The Plant Cell, Vol. 7, 773–783, June 1995

An N-Terminal Dimerization Domain Permits Homeodomain Proteins To Choose Compatible Partners and Initiate Sexual Development in the Mushroom *Coprinus cinereus*

Alison H. Banham,[a] Rachel N. Asante-Owusu,[a] Berthold Göttgens,[a] Sara A. J. Thompson,[a] Crawford S. Kingsnorth,[a] E. Jane C. Mellor,[b] and Lorna A. Casselton [a,1]

[a] Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, United Kingdom
[b] Microbiology Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, United Kingdom

The *A* mating-type locus of the mushroom *Coprinus cinereus* contains three or more paralogous pairs of genes encoding two families of homeodomain proteins (HD1 and HD2). A successful mating brings together different allelic forms of at least one gene, and this is sufficient to trigger initial steps in sexual development. Previous studies have suggested that development is regulated by heterodimerization between HD1 and HD2 proteins. In this report, we describe 5′ gene deletions and 5′ end exchanges showing that the N-terminal regions of the proteins are essential for choosing a compatible partner but not for regulating gene transcription. Using an in vitro glutathione *S*-transferase association assay, we demonstrated heterodimerization between HD1 and HD2 proteins and found that heterodimerization only occurs between compatible protein combinations. The N-terminal regions of the proteins were sufficient to mediate dimerization, and N-terminal swaps resulted in a predicted change in dimerization specificity. By analyzing the N-terminal amino acid sequences of HD1 proteins, we identified two potential coiled-coil motifs whose relative positions vary in paralogous proteins but are both required for in vivo function.

INTRODUCTION

Mating in the mushroom *Coprinus cinereus* commits cells to early events in sexual development; a sterile monokaryon with uninucleate cells is converted to a fertile binucleate-celled dikaryon that differentiates the fruit bodies. Somatic cell fusion initiates mating, but for mating to be successful, the partners must have different alleles of genes at the unlinked *A* and *B* mating-type loci (Casselton, 1978). There are an estimated 160 versions of the *A* locus and 79 of the *B* locus. Many versions of the genes exist, and their products can only be distinguished as compatible or not by intracellular events following cell fusion. This essential self/nonself-recognition step commits cells to development.

In this report, we focus on the function of the *A* mating-type genes. The *A* locus is complex, containing several functionally redundant, multiallelic genes separated into two subcomplexes (α and β) by a region of noncoding DNA (Figure 1). The genes encode a family of homeodomain proteins that are predicted to regulate development by controlling gene transcription. Our molecular analysis of several *A* loci (Kües et al., 1992, 1994a, 1994b) has shown that the number of genes is variable and that they encode two dissimilar classes of proteins, which we have designated HD1 and HD2 on the basis of conserved but distinct homeodomain sequences (Kües et al., 1992). The *A* locus has been called a haplotype (May et al., 1991; Kües et al., 1994c), and the archetype is considered to contain four divergently arranged pairs of paralogous *HD1* and *HD2* genes (Kües and Casselton, 1993). Various parts of the complex are missing in different *A* loci, as seen for *A42* and *A6* (Figure 1); each locus has only one complete gene pair (designated the *b* gene pair) but also solo genes representative of two (*A6*) or three (*A42*) other gene pairs (the *a*, *c*, and *d* pairs).

The HD1 and HD2 motifs of the *C. cinereus* A proteins are conserved in the mating-type proteins of other basidiomycete fungal species. The *Aα* locus of another mushroom, *Schizophyllum commune* (Stankis et al., 1992), has a single pair of divergently transcribed *HD1* and *HD2* genes encoding proteins very similar to those found in *C. cinereus*, as does the *b* mating-type locus of two smut fungi, *Ustilago maydis* and *U. hordei* (Gillissen et al., 1992; Bakkeren and Kronstad, 1993). The HD1 and HD2 motifs are related to homeodomain sequences in the a1 and α2 mating-type proteins of the budding yeast *Saccharomyces cerevisiae* (Astell et al., 1984). The *S. cerevisiae* proteins are known to heterodimerize following mating to generate a functional transcription factor complex that binds target sites upstream of developmentally regulated genes

[1] To whom correspondence should be addressed.

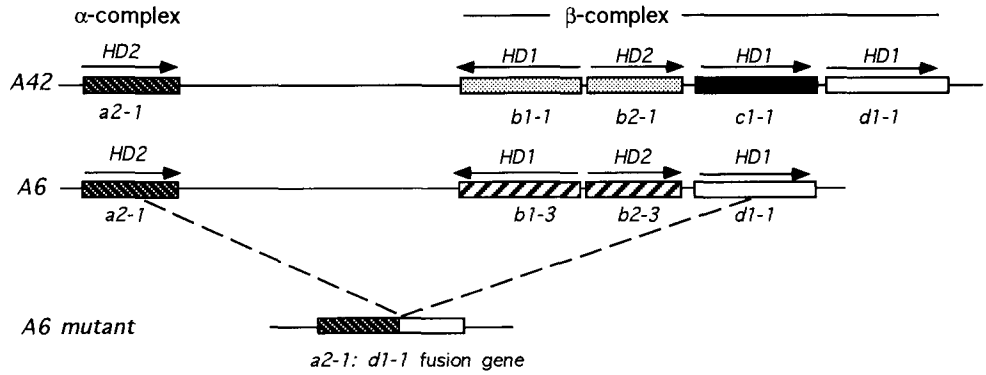

Figure 1. Organization of the Wild-Type *A42* and *A6* Loci of *C. cinereus* and a Mutant *A6* Locus.

The wild-type loci contain several *HD1* and *HD2* genes separated into the α and β complexes. Arrows indicate direction of transcription. The striped, stippled, black, and white rectangles indicate genes belonging to paralogous gene pairs (*a*, *b*, *c*, or *d*) and different alleles of the *b* gene pair. The *A6* mutant locus arose by deletion and fusion of the paralogous *a2-1* and *d1-1* genes.

(Herskowitz, 1988; Dranginis, 1990; Goutte and Johnson, 1993; Mak and Johnson, 1993).

Transformation studies using host cells that have genes either experimentally (Gillissen et al., 1992) or naturally deleted (Specht et al., 1992; Kües et al., 1994b) have led to the hypothesis that a similar interaction between HD1 and HD2 proteins might regulate development in the mushroom and smut fungi. Direct evidence has been provided by our analysis of a rare dominant mutation in the *A6* locus of *C. cinereus* that results in a single chimeric gene, part *HD2* and part *HD1*. This gene encodes a fused heterodimer that constitutively promotes *A*-regulated development. Fusion brings together essential regions of the two proteins, the HD2 homeodomain as the potential DNA binding domain and a predicted activation domain in the C terminus of the HD1 protein. The normal requirement that each protein recognize a compatible dimerization partner has been circumvented, and the chimeric protein is sufficient to promote the *A*-regulated program of development in the absence of any other A protein (Kües et al., 1994c).

Unlike *S. cerevisiae*, where individual cells express either a1 or α2 proteins, each *C. cinereus* cell has several HD1 and HD2 A proteins, and the cell must be able to discriminate between the incompatible proteins present before mating and the compatible proteins brought together by mating partners. By generating chimeric genes with 5′ exchanges, we have previously shown that the N-terminal domain of the proteins determines specificity within the same locus (see Figure 1). In *U. maydis*, where the mating-type locus contains only a single pair of multiallelic HD1 and HD2 genes (*bE* and *bW*), the N-terminal regions of both classes of protein have been implicated in determining allele specificity (Gillissen et al., 1992; Yee and Kronstad, 1993). The N-terminal regions of both classes of *C. cinereus* A proteins are predicted to be helical (Tymon et al., 1992; Kües et al., 1994a), and previously we suggested that these regions might constitute a dimerization domain. It is significant that the regions N-terminal to the homeodomains in all the HD1 and HD2 proteins and the a1 and α2 proteins of *S. cerevisiae* are similar in length (Astell et al., 1984; Kües et al., 1994a). Two domains have been implicated in heterodimerization between a1 and α2. One of the domains lies within the N-terminal regions of the proteins and is mediated by two leucine zipperlike coiled-coil motifs (Ho et al., 1994). An analogous dimerization domain in the *C. cinereus* HD1 and HD2 proteins is suggested by the presence of similar motifs (Gieser and May, 1994).

In this study, we investigated the role of the N-terminal domains of the *C. cinereus* HD1 and HD2 A proteins in regulating gene transcription and in mediating heterodimerization. Our data suggest that this domain is responsible for self/nonself-recognition by permitting heterodimerization only between compatible protein partners.

RESULTS

N-Terminal Domains of HD1 and HD2 Proteins Are Not Required for Regulating Gene Transcription in a Fused Heterodimer

Following cell fusion, the two nuclei derived from each mate remain associated but do not fuse. Their regular distribution in the dikaryotic mycelium is maintained by a complex cell division that results in the formation of a specialized structure known as a clamp connection. Following synchronized division in the tip cell, one of the daughter nuclei must pass through the clamp connection into a uninucleate subterminal cell. The *A*-regulated pathway of development leads to formation of the clamp cell, but its fusion to the subterminal cell requires activity of the *B*-regulated pathway (Swiezynski and Day, 1960). When only the *A*-regulated sequence is active, the clamp cells form but remain unfused; we can use this phenotype to assay

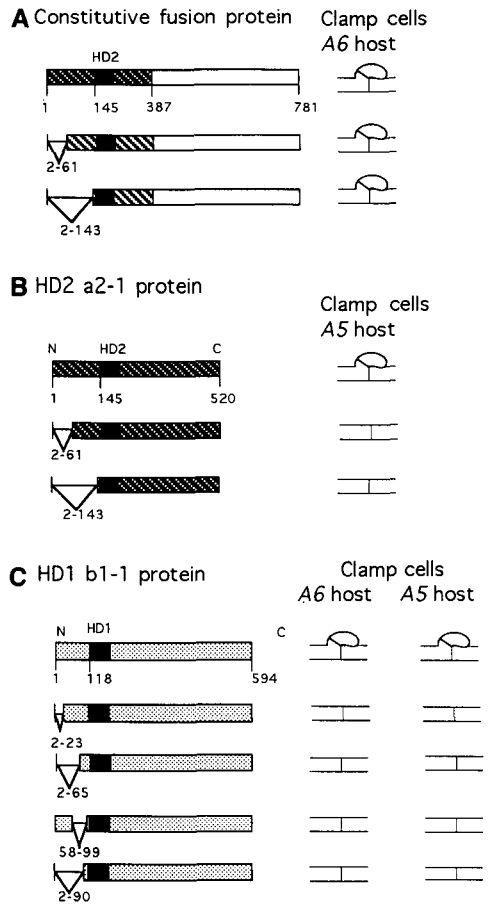

Figure 2. Effect of 5′ Deletions on the Ability of Genes To Promote Clamp Cell Development.

(A) The constitutive fusion gene.
(B) The *HD2* gene *a2-1*.
(C) The *HD1* gene *b1-1*.
Diagrams indicate the amino acid sequences of the proteins, and numbers indicate (below the bars) the amino acids N-terminal to the homeodomain, the position of the fusion between a2-1 and d1-1 in the mutant, amino acid coordinates of the N-terminal deletions, and the length of each protein. The *b1-1* gene used in these experiments lacked an inessential C-terminal 38 amino acids. Graphic symbols used to distinguish the proteins encoded by the different genes are as illustrated in Figure 1. Development of clamps is indicated by a clamp cell.

the function of cloned *A* genes, following transformation into a suitable host (Mutasa et al., 1990; Kües et al., 1992).

HD1 and HD2 proteins encoded by the paralogous genes in the same locus are normally unable to promote development. However, a rare mutation in the *A6* locus has fused the *a2-1* gene to the 3′ half of the *d1-1* gene (Kües et al., 1994c; see Figure 1) and overcome this normal incompatibility by generating a fused heterodimer that constitutively promotes clamp cell development. We used transformation into an *A6* wild-type background to demonstrate this constitutive function because there are no potentially compatible A proteins with which this fused heterodimer could interact.

The fusion protein lacks the N-terminal sequence of the HD1 protein (including the homeodomain); thus, this region of the fusion protein cannot be essential for regulating transcription of other genes (Figure 2). By generating 5′ deletions of the gene, we were able to show that the N-terminal sequence of the HD2 portion of the fusion protein is also not essential for its constitutive function (Figure 2A). Two versions of the fusion gene were generated by polymerase chain reaction in which the promoter sequences were retained, but sequences encoding amino acids 2 to 62 (pA6md1) or 2 to 143 (pA6md2) were deleted. The latter had only two amino acids in front of the HD2 homeodomain. The complete fusion gene and these two 5′ deleted versions were introduced into the *A6* test host, and all three were capable of constitutively promoting clamp cell development.

The strategy used to generate the 5′ deleted fusion gene was also used to create correspondingly deleted *a2-1 HD2* genes. These were also tested for their ability to promote clamp cell development. *a2-1* could not promote clamp cell development in the *A6* host because it was already present, but it found a compatible partner in a host with an *A5* locus that shares none of the genes in *A42* and *A6* (S.F. O'Shea, E.H. Pardo, and L.A. Casselton, unpublished data). The wild-type gene elicited clamps in this host, but neither of the deleted genes did (Figure 2B). Four 5′ truncated versions of the *HD1* gene *b1-1* were tested also. These were generated by polymerase chain reaction so that the promoter sequence was retained, but sequences encoding amino acids 2 to 23, 2 to 65, 2 to 90, and 58 to 99 were deleted (Figure 2C). *b1-1* found a compatible partner in both *A5* and *A6* hosts, but the truncated genes were inactive in both (Figure 2C). We concluded that the N-terminal regions of the HD1 and HD2 proteins have an essential function no longer required by the fused heterodimer.

N-Terminal Regions of the A Proteins Determine Allele Specificity

Previous studies with genes from *A42* (see Figure 1) have implicated the 5′ ends in determining gene specificity within the *A* locus (Kües et al., 1994a). We now asked whether the 5′ ends also determine allele specificity (Figure 3). By using conserved restriction sites within the homeodomain-encoding regions of *b1-1* and *b1-3* and of *b2-1* and *b2-3*, we generated two chimeric genes, namely, *b1-3:b1-1* and *b2-1:b2-3*. The wild-type and chimeric genes were introduced into both *A6* and *A42* hosts. Wild-type *b1-1* and *b2-1* elicited clamps in the *A6* host, and *b1-3* and *b2-3* elicited clamps in the *A42* host. The chimeric genes in both cases behaved like the gene from which they derived their 5′ ends; *b1-3:b1-1* behaved like *b1-3*, and *b2-1:b2-3* behaved like *b2-1*. This result is consistent with the conclusion that the 5′ ends of the genes, and hence the N-terminal

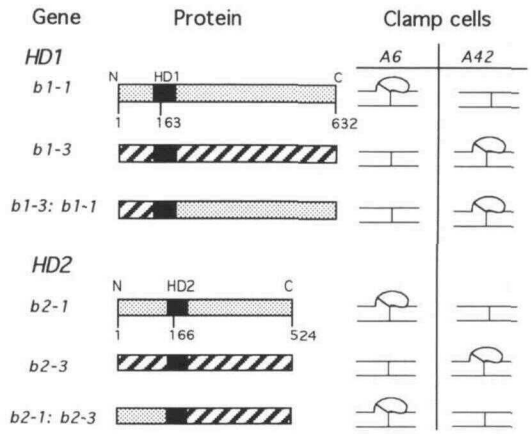

Figure 3. Effect of 5′ Exchanges on the Mating Specificity of Allelic Versions of the *HD1 b1* and *HD2 b2* Genes.

Shading is used to distinguish between the two allelic versions of each protein (as given for Figure 1). Numbers indicate the positions at which the amino acid sequences were exchanged.

regions of the proteins, determine allele specificity in addition to gene specificity.

Compatible HD1 and HD2 Proteins Can Heterodimerize in Vitro

Of the genes encoded by the *A42* and *A6* loci (Figure 1), the only two compatible *HD1* and *HD2* combinations predicted are *b1-1* plus *b2-3* and *b1-3* plus *b2-1.* Alleles of paralogous genes are either shared (*b2-1* and *d1-1*) or are inactive (*c1-1*). Several incompatible protein combinations are present and include b1-1 plus b2-1 or a2-1 and b1-3 plus b2-3 or a2-1. Having implicated the N-terminal regions in determining the specificity of the compatible interaction, we now asked whether these regions mediate protein–protein interactions and whether these interactions are specific.

We have used the in vitro glutathione *S*-transferase (GST) association assay first developed by Blackwood and Eisenman (1991) to demonstrate dimerization between A proteins. Six different proteins were expressed in *Escherichia coli* as GST fusion proteins. Following binding to glutathione *S*-Sepharose beads, the fusion proteins were tested for their ability to bind an in vitro–translated ^{35}S-labeled potential dimerization partner.

In the first experiment, illustrated in Figure 4A, we presented the full-length in vitro–translated HD1 b1-1 protein from *A42* with three potential HD2 partners. The HD2 proteins were not full length but contained the entire region N-terminal to the homeodomain, together with varying amounts of C-terminal sequence. The only fusion protein that specifically retained b1-1 was its compatible partner from *A6*, b2-3. This experiment was repeated using the full-length HD1 b1-3 protein from *A6* (Figure 4B), which was specifically retained by its compatible

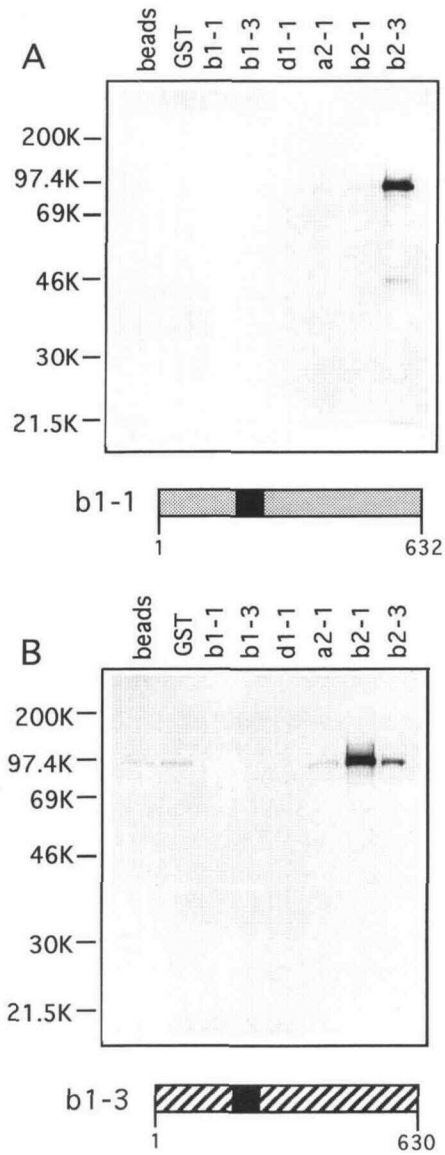

Figure 4. In Vitro Protein–Protein Interactions in Solution between *C. cinereus* A Mating-Type Proteins.

Full-length HD1 proteins were ^{35}S-labeled by translation in vitro and adsorbed to a range of GST fusion proteins coupled to glutathione *S*-Sepharose beads. The lanes contain the following: beads; the GST tag alone (as controls for nonspecific binding); and N-terminal mating-type protein GST fusions. The HD1 proteins are b1-1 (XCT.1, amino acids 1 to 163), b1-3 (pAB97, amino acids 1 to 158), and d1-1 (pAB76, amino acids 1 to 233). The HD2 proteins are a2-1 (pAB79, amino acids 1 to 257), b2-1 (pAB73, amino acids 1 to 166), and b2-3 (pAB47, amino acids 1 to 224).

(A) Interactions with b1-1.

(B) Interactions with b1-3.

Diagrams below the gels represent the in vitro–translated b1-1 and b1-3 proteins, distinguished by stippled and striped shading, respectively, and homeodomain sequences, indicated by a black box. Molecular mass markers are given at the left in kilodaltons (K).

partner from *A42* b2-1. A weak interaction was observed between the incompatible b1-3 and b2-3 proteins (only 20% of the binding observed to the compatible partner b2-1). We considered this interaction to be nonspecific because we observed no comparable interaction between the corresponding incompatible b1-1 and b2-1 proteins in the previous experiment. Moreover, the actual sequence mediating this weak interaction (which is also seen, to a lesser extent, in the controls) was at the C terminus of b1-3 and was no longer observed in a subsequent experiment when the protein was truncated (see later text).

In the yeast *S. cerevisiae*, the α2 protein, which is considered analogous to the *C. cinereus* HD1 proteins, has a role independent of its compatible partner a1. It homodimerizes via an N-terminal domain (Smith and Johnson, 1992; Vershon and Johnson, 1993) and binds other target sites with another transcription factor, MCM1. We have therefore looked for interactions between our HD1 proteins. Included in the experiments illustrated in Figures 4A and 4B are N-terminal b1-1, b1-3, and d1-1 HD1 GST fusion proteins. We found no evidence to suggest that N-terminal interactions occur in solution between proteins derived from the same *HD1* gene (b1-1/b1-1) or from different *HD1* genes (b1-1/b1-3 or b1-1/d1-1).

N-Terminal Specificity Is Conferred by the Ability To Form HD1/HD2 Dimers

By exchanging the 5′ ends of genes at conserved sites within the region encoding the homeodomain, we showed that allele specificity of both *HD1* and *HD2* genes in vivo is a function of the N terminus of the proteins (Figure 3). By exchanging the 5′ ends of the *b1-1* and *b1-3 HD1* genes at these same conserved sites, we generated constructs that translated chimeric proteins in vitro and presented these with the same potential HD2 partners as in the previous experiment, that is, a2-1, b2-1, and b2-3. The b1-1:b1-3 protein (Figure 5A) was retained by b2-3, and the b1-3:b1-1 protein (Figure 5B) was retained by b2-1. This result exactly reproduces the change in specificity observed in our in vivo experiment. Noting the positions at which the exchanges were made, we concluded that the regions that determine a compatible interaction lie within the N-terminal 158 amino acids of HD1 proteins. The weak nonspecific interaction of the b1-1:b1-3 protein with b2-1 (only 5.8% of that bound to b2-3) was observed (Figure 5A), and as in the previous experiment, this was attributed to a sequence at the C terminus of b1-3.

N-Terminal Regions of HD1 Proteins Are Sufficient To Mediate Dimerization

Suitable restriction sites in the genes allowed us to translate truncated versions of the b1-1 and b1-3 proteins to determine whether their N-terminal regions were sufficient to mediate specific heterodimerization. As seen in Figures 6A and 6B, the

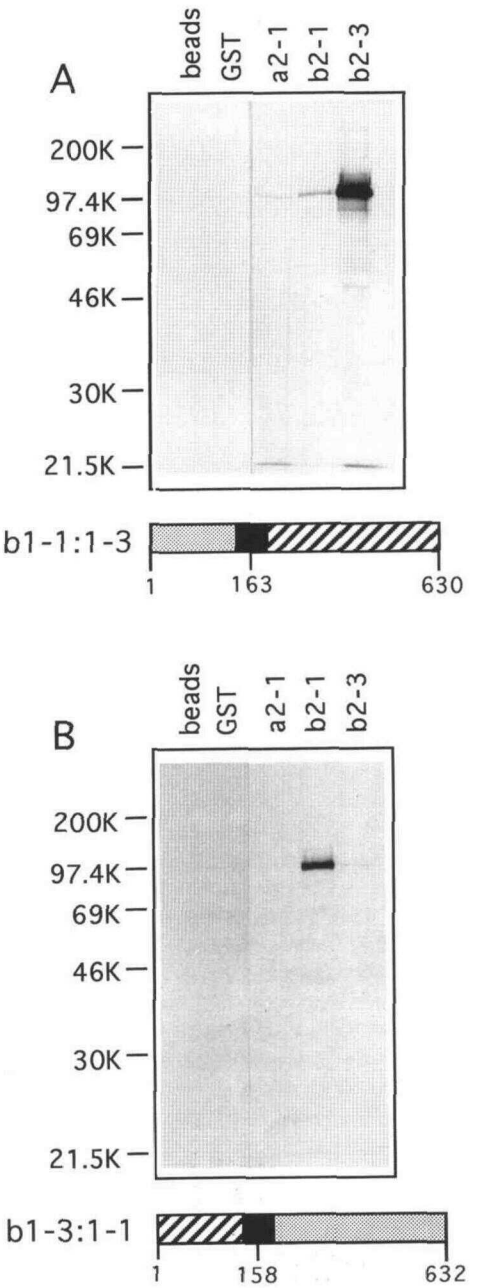

Figure 5. The Specificity of in Vitro HD1 and HD2 Protein Dimerization Mediated by Chimeric HD1 Proteins.

Constructs in which the 5′ ends of genes *b1-1* and *b1-3* were exchanged and used to translate the full-length HD1 chimeric proteins. Lane contents are beads, GST, and the HD2 proteins, as described in the legend to Figure 4.

(A) Interactions with b1-1:b1-3.

(B) Interactions with b1-3:b1-1.

Diagrams below the gels indicate the sequences derived from b1-3 (striped) and b1-1 (stippled) and the positions at which the amino acid sequences were exchanged.

Molecular mass markers are given at the left in kilodaltons (K).

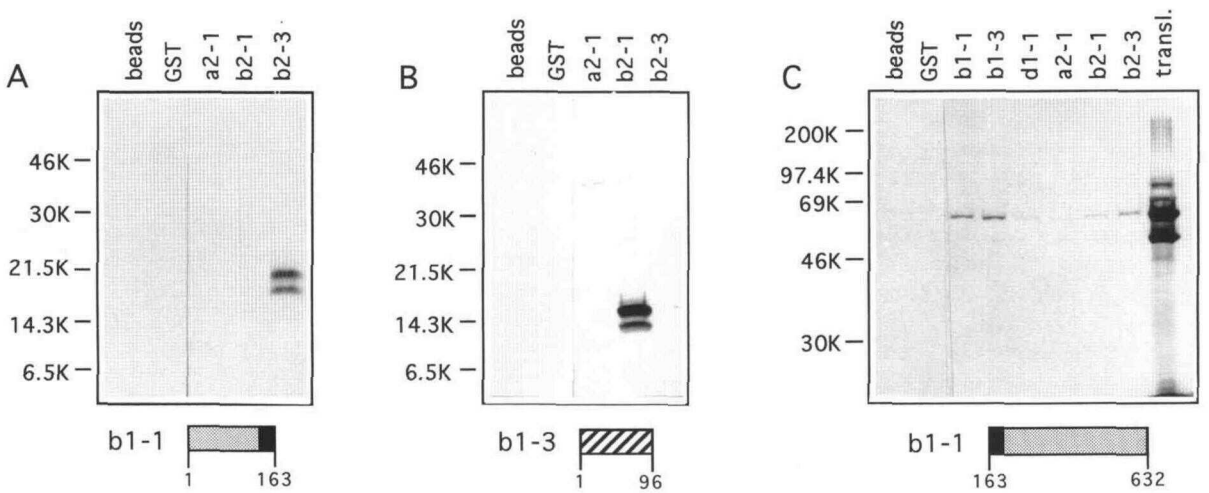

Figure 6. The N-Terminal Regions of HD1 Proteins Are Sufficient To Mediate Dimerization in Vitro.

(A) and **(B)** Interactions with the N-terminal 163 amino acids of b1-1 and N-terminal 96 amino acids of b1-3. Proteins were translated in vitro and adsorbed to the controls and N-terminal HD2 GST fusion proteins as described in the legend to Figure 4. Diagrams below the gels indicate the lengths of the proteins translated.

(C) Interactions with the C-terminal 470 amino acids of b1-1. The protein was translated in vitro and adsorbed to Sepharose beads and GST as controls or coupled GST fusion proteins. Full-length HD1 fusion proteins were b1-1 (pAB56), b1-3 (pAB100), and d1-1 (pAB70). The HD2 fusion proteins were a2-1 (pAB78, amino acids 1 to 358), b2-1 (pb2-1EX.2, amino acids 59 to 320), and b2-3 (pAB87, amino acids 1 to 249). The final lane (transl.) shows the amount of labeled protein presented to the pGEX fusion proteins. The diagram below the gel indicates the length of the b1-1 protein translated.

Molecular mass markers are given at the left in kilodaltons (K).

N-terminal 163 amino acids of b1-1 and 96 amino acids of b1-3 were sufficient to mediate dimerization with their compatible HD2 partners. For the shorter b1-3 fragment, we ruled out involvement of any homeodomain sequences promoting this interaction.

Other predicted helical regions lie immediately C-terminal to the homeodomains in both classes of proteins and at the extreme C terminus of HD1 proteins. It has been proposed that these could also mediate protein–protein interactions (Gieser and May, 1994; Kües et al., 1994a). In a separate experiment (Figure 6C), the C-terminal region of b1-1 (163 to 632 amino acids) was in vitro translated and tested for interactions with full-length HD1–GST fusions and longer HD2–GST fusions. (Full-length HD2 proteins have proved unstable in *E. coli*, but the expressed proteins contained all helical regions.) The final lane in Figure 6C shows the amount of translated protein presented to the fusion proteins. Unlike the compatible interactions observed in Figures 4 and 5, where consistently more than 50% of the in vitro–translated protein was retained on the beads, only a weak interaction (at most, 9%) was observed between b1-1 and the GST proteins, and this occurred to varying extents with all of them. We concluded that the C-terminal regions of the proteins are unlikely to mediate specific dimerizations in solution. Our overall conclusion from these in vitro experiments is that specific interactions in solution are only possible between compatible proteins.

Predicted Motifs within the N-Terminal Domains

Using the algorithm derived by Lupas et al. (1991), Gieser and May (1994) predicted the presence of two coiled coils within the N-terminal region of two allelic HD1 *C. cinereus* A proteins,

Table 1. Coiled-Coil Motifs Predicted in HD1 Mating-Type Proteins

	First Coil		Second Coil	
Protein	Amino Acids	Score[a]	Amino Acids	Score[a]
b1-1	15 to 33	1.49 to 1.62	73 to 99	1.57 to 1.82
b1-3	18 to 33	1.16	75 to 95	1.46 to 1.57
b1-2	14 to 33	1.31 to 1.41	75 to 99	1.21 to 1.51
d1-1	37 to 59	1.34 to 1.51	79 to 100	1.25 to 1.27
α2	16 to 29	1.48	60 to 103	1.41 to 1.68

Sequence analysis was performed using the method described by Lupas et al. (1991). In this method, amino acid sequences are analyzed in terms of empirically determined residue preferences for the seven positions (*a* through *g*) in a coiled-coil heptad repeat.

[a] The score is related to the probability of the occurrence of a coiled coil at a given position in the sequence. A score of 1.3 or above indicates a significant probability of coil formation.

b1-1, used in this study, and b1-2 from the *A43* locus. Using the same program (COILS), we extended this analysis to include the other HD1 proteins used in this study, b1-3 and d1-1 (Table 1).

N-terminal coiled-coil motifs have been implicated in a1/α2 heterodimerization in *S. cerevisiae* (Ho et al., 1994). Although some of the values obtained for the *C. cinereus* coiled coils predict only weak motifs, with a score below the threshold value of 1.5 (using a window of 14 residues), those in the *S. cerevisiae* α2 protein have comparable low scores. The HD1 b protein coiled coils are found in much the same N-terminal position as those in the α2 protein, that is, between amino acids 14 to 33 and 73 to 99. Interestingly, the paralogous HD1 protein d1-1 has the position of the first coil shifted to amino acids 37 to 59.

DISCUSSION

Sexual development in *C. cinereus*, as in *S. cerevisiae*, is regulated by two dissimilar classes of homeodomain proteins. Heterodimerization between the a1 and α2 proteins determines the developmental fate of mated *S. cerevisiae* cells by generating a diploid cell–specific transcription-factor complex. In this paper, we present evidence that development is similarly regulated by heterodimerization between the HD1 and HD2 homeodomain proteins encoded by the *A* mating-type locus of *C. cinereus*. Our in vitro studies indicate that HD1 and HD2 homeodomain proteins are able to heterodimerize in solution via N-terminal domains that have predicted structural similarity to corresponding dimerization domains in the a1 and α2 proteins.

Unlike *S. cerevisae*, the basidiomycete fungi have evolved multiple mating types (Kües and Casselton, 1992). In *C. cinereus*, this has been achieved by having several multiallelic, paralogous genes at each mating-type locus. An unmated cell has several functionally equivalent HD1 and HD2 proteins that must be prevented from promoting development. These same proteins, however, must be able to interact successfully with other versions of the proteins following mating. From the results of the experiments described in this study, we suggest that incompatible protein partners are distinguished from compatible ones in that they are unable to heterodimerize. We used a GST association assay in which we presented the b1-1 protein, encoded by an *HD1* gene in the *A42* locus, with two incompatible HD2 partners encoded by genes at the same locus, a2-1 and b2-1, and a compatible partner encoded by a different allele of *b2* at the *A6* locus, b2-3. Only the compatible protein pair, b1-1 and b2-3, dimerized. We showed, similarly, that the HD1 protein b1-3 encoded by the *A6 b1* gene only dimerized with its compatible partner b2-1 from *A42*.

By truncating the proteins used in the in vitro assay, we identified sequences N-terminal to the homeodomain as those mediating dimerization. These regions of both protein classes were implicated in determining the specificity of a compatible protein interaction in vivo by exchanging the 5′ ends of both paralogous genes (Kües et al., 1994a) and alleles of the same gene (this study). In all cases, chimeric genes had the mating-type specificity of the gene from which they derived their 5′ ends and thus the N-terminal domains of the proteins. The results of these in vivo experiments were the same as the results of in vitro experiments with similarly derived chimeric proteins. We exploited the unique properties of the *A6* mutant fusion gene to demonstrate, using 5′ gene deletions, that these normally essential specificity domains of both classes of proteins are not necessary for *A*-regulated transcription of other genes. We suggest that their major role is to allow HD1 and HD2 proteins to choose compatible partners by dimerization.

The Lupas et al. (1991) COILS program predicts two coiled-coil motifs in the N-terminal domains of HD1 proteins (Gieser and May, 1994; Table 1). Several potential coiled coils in the N-terminal regions of the HD2 proteins with which the HD1 coils could interact are also predicted by this analysis, but their positions are not so clearly defined. N-terminal coils are also predicted in the analogous HD1 proteins of *S. commune* (AαZ) and the bE proteins of *U. maydis* (data not shown), suggesting that these are conserved features of basidiomycete HD1 mating-type proteins. These motifs are found in the corresponding N-terminal regions of both α2 and a1 of *S. cerevisae* (Ho et al., 1994) and have been implicated in dimerization by mutation. Deletions and amino acid substitutions in the *d* positions of the 3,4-hydrophobic heptad repeat sequences of α2 protein disrupted heterodimer stability in vitro and reduced a1/α2 activity in vivo (Ho et al., 1994).

The N-terminal amino acid sequences of different *C. cinereus* HD1 and HD2 proteins are highly variable (Kües et al., 1994a), and the dimerization potential of proteins is undoubtedly influenced by different amino acids. Currently, we do not have sufficient protein sequences to make useful predictions about why, for example, b2-1 can heterodimerize with several allelic variants of b1 but cannot dimerize with its allelic partner in *A42*, b1-1. Failure to dimerize would appear to be a more specific interaction than the ability to dimerize. The *e* and *g* positions that flank the hydrophobic *a* and *d* positions of the 3,4-heptad repeat often contain charged amino acids that are thought to interact electrostatically (Cohen and Parry, 1990; Baxevanis and Vinson, 1993), and comparisons of dimer stabilities of other proteins that interact by means of coiled coils, such as Fos, Jun, and GCN4, suggest that it is the amino acids at these positions that determine the specificity of the interactions (O'Shea et al., 1992; Glover and Harrison, 1995). With so many genes encoding the HD1 and HD2 A proteins of *C. cinereus*, there may be more than one constraint on dimerization, and the actual positions of the N-terminal coils may be one of these. Although the second predicted coil in all four HD1 proteins we examined was at the same position (amino acids 73 to 100), the first coil in the paralogous d1-1 protein was displaced (amino acids 37 to 59) relative to that in the three allelic versions of b1 (amino acids 14 to 33). It would have been

interesting to look for this coil in other d1 proteins, but so far we have only identified this one allele in our stock collection.

The entire region N-terminal to the homeodomain of both HD1 and HD2 proteins is essential for a normal compatible interaction in vivo, because none of the 5′ deleted genes had function in vivo. Deletions of amino acids 2 to 65 would have removed the first predicted coil (amino acids 14 to 33) in the HD1 b1-1 proteins, and deletions of amino acids 58 to 99 would have removed the second (73 to 99). Interestingly, deletions of one or the other predicted coil from the α2 protein of *S. cerevisiae* only reduced the stability of the heterodimer (Ho et al., 1994). With the complex function of discriminating incompatible protein partners, dimerization through this region may have acquired a more essential function for the *C. cinereus* proteins.

Significance of Heterodimerization of Homeodomain Proteins

The homeodomain sequence is very conserved, and choosing a dimerization partner for cooperative DNA binding can confer specificity for target site selection. In the case of the α2 protein of *S. cerevisae*, association with either a1 or MCM1 imposes different spacing of otherwise similar homeodomain contact sites on DNA (Smith and Johnson, 1992) and leads to recognition of different targets and regulation of different sets of genes. Sequences adjacent to the homeodomains are critical in these protein–protein interactions, a C-terminal dimerization domain for a1/α2 and a sequence immediately N-terminal to the homeodomain for α2/MCM1, indicating that conformational effects may be important (Goutte and Johnson, 1993, 1994; Mak and Johnson, 1993; Stark and Johnson, 1994). Such interactions are not confined to fungal homeodomain proteins, and similar examples can be found in Drosophila, mammals, and *Caenorhabditis* (see reviews in Manak and Scott, 1993; White, 1994). The N-terminal dimerization domain in the *C. cinereus* A proteins is not required to promote target site selection because we have shown it to be dispensible in the chimeric fusion gene product. We have suggested that a conserved α-helical domain in both proteins just C-terminal to the homeodomains may be analogous to the C-terminal dimerization domain in the *S. cerevisiae* proteins (Kües et al., 1994a). If this C-terminal dimerization were DNA dependent, as it is with a1 and α2 (Dranginis, 1990; Mak and Johnson, 1993), we would have been unable to detect it in our in vitro assay in the absence of DNA.

The N-terminal domain of the mating-type proteins may be more analogous to an N-terminal domain of the POU-like homeodomain protein HNF-1, which also mediates dimerization in solution. This is thought to allow HNF-1 to choose different proteins as dimerization partners to diversify the activities associated with a single target site (Mendel et al., 1991). In the proteins we studied, choosing a dimerization partner is not required to diversify function or to select a different target site but to act as a mechanism that allows cells to distinguish self from nonself.

METHODS

Coprinus cinereus Strains

Strains used as transformation hosts were the tryptophan auxotrophs LN118 (*A42B42 ade-2 trp-1.1,1.6*), LT2 (*A6B6 trp-1.1,1.6*), and FA2222 (*A5B5 trp-1.1,1.6*). Plasmids containing *A* mating-type genes were cotransformed with plasmid pCc1001 (Binninger et al., 1987) containing the *C. cinereus trp-1* gene. At least 50 trp^+ transformants from each experiment were screened microscopically for the presence of clamp cells. Media and general methods of culture were those described by Lewis (1961) and Mutasa et al. (1990). The transformation procedure was based on Casselton and de la Fuente Herce (1989).

Plasmid Constructs Used for in Vivo Studies

pBluescript KS− clones containing genes used in *C. cinereus* transformations were pAMT1 (complete *b1-1* sequence) and pAMT2 and pAMT3 (a 3′ truncated version of *b1-1* lacking the terminal 120 bp), as described by Tymon et al. (1992). pESM1 and pUK4 (*b2-1*), pESM2 (*4-1*), and pUK2 (*a2-1*) are described by Kües et al. (1992). pA626 and pA625 contain the *b1-3* and *b2-3* genes on 2.2- and 5.5-kb XhoI fragments, respectively. pWRF1 contains the *A6* fusion gene (Kües et al., 1994c). Chimeric gene *b1-3/1-1* was constructed using a conserved ClaI site within the homeobox sequences of *b1-1* (in pAMT3) and *b1-3* (in pA626) to give pb1-3/1. Chimeric gene *b2-1/2-3* was constructed using a conserved NcoI site immediately 3′ to the homeodomain coding sequence using *b2-1* (in pUK4) and *b2-3* on a 1.9-kb XhoI-SacII fragment derived from pA625 to give pb2-1/3. 5′ Gene deletions were generated by polymerase chain reaction. Primers used to truncate *a2-1* (in pUK2) were 5′-CATGGTGAAGGCGATGTTGGTGG-3′, with either 5′-CTCGCCTACGATGTCCCTCG-3′ or 5′-TAGCTGCCCCTCCATCAAAG-3′ to delete nucleotides 4 to 183 and 4 to 429, respectively. Primers used to truncate *b1-1* (in pAMT2) were 5′-CATTGCTAGGGGATAGCCGCG-3′ with 5′-ACGCTGCGCGCGGATCTGTCTG-3′ (deleting nucleotides 4 to 66), with 5′-CTCGATCTGCTCTATTCTTTCTCC-3′ (deleting nucleotides 4 to 192), and with 5′-GCCGAATCAAACAGTTCACAGCG-3′ (deleting nucleotides 4 to 270). 5′-ATGGCAAGATTGGGCGAAACC-3′ and 5′-GCCATCCTTTCAGACAAGACC-3′ were used to delete nucleotides 153 to 299. NheI-SalI fragments encompassing the deletions in *a2-1* were subcloned for sequencing and then used to replace the appropriate nondeleted sequence in pUK2 (to give pA2d1 and pA2d2) and the *a2-1* sequence of the *A6* fusion gene in pWFR1 (to give pA6d1 and pA6d2). HindIII-EcoRV fragments encompassing the deletions in *b1-1* were similarly subcloned for sequencing and used to replace the appropriate nondeleted sequence in pAMT2.

Plasmid Constructs Used for in Vitro Assays

cDNAs for *A42* genes were derived from a λgt10 library (Mutasa et al., 1990), and cDNAs for *A6* genes were derived from a newly constructed library in λZAP II (Stratagene). Table 2 lists all plasmids

Table 2. Plasmid Constructs for Protein Expression Used in This Study

Plasmid	cDNA	Amino Acids	Vector
pAB52	*b1-1*	1 to 632	pTM1
pAB109	*b1-1*	1 to 163	pTM1
pAB81	*b1-1*	163 to 632	pTM1
pAB94	*b1-1:b1-3*	1 to 163:159 to 630	pTM1
pAB93	*b1-3*	1 to 630	pTM1
pAB95	*b1-3*	1 to 96	pTM1
pAB88	*b1-3:b1-1*	1 to 158:164 to 632	pTM1
pAB72	*d1-1*	1 to 632	pTM1
pAB56	*b1-1*	1 to 632	pGEX-2T
pXCT.1	*b1-1*	1 to 163	pGEX-2T
pAB100	*b1-3*	1 to 630	pGEX-3X
pAB97	*b1-3*	1 to 158	pGEX-3X
pAB70	*d1-1*	1 to 632	pGEX-1
pAB76	*d1-1*	1 to 233	pGEX-1
pAB78	*a2-1*	1 to 358	pGEX-2T
pAB79	*a2-1*	1 to 257	pGEX-2T
pb2-1EX.2	*b2-1*	59 to 320	pGEX-3X
pAB73	*b2-1*	1 to 166	pGEX-2T
pAB87	*b2-3*	1 to 249	pGEX-2T
pAB47	*b2-3*	1 to 224	pGEX-2T

cDNA sequences encoding fragments of the A proteins were cloned into the vector pTM1 for translation in vitro and into pGEX vectors for expression in *E. coli*. The amino acid numbers refer to the region of the protein translated, starting with 1 as the N-terminal methionine. A colon indicates a chimeric cDNA construct and corresponding amino acid sequences contributed by each gene.

generated for protein expression in *Escherichia coli* and for in vitro transcription and translation.

T7 in Vitro Transcription and Translation of Proteins

T7 expression plasmid DNA (1 to 2 μg) was transcribed and translated in vitro using a TNT T7 coupled reticulocyte lysate system (Promega), labeling with ^{35}S-methionine, according to the manufacturer's instructions. cDNAs for T7 expression were cloned into the T7 expression vector pTM1 (Elroy-Stein et al., 1989).

Expression of Glutathione *S*-Transferase Fusion Proteins and Preparation of Extracts

cDNAs for expression were cloned into pGEX vectors (Pharmacia Biotech Ltd.) and transformed into *E. coli* BL21 (DE3) (Studier et al., 1990). Cells were grown overnight at 37°C in Luria-Bertani medium plus 100 μg/mL ampicillin. Twenty milliliters of the overnight culture was used to inoculate 200 mL of Luria-Bertani medium plus 200 μg/mL ampicillin. After shaking at 37°C for 1 hr, expression of the fusion protein was induced by the addition of 1 mM isopropyl β-D-thiogalactopyranoside. After continuing the incubation for either 3 hr (pGEX, xCT.1, pAB97, pAB76, pAB79, pAB73, pAB47, pAB78, pb2-1 Ex.2, and pAB87) or 6 hr (pAB53, pAB56, pAB100, and pAB70), cells were pelleted and washed with 20 mL of PBS; the cell pellet was then frozen at −70°C for 1 hr. The cells were defrosted, resuspended in 20 mL of MTPBS (150 mM NaCl, 16 mM $Na_2HPO_4 \cdot 7H_2O$, 4 mM NaH_2PO_4, pH 7.3), and then lysed by sonication. Cell debris was removed by centrifugation, and the supernatant was aliquoted in 1-mL volumes and stored at −70°C. The presence of the fusion protein was confirmed by SDS-PAGE.

pGEX Dimerization Assay

Using this in vitro technique, protein–protein interactions were identified by the ability of a glutathione *S*-transferase (GST) fusion protein bound to Sepharose beads (Pharmacia Biotech Ltd.) to retain a labeled partner through a series of wash steps. The dimerization assay was based on Zappavigna et al. (1994). Twenty-microliter aliquots of glutathione *S*-Sepharose beads (Pharmacia Biotech Ltd.) were washed three times with 1 mL MTPBS. One-milliliter aliquots of the appropriate GST fusion protein were bound to the beads by incubation at room temperature for 10 min, then 10 min on ice. The beads were washed twice with 1 mL of MTPBS and once with 1 mL of HND buffer (20 mM Hepes, pH 7.2, 150 mM NaCl, 5 mM DTT, 0.1% Nonidet P-40, 10 mg/mL BSA). Washed beads were resuspended in 200 μL HND buffer; after 5 min on ice, 5 μL of the in vitro–translated protein from a single reaction was added to each tube of beads. After incubating at 4°C and being gently shaken for 1 hr, the beads were washed four times with 1 mL of MTPBS containing 0.1% Nonidet P-40. The beads were resuspended in 50 μL of Laemmli sample buffer, and 20 μL of each sample was run on an SDS–polyacrylamide gel. The gel was fixed for 30 min (10% methanol, 10% acetic acid), amplified for 15 min (Amplify, NAMP 100; Amersham), dried, and autoradiographed.

ACKNOWLEDGMENTS

We thank Ian Connerton for help with the PCR strategies used to truncate genes, Mark Sansom for valuable help with the COIL computer analysis, Liz Cowe for helping us run the program, and Michael Milner for isolating and helping to sequence the *b1-3* cDNA. This work was supported by Biotechnology and Biological Sciences Research Council Grant Nos. PG043/0564 and PO2352 to L.A.C. and E.J.C.M., a research studentship to R.N.A.-O., and a postdoctoral fellowship to L.A.C. B.G. received support from a Science and Engineering Research Council studentship and the Gatsby Charitable Foundation. E.J.C.M. is a Wellcome Trust Senior Research Fellow.

Received February 22, 1995; accepted April 12, 1995.

REFERENCES

Astell, C.R., Ahlstrom-Jonasson, L., Smith, M., Tatchell, K., Nasmyth, K.A., and Hall, B.J. (1984). The sequence of the DNAs coding for the mating-type loci of *Saccharomyces cerevisiae*. Cell **27**, 15–23.

Bakkeren, G., and Kronstad, J.W. (1993). Conservation of the *b* mating-type gene complex among bipolar and tetrapolar smut fungi. Plant Cell **5,** 123–136.

Baxevanis, A.D., and Vinson, C.R. (1993). Interactions of coiled coils in transcription factors: Where is the specificity? Curr. Opin. Genet. Dev. **3,** 278–285.

Binninger, D.M., Skrzynia, C., Pukkila, P.J., and Casselton, L.A. (1987). DNA-mediated transformation of the basidiomycete *Coprinus cinereus.* EMBO J. **6,** 835–840.

Blackwood, E.M., and Eisenman, R.N. (1991). Max: A helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. Science **251,** 1211–1217.

Casselton, L.A. (1978). Dikaryon formation in higher basidiomycetes. In The Filamentous Fungi, Vol. 3, J.E. Smith and D.R. Berry, eds (London: Edward Arnold), pp. 275–297.

Casselton, L.A., and de la Fuente Herce, A. (1989). Heterologous gene expression in the basidiomycete fungus *Coprinus cinereus.* Curr. Genet. **16,** 35–40.

Cohen, C., and Parry, D.A.D. (1990). Proteins: Structure, function and genetics. α-Helical coiled coils and bundles: How to design an α-helical protein. Protein **7,** 1–15.

Draginis, A.M. (1990). Binding of yeast a1 and α2 as a heterodimer to the operator DNA of a haploid-specific gene. Nature **347,** 682–685.

Elroy-Stein, O., Fuerst, T.R., and Moss, B. (1989). Cap-independent translation on mRNA conferred by encephalomyocarditis virus 5′ sequence improves the performance of the vaccinia virus/bacteriophage T7 hybrid expression system. Proc. Natl. Acad. Sci. USA **86,** 6126–6130.

Gieser, P.T., and May, G. (1994). Comparison of two *b1* alleles from within the *A* mating-type of the basidiomycete *Coprinus cinereus.* Gene **146,** 167–176.

Gillissen, B., Bergemann, J., Sandmann, C., Schroer, B., Bölker, M., and Kahmann, R. (1992). A two-component system for self/nonself recognition in *Ustilago maydis.* Cell **68,** 647–657.

Glover, J.N.M., and Harrison, S.C. (1995). Crystal structure of the heterodimeric bZIP transcription factor c-Fos–c-Jun bound to DNA. Nature **373,** 257–261.

Goutte, C., and Johnson, A.D. (1993). Yeast a1 and α2 homeodomain proteins form a DNA-binding activity with properties distinct from those of either protein. J. Mol. Biol. **233,** 359–371.

Goutte, C., and Johnson, A.D. (1994). Recognition of a DNA operator by a dimer composed of two different homeodomain proteins. EMBO J. **13,** 1434–1442.

Herskowitz, I. (1988). Life cycle of the budding yeast *Saccharomyces cerevisiae.* Microbiol. Rev. **52,** 536–553.

Ho, C.-Y., Adamson, J.G., Hodges, R.S., and Smith, M. (1994). Heterodimerization of the yeast MATa1 and MATα2 proteins is mediated by two leucine zipper-like coiled-coil motifs. EMBO J. **13,** 1403–1413.

Kües, U., and Casselton, L.A. (1992). Fungal mating type genes: Regulators of sexual development. Mycol. Res. **96,** 993–1006.

Kües, U., and Casselton, L.A. (1993). The origin of multiple mating types in mushrooms. J. Cell Sci. **104,** 227–230.

Kües, U., Richardson, W.V.J., Tymon, A.M., Mutasa, E.S., Göttgens, B., Gaubatz, S., Gregoriades, A., and Casselton, L.A. (1992). The combination of dissimilar alleles of the *A*α and *A*β gene complexes, whose proteins contain homeo domain motifs, determines sexual development in the mushroom *Coprinus cinereus.* Genes Dev. **6,** 568–577.

Kües, U., Asante-Owusu, R.N., Mutasa, E.S., Tymon, A.M., Pardo, E.H., O'Shea, S.F., Göttgens, B., and Casselton, L.A. (1994a). Two classes of homeodomain proteins specify the multiple *A* mating types of the mushroom *Coprinus cinereus.* Plant Cell **6,** 1467–1475.

Kües, U., Tymon, A.M., Richardson, W.V.J., May, G., Gieser, P.T., and Casselton, L.A. (1994b). *A* mating-type factors of *Coprinus cinereus* have variable numbers of specificity genes encoding two classes of homeodomain proteins. Mol. Gen. Genet. **245,** 45–52.

Kües, U., Göttgens, B., Stratmann, R., Richardson, W.V.J., O'Shea, S.F., and Casselton, L.A. (1994c). A chimeric homeodomain protein causes self-compatibility and constitutive sexual development in the mushroom *Coprinus cinereus.* EMBO J. **13,** 4054–4059.

Lewis, D. (1961). Genetical analysis of methionine suppressors in *Coprinus.* Genet. Res. **2,** 141–155.

Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. Science **252,** 1162–1164.

Mak, A., and Johnson, A.D. (1993). The carboxy-terminal tail of the homeodomain protein α2 is required for function with a second homeodomain protein. Genes Dev. **7,** 1862–1870.

Manak, J.R., and Scott, M.P. (1993). Able assistants for homeodomain proteins. Curr. Biol. **3,** 318–320.

May, G., Le Chevanton, L., and Pukkila, P.J. (1991). Molecular analysis of the *Coprinus cinereus* mating type *A* factor demonstrates an unexpectedly complex structure. Genetics **128,** 529–538.

Mendel, D.B., Hansen, L.P., Graves, M.K., Conley, P.B., and Crabtree, G.R. (1991). HNF-1α and HNF-1β (vHNF-1) share dimerization and homeodomains, but not activation domains, and form heterodimers in vitro. Genes Dev. **5,** 1042–1056.

Mutasa, E.S., Tymon, A.M., Göttgens, B., Mellon, F.M., Little, P.F.R., and Casselton, L.A. (1990). Molecular organization of an *A* mating type factor of the basidiomycete fungus *Coprinus cinereus.* Curr. Genet. **18,** 223–229.

O'Shea, E.K., Rutowski, R., and Kim, P.S. (1992). Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. Cell **68,** 699–708.

Smith, D.L., and Johnson, A.D. (1992). A molecular mechanism for combinatorial control in yeast: MCM1 protein sets the spacing and orientation of the homeodomains of an α2 dimer. Cell **68,** 133–142.

Specht, C.A., Stankis, M.M., Giasson, L., Novotny, C.P., and Ullrich, R.C. (1992). Functional analysis of the Aα homeotic mating-type genes of *Schizophyllum commune.* Proc. Natl. Acad. Sci. USA **81,** 7174–7178.

Stankis, M.M., Specht, C.A., Yang, H., Giasson, L., Ullrich, R.C., and Novotny, C.P. (1992). The *A*α mating type locus of *Schizophyllum commune* encodes two dissimilar, multiallelic, homoedomain proteins. Proc. Natl. Acad. Sci. USA **89,** 7169–7173.

Stark, M.P., and Johnson, A.D. (1994). Interaction between two homeodomain proteins is specified by a short C-terminal tail. Nature **371,** 429–432.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185,** 60–89.

Swiezynski, K.M., and Day, P.R. (1960). Heterokaryon formation in *Coprinus lagopus.* Genet. Res. **1,** 114–128.

Tymon, A.M., Kües, U., Richardson, W.V.J., and Casselton, L.A. (1992). A fungal mating type protein that regulates sexual and asexual development contains a POU-related domain. EMBO J. **11,** 1805–1813.

Vershon, A.K., and Johnson, A.D. (1993). A short, disordered protein region mediates interactions between the homeodomain of the yeast α2 protein and the MCM1 protein. Cell **72,** 105–112.

White, R. (1994). Homeotic genes seek partners. Curr. Biol. **4,** 48–50.

Yee, A.R., and Kronstad, J.W. (1993). Construction of chimeric alleles with altered specificity at the *b* incompatibility locus of *Ustilago maydis*. Proc. Natl. Acad. Sci. USA **90,** 664–668.

Zappavinga, V., Sartori, D., and Mavilio, F. (1994). Specificity of Hox protein function depends on DNA–protein and protein–protein interactions, both mediated by the homeodomain. Genes Dev. **8,** 732–744.