construct displayed an identical *het-c* specificity to a canonical *het-c*^{GR} allele in transformation experiments, although the severity of the incompatibility response was affected (Table 4). The introduction of *pd4* into C2-2-9 (OR) yielded class 3, severely incompatible transformants, but the introduction of *het-c*^{GR} into C2-2-9 (OR) yielded class-2-incompatible transformants. Similar to *het-c*^{GR}, the introduction of *pd4* into C9-2 (PA) yielded class-2-incompatible transformants and compatible transformants in FGSC2193 (GR). These data show that although *het-c* specificity is dependent upon indel motif, amino acid differences within the indel motif can affect the severity of the incompatibility phenotype.

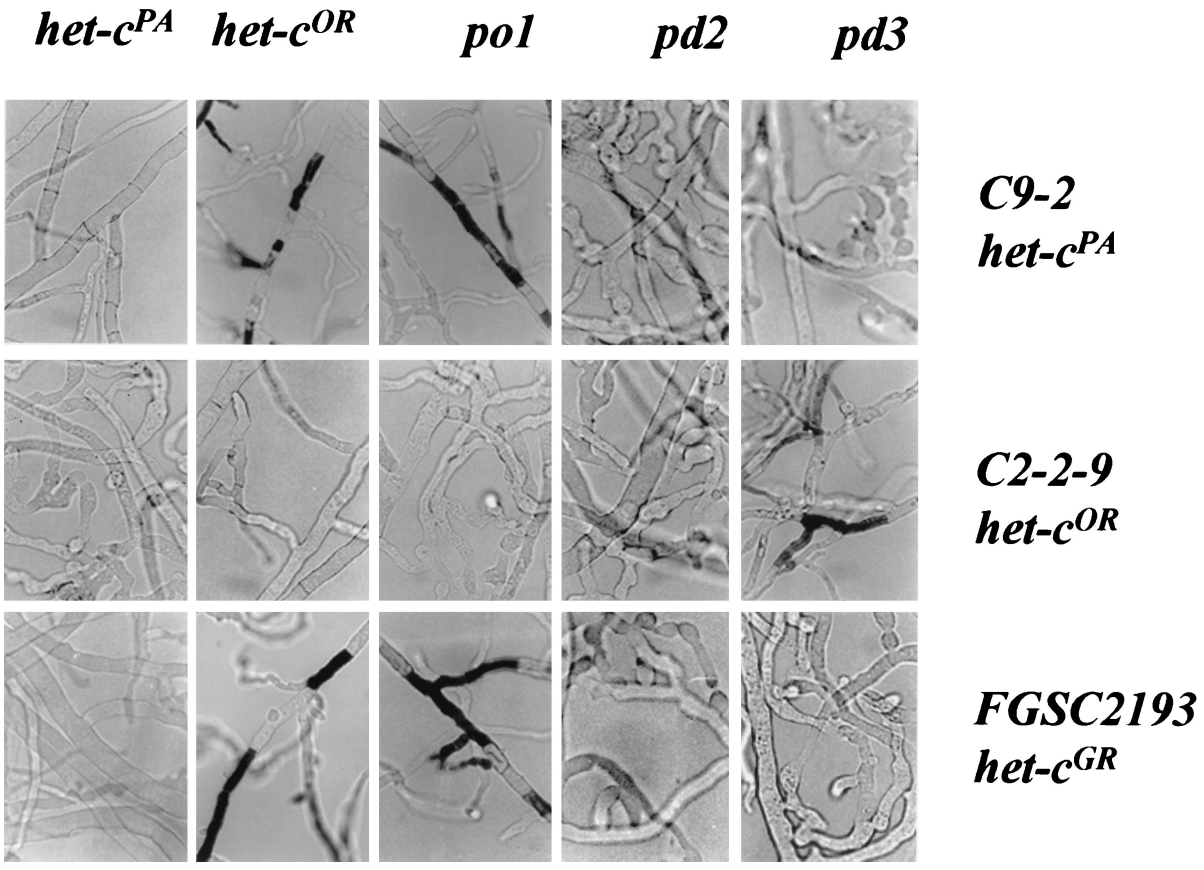

FIG. 6. Phenotypes of C9-2 (*het-c*^{PA}), C2-2-9 (*het-c*^{OR}), and FGSC2193 (*het-c*^{GR}) transformants containing *het-c*^{PA}, *het-c*^{OR}, *po1, pd2*, and *pd3* constructs observed after 1 day of growth. Hyphae were treated with the vital dye Evan's blue (16), which stains dead hyphal compartments. Magnification, ×64.

Similar to *Ndi5923*, incompatible transformants were obtained when *pd1*, *pd2*, and *pd3* were introduced into C9-2 (PA), C2-2-9 (OR), and FGSC2193 (GR) (Table 4 and Fig. 6). The *pd1* allele has a size identical to but an amino acid composition different from those of *Ndi5923* and produced class-2-incompatible transformants in C9-2 (PA) and FGSC2193 (GR) but produced severely incompatible transformants in C2-2-9 (OR). The *pd2* construct, which has an indel size that is different from those of all other alleles, yielded transformants that displayed a class-2-incompatible phenotype. The *pd3* construct has a variable domain that is identical in size to that of a *het-c*^{OR} allele (Fig. 7). Secondary-structure predictions of the predicted variable region from PD3 and HET-C^{OR} showed different profiles in the loop region between the antiparallel β-strands. Unlike a standard *het-c*^{OR} allele, the introduction of *pd3* into C2-2-9 (OR), C9-2 (PA), and FGSC2193 yielded a spectrum of incompatible transformants identical to that of *pd1* (Table 4). These data indicated that amino acid composition variation in the indel motif can confer novel *het-c* specificity, perhaps by altering spatial characteristics of the loop domain formed by the indel motif.

Novel *het-c* specificity can be generated by increasing length of indel motif. Our results indicated that variations in the pattern and length of the *het-c*^{PA}-like insertion could generate alleles with novel *het-c* specificity. To examine if increasing the length of the variable domain affected *het-c* specificity, two

```
A.

Ndi5923   EALRCLGQALHTLEDFPAHSNYCELALIDIHQKEDRSKSRIFPHVGTRTKLKRENGMFVPAPYEERDYVWPLVTGTFGG
PD1       EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGH-SPVFPHVGTATKLKLENRQFRRVRPGEGKYAWPLVTGTFGG
          ************************* ***         ** ****** **** **  *      *  * **********

B.

HET-C^OR  EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGHSPVFPHVGTDTRVTLRNDTRNNGKSVWPLVTGTFGG
PD3       EALRCLGQALHTLEDFPAHSNYCELVLIDMEERRGGHSPVFPHVGTATKLKLENRQFRRVKYAWPLVTGTFGG
          ******************************************** *   *          **********
```

FIG. 7. Amino acid comparison between predicted products of alleles that confer identical *het-c* specificity, Ndi5923 and PD1 (A), and predicted products of alleles that confer different *het-c* specificity, HET-C^{OR} and PD3 (B). *, amino acid identity.

artificial alleles were constructed, *po1* and *po2*, which contain both of the *het-c*^{OR}-specific and *het-c*^{PA}-specific insertions. These two constructs are 15 bp longer than a typical *het-c*^{PA} allele. The *po1* and *po2* constructs are identical, except that *po1* has the predicted NNG sequences from *het-c*^{OR} while *po2* has the predicted ENR sequences from *het-c*^{PA} (Fig. 5).

The introduction of both *po1* and *po2* into C9-2, C2-2-1, and FGSC2193 produced transformants that displayed incompatible phenotypes (Table 4 and Fig. 6). Both C9-2 (PA) and FGSC2193 (GR) strains that contained either *po1* or *po2* constructs displayed severely incompatible phenotypes (LGR of <0.5 cm/day); dead hyphal compartments could be observed after 24 h of growth. The C2-2-9 (OR) transformants that contained either *po1* or *po2* constructs showed a class-2-incompatible phenotype. Secondary-structure predictions showed that the indel region of PO1 and PO2 contained two loops with an additional short region predicted to form a β-sheet between the two conserved antiparallel β strands. The fact that *po1* and *po2* constructs showed an identical novel *het-c* specificity and similar spectrum of phenotypes suggested that they might also confer an identical *het-c* specificity.

Alleles conferring novel *het-c* specificity are not self-incompatible. The *po1*, *po2*, *pd1*, *pd2*, *pd3*, and *Ndi5923* constructs displayed a novel *het-c* specificity and produced incompatible transformants when introduced into C9-2 (PA), C2-2-9 (OR), and FGSC2193 (GR) strains. It is possible that the incompatible phenotypes that we observed were the result of self-incompatibility of these alleles rather than of incompatibility triggered by an interaction with the resident *het-c* allele. To determine whether *po1*, *po2*, *pd1*, *pd2*, *pd3*, and *Ndi5923* conferred self-incompatibility, each construct was individually introduced into a *het-c* deletion strain, CJ44 (Table 1). In all cases, only compatible, wild-type transformants were obtained (Table 5). In particular, *po1* and *po2*, which have both *het-c*^{OR}-like and *het-c*^{PA}-like insertions, were not self-incompatible. These data indicate that *po1*, *po2*, *pd1*, *pd2*, *pd3*, and *pd4* confer a novel *het-c* specificity in *N. crassa* that is different from *het-c*^{PA}, *het-c*^{OR}, and *het-c*^{GR} allelic specificities.

Alleles with variable specificity domains confer four novel and different *het-c* specificities. The introduction of *po1*, *po2*, *pd1*, *pd2*, *pd3*, and *Ndi5923* into C2-2-9 (OR), C9-2 (PA), and FGSC2193 (GR) produced transformants that displayed all of the hallmarks of *het-c* incompatibility (growth inhibition, suppression of conidiation, and HCD), indicating that these constructs conferred a novel *het-c* specificity. However, it was unclear whether these constructs each conferred an identical *het-c* specificity or defined new and different *het-c* allelic specificities. To distinguish these two possibilities, *po1*, *po2*, *pd1*, *pd2*, *pd3*, and *Ndi5923* were cotransformed in pairwise combinations into CJ44 (Table 5); the presence of alternative alleles was coselected by growth on hygromycin media which lacked pantothenic acid (see Materials and Methods). As expected, the introduction by cotransformation of constructs of identical type yielded only compatible transformants (e.g., *het-c*^{OR} plus *het-c*^{OR}, *pd1* plus *pd1*, or *po1* plus *po1*) (Table 5), while the cotransformation of *het-c*^{PA} plus *het-c*^{OR} and *het-c*^{GR} plus *het-c*^{OR} constructs produced severely incompatible transformants.

Only two pairwise combinations of novel alleles yielded compatible transformants. The *pd1* construct has an indel motif of the same size as that of *Ndi5923*, although only four predicted amino acid positions are conserved (Fig. 7, E, N, F, and E). However, cotransformation of *pd1* and *Ndi5923* into CJ44 yielded only compatible transformants, indicating that these two constructs conferred an identical *het-c* specificity (Table 5). Similarly, *po1* and *po2* have an indel motif of identical size but differ in predicted amino acid sequence within the NNG or ENR motif (Fig. 5). The cotransformation of *po1* and *po2* alleles into CJ44 also yielded compatible transformants, indicating that these two alleles have identical *het-c* specificity (Table 5).

TABLE 5. Phenotypes and growth rates of CJ44 transformants containing different combinations of *het-c* allele pairs

Allele pair	Phenotype[a]	Growth rate (cm/day) (mean ± SD)	HCD[b]
pCB1004 + pOKE	C	4.6 ± 0.20	−
het-c^{OR} + *het-c*^{OR}	C	4.7 ± 0.20	−
het-c^{PA} + *het-c*^{PA}	C	4.8 ± 0.20	−
het-c^{OR} + *het-c*^{PA}	I	0.3 ± 0.10	+++
het-c^{OR} + *het-c*^{GR}	I	0.4 ± 0.10	+++
Ndi5923 + Ndi5923	C	4.6 ± 0.17	−
pd1 + pd1	C	4.3 ± 0.30	−
pd2 + pd2	C	5.4 ± 0.60	−
pd3 + pd3	C	4.4 ± 0.30	−
po1 + po1	C	4.7 ± 0.57	−
po2 + po2	C	4.8 ± 0.46	−
pd1 + Ndi5923	C	4.5 ± 0.35	−
pd1 + pd2	I	0.5 ± 0.12	+++
pd1 + pd3	I	0.4 ± 0.05	+++
pd1 + po1	I	0.3 ± 0.10	+++
pd2 + pd3	I	0.6 ± 0.10	+++
pd2 + po1	I	0.4 ± 0.06	+++
pd3 + po1	I	0.4 ± 0.10	+++
pd3 + het-c^{OR}	I	0.6 ± 0.15	+++
po1 + po2	C	4.7 ± 0.30	−

[a] Transformants were compatible (C) or incompatible (I).
[b] −, <1% HCD was observed in the growing colony. +++, 20 to 30% HCD was observed in the colony after 1 day of growth.

Pairwise cotransformation of all other allele combinations resulted in severely incompatible transformants (Table 5). These transformants displayed a typical *het-c* incompatibility phenotype with an LGR of approximately 0.5 cm/day and showed 20 to 30% HCD throughout the colony. In particular, the introduction of *pd3* and *het-c*^{OR} into CJ44 yielded incompatible transformants, even though they had specificity domains with identical size (Fig. 7). Thus, four additional novel *het-c* specificities were identified in this study. One, *Ndi5923*, was identified from a survey of naturally occurring alleles. The other three were identified from artificially constructed alleles that contained length variations in the indel motif (*pd2* and *po1*/*po2*) or variations in amino acid sequence of the indel motif (*pd3*).

DISCUSSION

***het-c* specificity is dependent upon indel motif.** The *het-c* variable domain characterized from species and genera related to *N. crassa* showed that the alleles could be grouped into one of the three *het-c* specificities identified in *N. crassa*, based on the indel motif; the indel motif pattern exhibits transspecies polymorphisms (reference 48 and Fig. 2). Phylogenetic analysis (either including or excluding the indel motif) supported this

grouping but also split each group into additional clades. Transspecies polymorphisms were also apparent in these clades; clade 5 contains a *het-c* allele from *Neurospora pannonica*, *Sordaria sclerogenia*, and a *Gelasinospora* sp. isolate. These data suggest that multiple specificities might occur at *het-c*, in addition to *het-c*^{OR}, *het-c*^{PA}, and *het-c*^{GR}-types. However, the results of this study indicate that naturally occurring variations in amino acid sequences in the variable domain do not affect *het-c* specificity; *het-c* specificity is dependent upon the indel motif. The neutral amino acid variation surrounding the indel motif has presumably been retained in populations because of selection for alleles conferring alternative *het-c* specificity, which is dependent upon the indel motif. Transspecies polymorphism associated with two different indel motifs has also been reported in allelic lineages in the MHC Aβ1 locus of mice and rats, which have been preserved for over 10 million years (15). Interestingly, intragenic recombination surrounding the indel motif of the Aβ1 locus was also observed, thus creating hybrid genes. Occasional recombination within the *het-c* variable region could also link different neutral polymorphisms (such as those found in regions I and II) with identical indel motifs, resulting in apparent transspecies polymorphisms within an indel type. Phylogenetic analysis treats insertions and deletions as single events, and therefore, mutations that result in amino acid variability outside of the indel motif significantly affect tree topology (48). In this study, we determined that chimeric constructs containing variable domains from naturally occurring *het-c* alleles fell into one of three *het-c* allelic specificities identified in *N. crassa*. The only exception was a naturally occurring allele from *Neurospora discreta*. If this alternative specificity is also under balancing selection, it should increase in frequency within the *N. discreta* population, similar to what has been hypothesized for both *S* alleles and polymorphic loci in the MHC (27).

DNA sequences of various *het* loci from *N. crassa* and *P. anserina* show that alleles conferring alternative specificities are polymorphic. For example, the alternative *het-s* polypeptides of *P. anserina* differ by 12 aa substitutions, although only a single amino acid change is sufficient to switch allelic specificity (44). Similarly, 16 polymorphic positions were identified in predicted *P. anserina het-c* polypeptides (not related to *N. crassa het-c*) (36). As with *het-s*, allelic specificity at *P. anserina het-c* was dependent upon a single amino acid difference and constructs conferring novel *het-c* allelic specificities were generated by chimeric allele construction. In *N. crassa*, the predicted *het-6* polypeptides show only 68% amino acid identity, with polymorphic positions scattered throughout the open reading frame (42). Although alternative alleles at *het* loci are polymorphic, it is unclear whether selection mechanisms are maintaining polymorphisms at *het* loci other than the *N. crassa het-c* locus. The introduction of the *N. crassa het-c*^{OR}, *het-c*^{PA}, and *het-c*^{GR} alleles into *P. anserina* resulted in growth inhibition and HCD (38). These data indicate that the mechanism of *het-c*-mediated vegetative incompatibility is well conserved among filamentous fungi, but whether it occurs in a particular species may be dependent on the presence of polymorphisms within the indel motif.

Secondary-structure predictions and *het-c* specificity. The indel motif is predicted to form a coiled-loop structure between conserved antiparallel β-strands. The allele-specific indel motif may form a protruding loop that mediates *het-c* allele specificity, perhaps by protein-protein interactions (between alternative HET-C proteins or with other proteins). The predicted secondary structure of HET-C is reminiscent of immunoglobulin molecules; crystallographic studies with immunoglobulin molecules show that the antigen interface consists of three hypervariable loops known as the complementarity determining region (CDR) which are formed on nine antiparallel β-strands of the variable domain (9). Both length variations and amino acid substitutions occur in the CDRs and are correlated with antigen binding specificity (5). The large number of sequence variations in the loop region of antibodies shows that the structural framework of the variable domain is fairly insensitive to changes in amino acid composition and length of the CDR loop regions. In this study, novel *het-c* specificities were generated either by alterations in amino acid and/or length of the indel motif. Only one construct, *del3*, which removed the loop between the conserved antiparallel β-strands, destroyed *het-c* function. Although many of the alleles that confer alternative *het-c* specificities give significantly different secondary-structure profiles in the indel region (*het-c*^{OR}, *het-c*^{PA}, *het-c*^{GR} and *po1*/*po2*), it is not obvious from such predictions what differences are crucial for conferring alternative *het-c* specificities. Future structural studies using the predicted products of the seven different *het-c* specificities identified in this study will provide useful tools to address this issue.

Relationship between allelic specificity and recognition. An allelic specificity region functioning in heteromeric complex formation has been described for several fungal mating systems. In *U. maydis*, the *b* mating-type genes (*bE* and *bW*) have variable and constant regions (28). Analysis of chimeric *bE* and *bW* alleles showed that single-amino-acid alterations in the N-terminal variable region were sufficient to generate novel mating specificities. It was hypothesized that these amino acid differences affected the capacity of different bE and bW proteins to heterodimerize (26). In *P. anserina*, alternative *het-s* polypeptides (HET-S and HET-s) have been shown to form both heterodimers and homodimers via yeast two-hybrid experiments (8). The results in this study also support the possibility that vegetative incompatibility is mediated by HET-C heteromeric complex formation. In an immunoprecipitation study, alternative HET-C proteins formed a heteromeric complex during vegetative incompatibility (47; G. Iyer and N. L. Glass, unpublished data). From the results generated in this study, we would predict that the amino acid composition and length differences in the indel motif in the chimeric constructs that displayed novel specificity (*Ndi5923*/*pd1*, *pd2*, *pd3*, and *po1*/*po2*) may facilitate heterocomplex formation between alternative HET-C proteins. The specificity domain may directly mediate protein-protein interactions or, alternatively, affect the conformation of a different region of HET-C that mediates physical interactions between alternative HET-C polypeptides.

Models for *het-c* vegetative incompatibility. Vegetative incompatibility reactions appear to have conserved features in filamentous fungi. In *N. crassa* and *P. anserina*, the vegetative incompatibility response involves common stages of hyphal compartmentation, vacuolization, and death (3, 18, 25). Forced heterokaryons or transformants containing alternative *het-c* alleles show growth inhibition, growth arrest, suppression of conidiation, and HCD. The morphological phenotype and se-

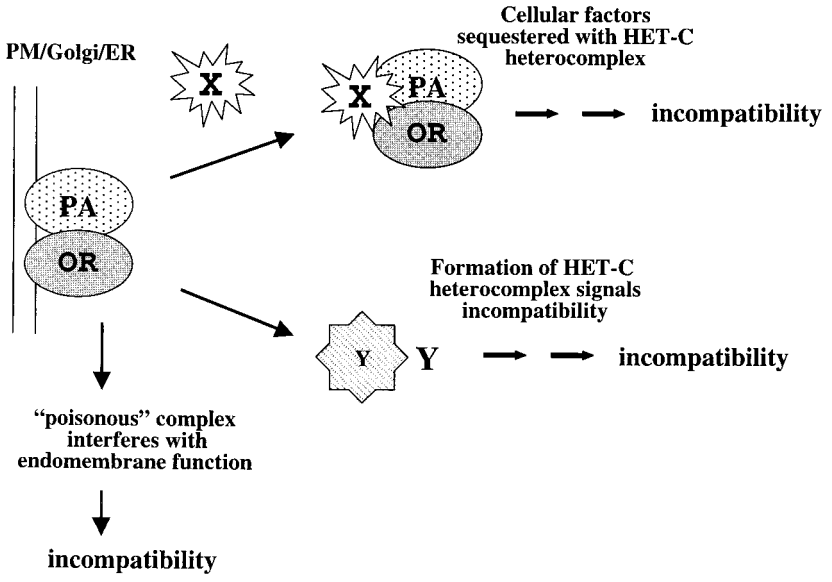

FIG. 8. Alternative models for *het-c*-mediated vegetative incompatibility. The model depicts the heteromeric complex formation between the products of incompatible *het-c* alleles, the formation of which is dependent upon variations in the indel motif in the variable domain. The HET-C heteromeric complex may trigger vegetative incompatibility by a poison effect on function of the endoplasmic reticulum (ER) and golgi or plasma membrane (PM) or by a dominant-negative effect on cell growth by an interaction with a modulator (X) or by regulating the activity or expression of a separate component (Y) that triggers the activation of the downstream effectors of vegetative incompatibility. OR, HET-C^{OR}; PA, HET-C^{PA}.

verity of the incompatibility phenotype can be affected by the genetic background of recipient strains (reference 39 and this study). In addition, it is possible that variations in the indel motif affect the thermodynamics of heterocomplex formation and thus affect the phenotypic output of vegetative incompatibility. Further genetic experiments and experiments that assess the affinity of alternative HET-C proteins to form a heterocomplex will be required to differentiate these two possibilities.

The HET-C protein is predicted to enter the endomembrane system and to ultimately reside within either the golgi or the plasma membrane. The presence of alternative HET-C proteins within a common cytoplasm results in growth inhibition and HCD, presumably by heterocomplex formation (Fig. 8). Chimeric *het-c* allele construction and secondary-structure prediction in this study suggested that the specificity domain of *het-c* may be involved in the mechanism of stable HET-C heteromeric complex formation. An interaction of alternative HET-C proteins may mediate vegetative incompatibility directly, perhaps by interfering with the normal functioning of HET-C within the cell, although mutational analysis has shown that *het-c* is not an essential gene; *het-c* mutants are indistinguishable from the wild type in morphology (40). Alternatively, heterocomplex formation between HET-C proteins may result in the formation of a "poison" complex that affects normal cellular function of either the plasma membrane or endomembrane system. In *P. anserina*, a poison heteromeric complex model, in which heteromeric complexes between the products of incompatible genes are lethal to the cell, has been proposed for mediating vegetative incompatibility (2). The formation of such a complex may result in general growth inhibition and morphological changes, until a threshold of HET-C heterocomplex is formed and HCD is triggered. Alternatively, HET-C heterocomplex formation may trigger vegetative incompatibility via signaling mechanisms that result in growth inhibition and HCD (Fig. 8). The results of this study have provided tools and information on the molecular basis of allelic specificity that will be extremely useful in deciphering the structural mechanism of recognition, which ultimately leads to the phenotypic manifestations of vegetative incompatibility in filamentous fungi.

ACKNOWLEDGMENTS

We thank S. Sarkar and G. Iyer and the members of the N. L. Glass laboratory for critical reading of the manuscript. We also thank the members of J. Wu's graduate committee, Michel Roberge, Jim Kronstad, and Sally Otto, for their advice and suggestions.

We gratefully acknowledge financial support from the Canadian Natural Sciences and Engineering Research Council (NSERC) and a National Institutes of Health GM60468 grant to N.L.G.

REFERENCES

1. **Asante-Owusu, R. N., A. H. Banham, H. U. Böhnert, E. J. Mellor, and L. A. Casselton.** 1996. Heterodimerization between two classes of homeodomain proteins in the mushroom *Coprinus cinereus* brings together potential DNA-binding and activation domains. Gene **172:**25–31.
2. **Begueret, J., B. Turcq, and C. Clave.** 1994. Vegetative incompatibility in filamentous fungi—*het* genes begin to talk. Trends Genet. **10:**441–446.
3. **Beisson-Schecroun, J.** 1962. Incompatibilité cellulaire et interactions nucleocytoplasmiques dans les phénomènes de barrage chez *Podospora anserina*. Ann. Genet. **4:**3–50.
4. **Berges, T., and C. Barreau.** 1989. Heat shock at an elevated temperature improves transformation efficiency of protoplasts from *Podospora anserina*. J. Gen. Microbiol. **135:**601–604.
5. **Branden, C., and J. Tooze.** 1991. Introduction to protein structure, p. 179–199. Garland Publishing, Inc., New York, N.Y.
6. **Carroll, A. M., J. A. Sweigard, and B. Valent.** 1994. Improved vectors for selecting resistance to hygromycin. Fungal Genet. Newsl. **41:**22.
7. **Caten, C. E.** 1972. Vegetative incompatibility and cytoplasmic infection in fungi. J. Gen. Microbiol. **72:**221–229.
8. **Coustou, V., C. Deleu, S. Saupe, and J. Bégueret.** 1997. The protein product of the *het-s* heterokaryon incompatibility gene of the fungus *Podospora*

anserina behaves as a prion analog. Proc. Natl. Acad. Sci. USA **94:**9773–9778.
9. **Davies, D. R., E. A. Padlan, and S. Sheriff.** 1990. Antibody-antigen complexes. Annu. Rev. Biochem. **59:**439–473.
10. **Davis, R. H., and F. J. De Serres.** 1970. Genetic and microbial research techniques for *Neurospora crassa*. Methods Enzymol. **17A:**79–143.
11. **Debets, A. J. M., and A. J. F. Griffiths.** 1998. Polymorphism of *het*-genes prevents resource plundering in *Neurospora crassa*. Mycol. Res. **102:**1343–1349.
12. **Debets, F., X. Yang, and A. J. F. Griffiths.** 1994. Vegetative incompatibility in *Neurospora*—its effect on horizontal transfer of mitochondrial plasmids and senescence in natural populations. Curr. Genet. **26:**113–119.
13. **Deléage, G., and B. Roux.** 1987. An algorithm for protein secondary structure prediction based on class prediction. Protein Eng. **1:**289–294.
14. **Deleu, C., C. Clavé, and J. Bégueret.** 1993. A single amino acid difference is sufficient to elicit vegetative incompatibility in the fungus *Podospora anserina*. Genetics **135:**45–52.
15. **Figueroa, F., E. Gunther, and J. Klein.** 1988. MHC polymorphisms predating speciation. Nature **335:**265–271.
16. **Gaff, D. F., and O. Okong'O-Ogola.** 1971. The use of non-permeating pigments for testing the survival of cells. J. Exp. Bot. **22:**756–758.
17. **Garnjobst, L.** 1953. Genetic control of heterocaryosis in *Neurospora crassa*. Am. J. Bot. **40:**607–614.
18. **Garnjobst, L., and J. F. Wilson.** 1956. Heterocaryosis and protoplasmic incompatibility in *Neurospora crassa*. Proc. Natl. Acad. Sci. USA **42:**613–618.
19. **Geourjon, C., and G. Deléage.** 1994. SOPM: a self-optimized method for protein secondary structure prediction. Protein Eng. **7:**157–164.
20. **Geourjon, C., and G. Deléage.** 1995. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput. Appl. Biosci. **11:**681–684.
21. **Gibrat, J. F., J. Garnier, and B. Robson.** 1987. Further developments of protein secondary structure prediction using information theory. New parameters and consideration of residue pairs. J. Mol. Biol. **198:**425–443.
22. **Glass, N. L., D. J. Jacobson, and K. T. Shiu.** 2000. The genetics of hyphal fusion and vegetative incompatibility in filamentous ascomycetes. Annu. Rev. Genet. **34:**165–186.
23. **Horton, R. M., H. D. Hunt, S. N. Ho, J. K. Pullen, and L. R. Pease.** 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene **77:**61–68.
24. **Howlett, B., J. F. Leslie, and D. D. Perkins.** 1993. Putative multiple alleles at the vegetative (heterokaryon) incompatibility loci *het-c* and *het-8* in *Neurospora crassa*. Fungal Genet. Newsl. **40:**40–42.
25. **Jacobson, D. J., K. Beurkens, and K. L. Klomparens.** 1998. Microscopic and ultrastructural examination of vegetative incompatibility in partial diploids heterozygous at *het* loci in *Neurospora crassa*. Fungal Genet. Biol. **23:**45–56.
26. **Kämper, J., M. Reichmann, T. Romeis, M. Bölker, and R. Kahmann.** 1995. Multiallelic recognition: nonself-dependent dimerization of the bE and bW homeodomain proteins in *Ustilago maydis*. Cell **81:**73–83.
27. **Klein, J., A. Sato, S. Nagl, and C. O'hUigin.** 1998. Molecular trans-species polymorphism. Annu. Rev. Ecol. Syst. **29:**1–21.
28. **Kronstad, J. W., and S. A. Leong.** 1990. The b mating-type locus of *Ustilago maydis* contains variable and constant regions. Genes Dev. **4:**1384–1395.
29. **Leslie, J. F.** 1993. Fungal vegetative compatibility. Annu. Rev. Phytopathol. **31:**127–150.
30. **Levin, J. M., B. Robson, and J. Garnier.** 1986. An algorithm for secondary structure determination in proteins based on sequence similarity. FEBS Lett. **205:**303–308.
31. **Mylyk, O. M.** 1975. Heterokaryon incompatibility genes in *Neurospora crassa* detected using duplication-producing chromosome rearrangements. Genetics **80:**107–124.
32. **Perkins, D. D.** 1997. Chromosome rearrangements in *Neurospora* and other filamentous fungi. Adv. Genet. **36:**239–398.
33. **Perkins, D. D.** 1982. Main features of vegetative incompatibility in *Neurospora crassa*. Fungal Genet. Newsl. **35:**44–46.
34. **Perkins, D. D.** 1975. The use of duplication-generating rearrangements for studying heterokaryon incompatibility genes in *Neurospora*. Genetics **80:**87–105.
35. **Perkins, D. D., A. Radford, and M. S. Sachs.** 2000. The *Neurospora* compendium: chromosomal loci. Academic Press, San Diego, Calif.
36. **Saupe, S., B. Turcq, and J. Bégueret.** 1995. Sequence diversity and unusual variability at the *het-c* locus involved in vegetative incompatibility in the fungus *Podospora anserina*. Curr. Genet. **27:**466–471.
37. **Saupe, S. J.** 2000. Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes. Microbiol. Mol. Biol. Rev. **64:**489–502.
38. **Saupe, S. J., C. Clavé, M. Sabourin, and J. Bégueret.** 2000. Characterization of *hch*, the *Podospora anserina* homolog of the *het-c* heterokaryon incompatibility gene of *Neurospora crassa*. Curr. Genet. **38:**39–47.
39. **Saupe, S. J., and N. L. Glass.** 1997. Allelic specificity at the *het-c* heterokaryon incompatibility locus of *Neurospora crassa* is determined by a highly variable domain. Genetics **146:**1299–1309.
40. **Saupe, S. J., G. A. Kuldau, M. L. Smith, and N. L. Glass.** 1996. The product of the *het-C* heterokaryon incompatibility gene of *Neurospora crassa* has characteristics of a glycine-rich cell wall protein. Genetics **143:**1589–1600.
41. **Schweizer, M., M. E. Case, C. C. Dykstra, N. H. Giles, and S. R. Kushner.** 1981. Identification and characterization of recombinant plasmids carrying the complete *qa-2*+ cluster from *Neurospora crassa* including the *qa-1*+ regulatory gene. Proc. Natl. Acad. Sci. USA **78:**5086–5090.
42. **Smith, M. L., O. C. Micali, S. P. Hubbard, N. Mir-Rashed, D. J. Jacobson, and N. L. Glass.** 2000. Vegetative incompatibility in the *het-6* region of *Neurospora crassa* is mediated by two linked genes. Genetics **155:**1095–1104.
43. **Smith, M. L., C. J. Yang, R. L. Metzenberg, and N. L. Glass.** 1996. Escape from *het-6* incompatibility in *Neurospora crassa* partial diploids involves preferential deletion within the ectopic segment. Genetics **144:**523–531.
44. **Turcq, B., C. Deleu, M. Denayrolles, and J. Begueret.** 1991. Two allelic genes responsible for vegetative incompatibility in the fungus *Podospora anserina* are not essential for cell viability. Mol. Gen. Genet. **228:**265–269.
45. **Vallette, F., E. Mege, A. Reiss, and M. Adesnik.** 1989. Construction of mutant and chimeric genes using polymerase chain reaction. Nucleic Acids Res. **17:**723–733.
46. **Vogel, H. J.** 1964. Distribution of lysine pathways among fungi: evolutionary implications. Am. Nat. **98:**435–446.
47. **Wu, J.** 2000. Ph.D. thesis. Non-self recognition in filamentous fungi—the *het-c* mediated vegetative incompatibility in *Neurospora crassa*. University of British Columbia, Vancouver, Canada.
48. **Wu, J., S. J. Saupe, and N. L. Glass.** 1998. Evidence for balancing selection operating at the *het-c* heterokaryon incompatibility locus in a group of filamentous fungi. Proc. Natl. Acad. Sci. USA **95:**12398–12403.
49. **Yee, A. R., and J. W. Kronstad.** 1993. Construction of chimeric alleles with altered specificity at the *b* incompatibility locus of *Ustilago maydis*. Proc. Natl. Acad. Sci. USA **90:**664–668.
50. **Yee, A. R., and J. W. Kronstad.** 1998. Dual sets of chimeric alleles identify specificity sequences for the *bE* and *bW* mating and pathogenicity genes of *Ustilago maydis*. Mol. Cell. Biol. **18:**221–232.