Molecular Microbiology (2002) **44**(2), 585–597

PrnA, a Zn_2Cys_6 activator with a unique DNA recognition mode, requires inducer for *in vivo* binding

Dennis Gómez,[1] Beatriz Cubero,[1†] Gianna Cecchetto[1‡] and Claudio Scazzocchio[1,2*]
[1]*Institut de Génétique et Microbiologie, Université Paris-Sud, Bâtiment 409, UMR 8621 CNRS, 91405 Orsay Cedex, France.*
[2]*Institut Universitaire de France.*

Summary

The PrnA transcriptional activator of *Aspergillus nidulans* binds as a dimer to CCGG-N-CCGG inverted repeats and to CCGG-6/7N-CCGG direct repeats. The binding specificity of the PrnA Zn cluster differs from that of the Gal4p/Ppr1p/UaY/Put3p group of proteins. Chimeras with UaY, a protein that strictly recognizes a CGG-6N-CCG motif, show that the recognition of the direct repeats necessitates the PrnA dimerization and linker elements, but the recognition of the CCGG-N-CCGG inverted repeats depends crucially on the PrnA Zn binuclear cluster and/or on residues amino-terminal to it. Three high-affinity sites in two different promoters have been visualized by *in vivo* methylation protection. Proline induction is essential for *in vivo* binding to these three sites but, as shown previously, not for nuclear entry. Simultaneous repression by ammonium and glucose does not affect *in vivo* binding to these high-affinity sites. PrnA differs from the isofunctional *Saccharomyces cerevisiæ* protein Put3p, both in its unique binding specificity and in the requirement of induction for *in vivo* DNA binding.

Introduction

The Zn binuclear cluster motif is a DNA-binding domain characteristic of and exclusive to the fungi. A subset of these binding domains [Gal4p, Put3p, Ppr1p, UaY, Leu3p, Pdr1, Hap1(Cyp1)p] bind inverted, everted or direct repeats of a CGG triplet (Carey *et al*., 1989; Marmorstein *et al*., 1992; Marmorstein and Harrison, 1994; Zhang and Guarente, 1994; Suárez *et al*., 1995; Hellauer *et al*., 1996; Swaminathan *et al*., 1997; King *et al*., 1999). The amino acid side-chains of the first loop of the Zn complex determine basepair recognition, whereas the distance separating and the orientation of the repeats are determined by a linker that lies between the DNA-binding motif *sensu strictu* and a dimerization element (Reece and Ptashne, 1993). At least two proteins fail to conform to this model. These are the NirA and AlcR transcription factors of *Aspergillus nidulans*. NirA is a dimer that recognizes an asymmetric sequence that is not a repeat of a CGG (or any other) motif (Strauss *et al*., 1998). AlcR presents a unique mode of binding. It binds as a monomer, and both mutational and nuclear magnetic resonance (NMR) studies show that residues outside the first basic loop are essential for sequence recognition (Nikolaev *et al*., 1999; Cerdan *et al*., 2000; Cahuzac *et al*., 2001). The FacB activator recognizes two sets of very different sequences, both possibly imperfect inverted repeats (Todd *et al*., 1998), but no details are known about its mode of binding. Thus, this group of proteins shows greater plasticity that was once supposed.

PrnA is necessary for the induction of all the genes involved in proline utilization (Sharma and Arst, 1985; Cazelle *et al*., 1998). These are clustered as shown in Fig. 1A. Proline utilization involves a two-step conversion to glutamate and is identical in *A. nidulans* and *Saccharomyces cerevisiæ*. However, both the organization of the cognate genes and the patterns of regulation differ between the two model ascomycetes. In *A. nidulans*, the *prn* genes are clustered in chromosome VII (Fig. 1). In *S. cerevisiæ*, they are scattered. Proline is both a nitrogen and a carbon source in *A. nidulans*; it is only a nitrogen source in *S. cerevisiæ*. Although gene expression necessitates proline induction in both organisms, the pathway is subject to both nitrogen and carbon metabolite repression in *A. nidulans*, but only to nitrogen metabolite repression in *S. cerevisiæ* (Arst and Cove, 1973; Arst and MacDonald, 1975; Brandriss and Magasanik, 1979; Arst *et al*., 1981; Hull *et al*., 1989; Daugherty *et al*., 1993; Xu *et al*., 1995).

The cluster has been sequenced (Sophianopoulou and Scazzocchio, 1989; Cazelle *et al*., 1998; S. Demais, V. Gavrias, R. Gonzalez and C. Scazzocchio, unpublished data; database entries: *prnA–prnX–prnD*, AJ 223459; *prnC*, AF 252630). *prnB* encodes the specific proline transporter (Arst and MacDonald, 1975; Sophianopoulou

Accepted 14 February, 2002. *For correspondence. E-mail: scazzocchio@igmors.u-psud.fr; Tel. (+33) 1 69 15 63 56; Fax (+33) 1 69 15 78 08. Present addresses: †Instituto de Recursos Naturales y Agrobiología, Consejo Superior de Investigaciones Científicas, Apdo 1052, Sevilla 41080, Spain. ‡Cátedra de Microbiología, Facultad de Química, Universidad de la República, Casilla de Correos 1157, Montevideo, Uruguay.

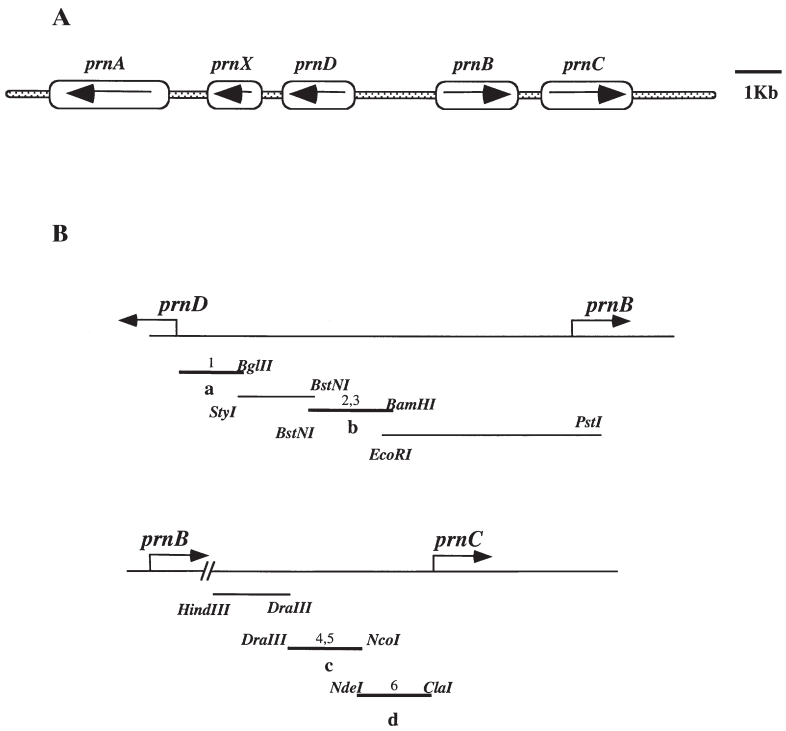

Fig. 1. A. Structure of the *prn* gene cluster in chromosome VII of *A. nidulans* (see text). B. The *prnD–prnB* and *prnB–prnC* intergenic regions and the probes used in the gel retardation experiments. The bold lines correspond to the probes that are retarded by the PrnA full-length and truncated proteins and the lean lines to those that are not. The numbers above the bold lines refer to the binding sites contained in each probe as shown in Fig. 3. For simplicity, not all the overlapping probes used are shown. These completely overlap both intergenic regions.

and Scazzocchio, 1989), *prnD* encodes proline oxidase, and *prnC* encodes Δ1-pyrroline carboxylate dehydrogenase (Arst *et al*., 1981). The *prnX* gene has been detected by transcript analysis and sequencing and has an unknown physiological function (V. Gavrias, S. Demais and C. Scazzocchio, unpublished). The expression of *prnB*, *prnD* and *prnC* depends absolutely on the PrnA protein and on proline induction (Sharma and Arst, 1985; Cazelle *et al*., 1998). Although PrnA is similar in sequence to Put3p, the isofunctional protein of *S. cerevisiæ*, its closest homologue is Thi1p, the regulator of the thiamine biosynthesis pathway of *Schizosaccharomyces pombe* (Cazelle *et al*., 1998). Put3p binding sites are inverted repeats of two CGG triplets separated by 10 bp (Swaminathan *et al*., 1997). Put3p binds to these sequences *in vivo* in the absence of inducer (Axelrod *et al*., 1991). In this study, we determine the binding sites for PrnA and show that induction by proline is absolutely necessary for its *in vivo* binding. The effects of PrnA binding to each of the sites, at the level of both transcription and chromatin structure, are complex. These effects can only be understood in their interaction with the other factors operating in the cluster. These are AreA, mediating nitrogen metabolite repression (Arst and Cove, 1973; Kudla *et al*., 1990; for a review, see Scazzocchio, 2000), and CreA, mediating carbon catabolite repression (Bailey and Arst, 1975; Dowzer and Kelly, 1991; Gonzalez *et al*., 1997). These interactions have been investigated and will be described in detail in a separate publication.

Results

The PrnA binding sites of the prn *cluster*

We searched both the *prnD–prnB* and the *prnB–prnC* intergenic regions for PrnA binding sites. We carried out gel retardation experiments with two different PrnA proteins. A His-tagged 1–163 protein was prepared by expression in *Escherichia coli*. This protein carries the whole PrnA Zn cluster and putative dimerization element (see below). The full-length PrnA(1–818) protein was synthesized in an *in vitro*-coupled transcription–translation system. These proteins were tested for binding to overlapping probes from both intergenic regions in gel retardation experiments. The two proteins gave qualitatively identical results. The results of the gel shifts are summarized in Fig. 1B. Fragments b and c must contain at least two sites of different affinity, as a high-mobility and a low-mobility complex are revealed with increasing amounts of protein. Fragments a and d showed only one complex (not shown). Thus, a minimum of three binding sites exists in each of the intergenic regions.

Methylation interference experiments were carried out to identify PrnA binding sites in each of the retarded probes. These are shown in Fig. 2 (and for probe c in Fig. 7, see below). The collated results of all methylation interference experiments are shown in Fig. 3. These reveal six sites belonging to three classes. These sites (with the exception of sites 1 and 2, see below) contain a CCGG repeat. The CCGG module is itself an inverted repeat and, thus, binding to an inverted repeat or a direct

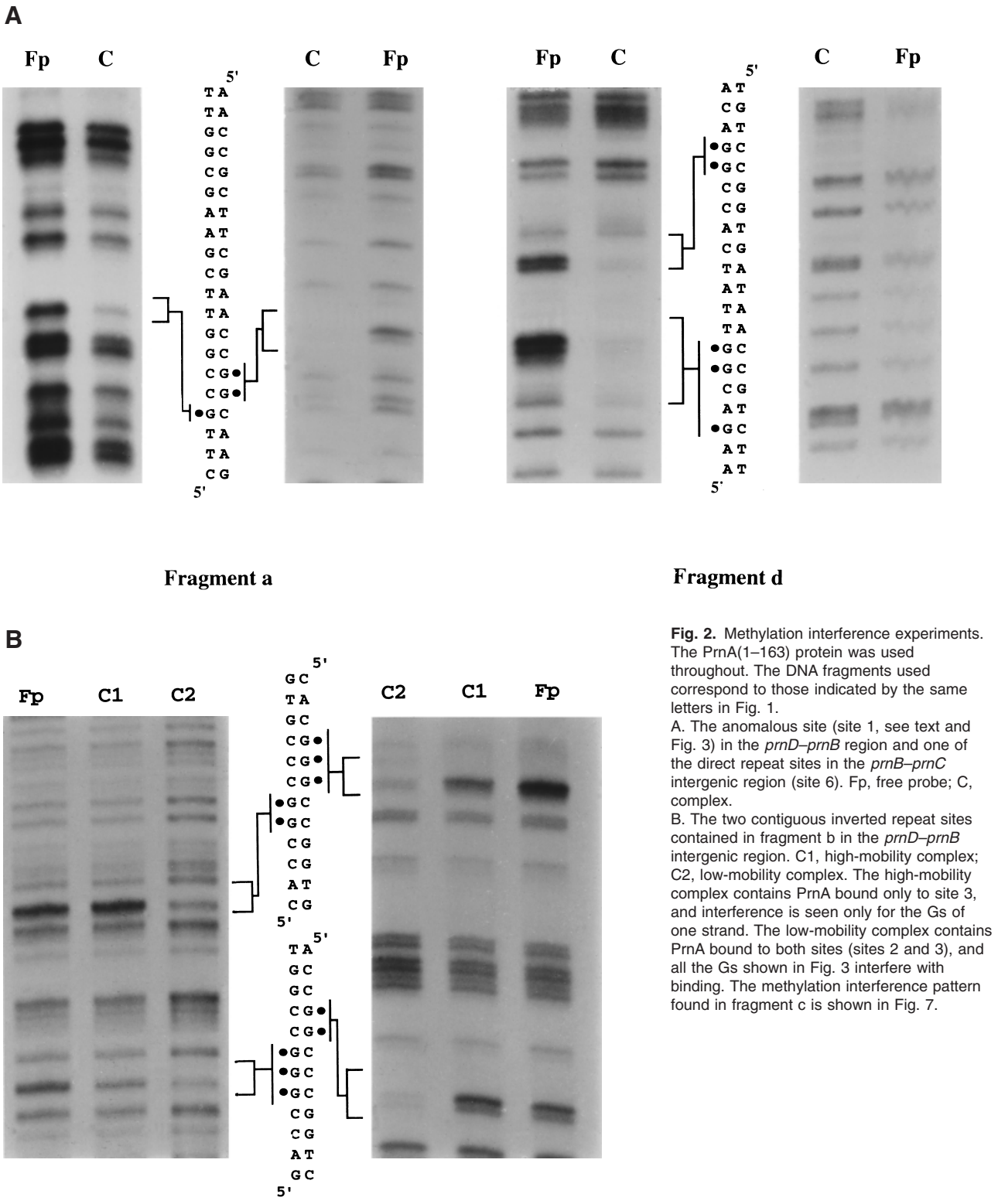

Fig. 2. Methylation interference experiments. The PrnA(1–163) protein was used throughout. The DNA fragments used correspond to those indicated by the same letters in Fig. 1.
A. The anomalous site (site 1, see text and Fig. 3) in the *prnD–prnB* region and one of the direct repeat sites in the *prnB–prnC* intergenic region (site 6). Fp, free probe; C, complex.
B. The two contiguous inverted repeat sites contained in fragment b in the *prnD–prnB* intergenic region. C1, high-mobility complex; C2, low-mobility complex. The high-mobility complex contains PrnA bound only to site 3, and interference is seen only for the Gs of one strand. The low-mobility complex contains PrnA bound to both sites (sites 2 and 3), and all the Gs shown in Fig. 3 interfere with binding. The methylation interference pattern found in fragment c is shown in Fig. 7.

repeat sequence can only be distinguished by the results of the interference experiments. In the inverted repeat mode, we should find symmetrical methylation interference in both strands, whereas in direct repeat mode, we should find methylation interference in one strand. Sites 2, 3 and 4 could be interpreted as either typical CGG inverted repeats separated by 1 bp or extended CCGG inverted repeats, also separated by 1 bp (but note that site

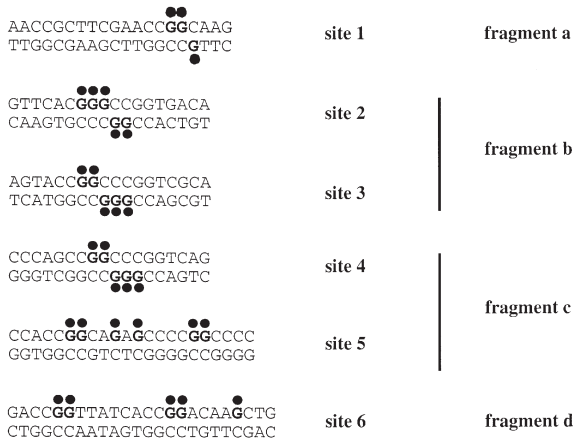

Fig. 3. Summary of all methylation interference experiments. Symbols as in Fig. 2. The data for sites 1, 2, 3 and 6 are shown in Fig. 2; the data for sites 4 and 5 in Fig. 7. The interfering Gs are shown in bold and indicated by ah dot.

2 does not conform exactly to this extended sequence, see next section). In the three sites, the separating basepair is a C/G. The methylation interference of the middle G, which does not belong to the inverted repeat, shows that the site is pseudosymmetrical, as the axis of symmetry passes through the middle basepair. Site 5 can be interpreted as a CCGG direct repeat separated by 7 bp or as a CGG direct repeat separated by 8 bp. Site 6 can be interpreted as a CCGG direct repeat separated by 6 bp or as a CGG direct repeat separated by 7 bp. As methylation of the Gs in the 'bottom' strand does not interfere with binding, we cannot decide as yet between the two interpretations. We should refer to these sites as CCGG direct repeats separated by, respectively, 6 and 7 bp throughout this article. Site 1, which has been confirmed with a number of other PrnA constructs, including a GST-tagged PrnA protein (not shown), is obviously anomalous, and this suggests that binding could occur to a half-site if a G is present in the opposite strand. Whether this is a sufficient condition for this apparent anomalous binding has not been investigated. A CCGG half-site is not sufficient on its own for binding, as many such sites occur in the *prnD–prnB* and *prnB–prnC* intergenic regions and were not revealed by the gel shift or footprinting experiments.

Establishing the PrnA optimal binding sequence with synthetic oligonucleotides

Apparent K_{diss} are shown in Fig. 4 for inverted, direct repeat sites, for the anomalous site (site 1) and for a number of mutant sites. The CCGG-N-CCGG inverted repeat is the preferred sequence. We have carried out a mutational analysis of the inverted repeat sites, as work to be published elsewhere (I. García, D. Gómez and C.

Oligonucleotide name	Oligonucleotide sequence			K_{diss} (nM)
stp1	gtcagta	A**CCGG**C**CCGG**T	actaacg	**2.5**★ (site 3)
stp2	gtcagta	AC**t**GGCCC**a**GT	actaacg	**nb**
stp3	gtcagta	ACC**t**GCC**a**GGT	actaacg	**nb**
stp5	gtcagta	AC**t**GGCCCGGT	actaacg	**50**
stp6	gtcagta	ACC**t**GCCCGGT	actaacg	**75**
stp7	gtcagta	A**t**CGGCCCG**a**T	actaacg	**75**
stp8	gtcagta	ACCG**t**C**a**CGGT	actaacg	**75**
stp9	gtcagta	ACCGG-CCGGT	actaacg	**75**
stp10	gtcagta	ACCGGC**a**CCGGT	actaacg	**13**
stp11	gtcagta	A**CCGG**C**CCGG**T	cgcaata	**2.5**★ (site 3)
stp12	gtcagta	AC-GGGCCGGT	cgcaata	**7.5**★ (site 2)
stp13	gtcagta	ACC**tt**CCCGGT	cgcaata	**50**
stp14	gtcagta	ACCGGC**aa**GGT	cgcaata	**75**
stp15	gtcagta	A**t**CGGCCCGGT	cgcaata	**18**
stp16	gtcagta	**g**CCGGCCCGGT	cgcaata	**2.5**★ (site 4)
stp17	gtcagta	ACCGG**t**CCGGT	cgcaata	**18**
stp18	gtcagta	ACCGGC**c**CCGGT	cgcaata	**13**
stp19	gtcagta	ACCGGC**a**CCGGT	cgcaata	**13**
stp21	gtcagta	ACCGCTTCGAA**CCGG**CA	cgcaata	**7.5**★ (site 1)
stp20	gtcagta	A**CCGG**TTATCA**CCGG**A	caagctg	**13** ★ (site 6)
stp22	gtcagta	ACCGGTTA**c**TCACCGGA	caagctg	**13**

Fig. 4. Apparent K_{diss} of PrnA(1–163) for synthetic oligonucleotides. In bold capitals, the CCGG modules involved in binding (see Figs 2 and 3 and text). Stp2 to stp10 are mutants of stp1; stp12 to stp19 are mutants of stp11. In these mutant oligonucleotides, the mutated, inserted or deleted bases are indicated in bold small letters. K_{diss}sa calculated as described in *Experimental procedures*. nb, no binding detected. Asterisks show sites that are naturally present in the *prn* cluster (not necessarily with the same flanking sequences, which have been kept constant for a more reliable comparison of relative affinities).

Scazzocchio, unpublished), which shows that these sites are the most important in the regulation of gene expression. The optimal distance between the inverted repeats is 1 bp, and mutation of any base in the CGG triplet results in decreased affinity (stp2, 3, 5, 6). Binding is lost when mutations affect both repeats (stp2 and stp3). Stp7 and Stp15 show that the external C/G pairs are important but not essential for binding. Oligonucleotide stp8, in which the most internal G/C pairs of the inverted repeat are mutated behaves identically to stp7. The addition of 1 bp in between the CCGG inverted repeat has a mild effect on affinity (stp10, stp18 and stp19), whereas a deletion (stp9) has a more drastic effect. A T/A pair can substitute for the internal asymmetrical C/G pair, with a clear but not drastic effect on binding (stp17). Thus, the optimal sequence is a CCGG inverted repeat separated by 1 bp and, in line with the interference studies, a C/G is better than a T/A. Site 2 (stp12) can be equated to a 1 bp deletion of site 3. The anomalous site 1 (stp21) has surprisingly high affinity, particularly if compared with stp9 and stp14, which also contain a CCGGC sequence. A G in the other strand and following the CCGG sequence appears to be the minimal requirement for binding in the 'inverted repeat mode' (but not in the 'direct repeat mode'). The direct repeat sites 5 and 6 (CCGG 6/7N-CCGG) have very similar affinities when tested in the context of fragments c and d (not shown). This is confirmed by comparing oligonucleotides stp20 and stp22. These differ, as do sites 5 and 6, only by the addition of an internal C/G basepair.

PrnA binds as a dimer

A hybrid between the 1–126 and the 1–163 protein, co-translated *in vitro*, can clearly be detected in gel shift experiments. This is shown in Fig. 5B. A 1–107 protein

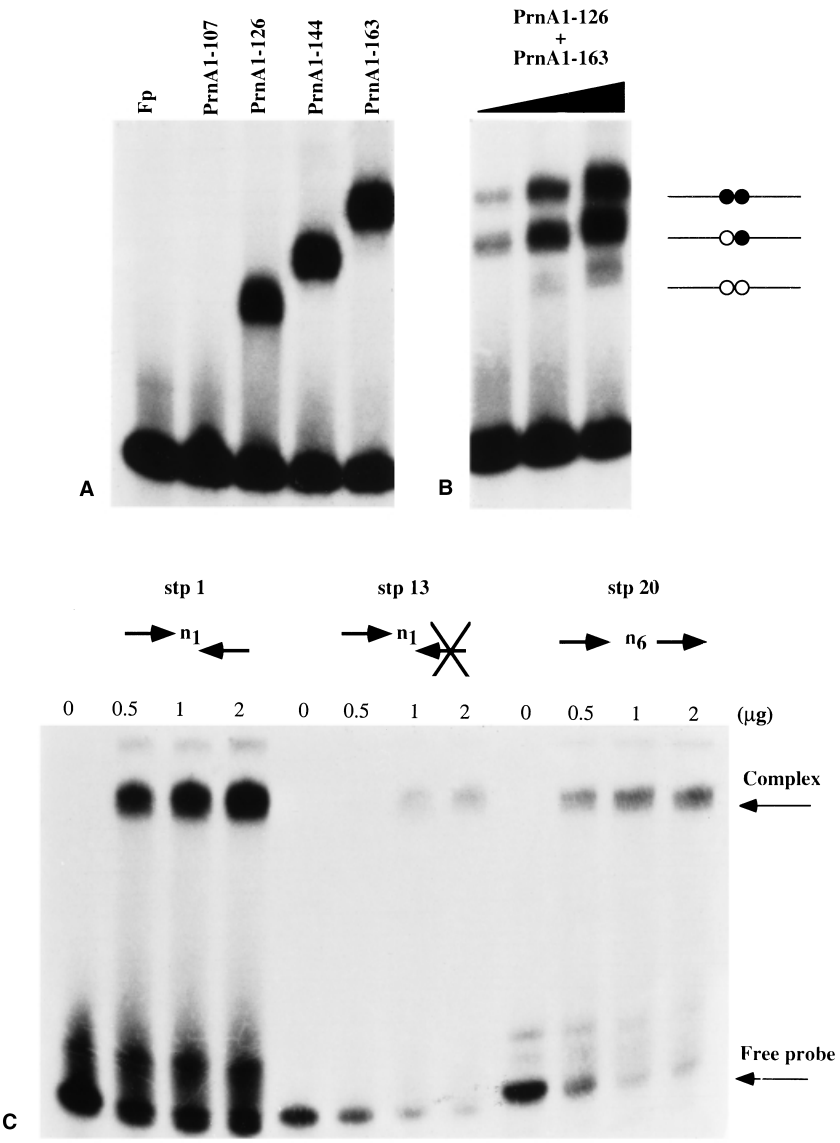

Fig. 5. PrnA binds as dimer.
A. PrnA–DNA complexes. The probe is a *Dra*III–*Nci*I fragment from the *prnB–prnC* intergenic region that contains binding site 4 (Fig. 3). The PrnA proteins of different lengths were produced by an *in vitro* transcription–translation system (see *Experimental procedures*).
B. The same probe is retarded by increasing amounts of PrnA(1–126) and PrnA(1–163) co-transcribed and co-translated *in vitro*. On the right, the two homodimers and the heterodimer are shown; two white circles, PrnA(1–126) homodimer; two black circles, PrnA(1–163) homodimer; one white and one black circle, heterodimer.
C. Gel retardation using the PrnA(1–163) protein with an inverted repeat probe (left), an inverted repeat probe in which one of the repeats is mutated (see text, middle) and a direct repeat probe (right). Stp1, stp13 and stp20 are described in the legend to Fig. 4. The number following the 'n' indicates the number of basepairs between the CCGG motifs (see text). The μgs of recombinant protein used in each experiment are indicated above (final volume 20 μl).

shows no DNA binding, suggesting that this protein lacks the dimerization sequence (Fig. 5A). Figure 5C shows that PrnA always binds as a dimer, even when it binds to half-sites, which can be obtained *in vitro* at high concentrations of protein. An oligonucleotide in which the two interfering Gs of one of the half-sites of the inverted repeat are mutated to As (stp13) shows a radically diminished affinity but gives a complex of exactly the same size as the wild-type probe. This result implies that PrnA(1–163) dimerization is independent of DNA binding. Moreover, a probe of almost identical size containing the direct repeat sequence separated by 6 bp (stp20) gives a complex of identical mobility to that found with the inverted repeat probe. A similar experiment (not shown) demonstrates that the anomalous site 1 is also recognized by a PrnA dimer, as probes stp11 (canonical inverted repeat) and stp21 (anomalous site; for sequences see Fig. 3) migrate identically when complexed with PrnA(1–163). It is thus demonstrated that PrnA binds as a dimer to every one of the recognized sequences.

Chimeric PrnA–UaY proteins

Sequences in the Zn binuclear cluster are believed to determine the nature of the DNA bases recognized, and the linker and dimerization elements the orientation of and

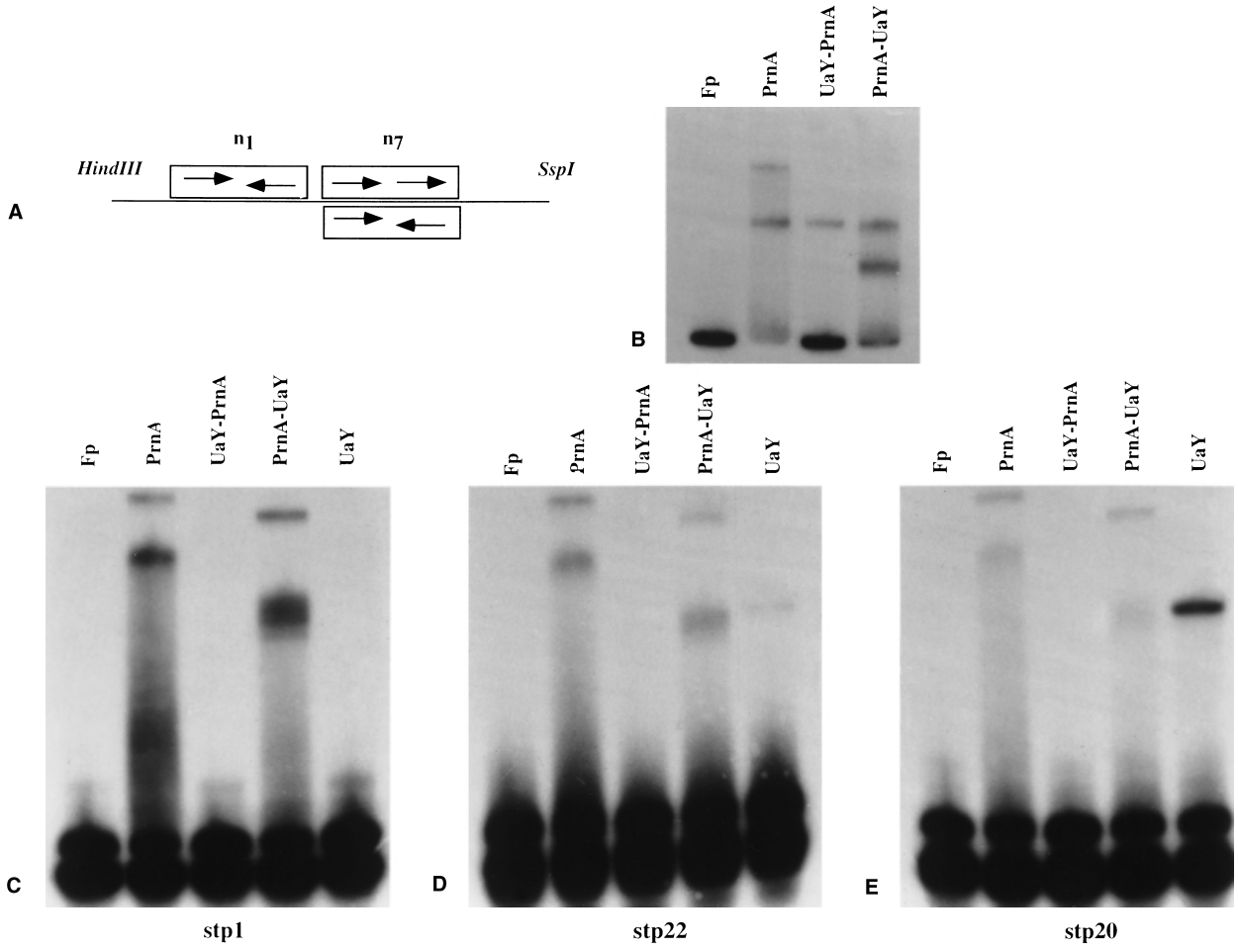

Fig. 6. Binding of PrnA–UaY chimeric proteins.
A. A scheme of the probe (from the *prnB–prnC* intergenic region) used in the experiment in (B). This probe carries two different binding sites for PrnA, an inverted and a direct repeat (sites 4 and 5, respectively, in Fig. 3; shown above the line) and also an aberrant UaY binding site (shown below the line; see text). The number following the 'n' indicates the number of basepairs between the CCGG motifs (see text).
B. The gel retardation obtained with this probe and the PrnA(1–163) and the chimeric proteins UaY–PrnA and PrnA–UaY (described in the text).
C–E. A similar experiment carried out with oligonucleotides and also including retardation with a UaY(1–153) protein. Oligonucleotides named as in the legend to Fig. 4. Fp, free probe. The proteins used are shown above each track. In the experiment at the top, 40 ng of protein μl^{-1} was used (about 2 μM for each protein, which do not have exactly the same molecular weight). In the experiment shown at the bottom, the concentration of all proteins was adjusted to 2.5 μM. Here, a minor low-mobility complex can be seen in some tracks. These probes contain only one binding site, and the minor band results from protein–protein aggregation at the high concentrations used.

the interval between the base repeats (see *Introduction*; these three elements comprise the specific DNA-binding motif). It should thus be possible to predict exactly the DNA sequence recognized by a chimeric protein constructed from the discrete elements of two different DNA-binding motifs. If, for a specific set of chimeric proteins, these predictions are falsified, it can be concluded that this naive model does not apply to at least one of the binding motifs involved. We have thus constructed chimeras involving the DNA-binding domain of PrnA and that of another protein of *A. nidulans*, UaY, which clearly binds according to the 'naive' model (Suárez *et al*., 1995). We can thus determine whether the DNA-binding motif of PrnA also complies with this model.

The UaY protein is a close homologue of Ppr1p. The similarities comprise the Zn cluster, the linker and dimerization elements and sequences outside the DNA-binding domain. Both proteins exclusively recognize CGG inverted repeats separated by 6 bp (Suárez *et al*., 1995). UaY binds as a dimer. The structure of the Prp1p–DNA complex is known. Striking sequence identities of the DNA-binding domains and the patterns of interference (including a strictly symmetric pattern of methylation interference) show that these two proteins bind identically (Suárez *et al*., 1995; G. Cecchetto, C. Scazzocchio and N. Oestreicher, unpublished). By similarity to Ppr1p, the UaY dimerization element must lie from Val-109 to Gly-132. It has been determined experimentally using a λ phage assay (Hu *et al*., 1990; Battaglia *et al*., 1994; Strauss *et al*., 1998) to be contained within Arg-103 and Ala-147 (G. Cecchetto, C. Scazzocchio and N. Oestreicher, unpublished).

We have constructed two chimeric proteins. One contains the UaY binuclear cluster and putative linker and dimerization elements of PrnA (UaY1–96, PrnA59–163). These sequences should be sufficient to ensure dimerization, as PrnA(1–126) is able to dimerize (Fig. 5). The second contains the PrnA binuclear cluster and the linker and dimerization elements of UaY (PrnA1–57, UaY98–153). The 'naive' model predicts that, if the Zn cluster sequences are interchangeable and able to recognize the CGG motifs contained in both UaY and PrnA cognate DNA sequences, the differences in specificity (of the number of bases separating the repeats and their orientation) should depend only on the linker and dimerization elements. We show below that this is not the case.

We have analysed the constructs described above with fragment c (Fig. 1), which contains both inverted repeat and direct repeat PrnA binding sequences. It should be noticed that, overlapping with the CCGG-7N-CCGG direct repeat sequence, there is an aberrant UaY CGG-7N-CCG binding sequence. This sequence, cCGGcagagccCCGg shows about 250 times lower affinity for a UaY(1–153) protein than a physiological sequence, tCGGaggtggC CGa, found in the *uaZ* promoter (not shown). This confirms further the rigidity of the UaY linker and dimerization elements, as inferred from their strong similarity to those of Ppr1p (Suárez *et al*., 1995). Methylation interference to PrnA binding to direct repeat sites is seen only on one strand (Fig. 2), whereas methylation interference to UaY binding is on both strands and strictly symmetrical (Suárez *et al*., 1995). The UaY–PrnA protein has a low affinity for the probe and gives only one complex (Fig. 6B). Methylation interference (not shown) demonstrates that this corresponds to the PrnA inverted repeat. However, experiments with oligonucleotides (see below, Figs 6C–E) demonstrate that this protein has equally low affinity for all sites tested. Thus, the first prediction is falsified: the UaY binuclear cluster sequences cannot replace the cognate PrnA sequences

The PrnA–UaY construct is quite surprising. It results in two complexes in gel retardation experiments with fragment c (Fig. 6B). Methylation interference (Fig. 7) shows that these complexes correspond to the PrnA inverted repeat and the anomalous UaY inverted repeat but not to the CCGG-7N-CCGG direct repeat (this is shown on the right, with the control PrnA protein on the left; methylation interference is seen only on one strand). For the latter sequence, interference to binding of the PrnA–UaY chimera is now seen on both strands, even if is not strictly symmetrical. Everything happens as if the capacity to recognize the UaY inverted repeat and the PrnA direct repeat depended, as expected, on the linker and dimerization elements, whereas the capacity to recognize the PrnA CCGG-N-CCGG inverted repeat resided in the Zn binuclear cluster and/or in sequences amino-terminal to it. We have compared the binding of the PrnA, UaY and the two chimeric proteins to three different oligonucleotides, one containing the optimal CCGG-N-CCGG PrnA binding site and two others, containing CCGG-6N-CCGG and CCGG-7N-CCGG direct repeats (Fig. 6, bottom). The latter two overlap with CGG-6N-CCG and CGG-7N-CCG UaY canonical and aberrant binding sites respectively. The binding of PrnA and the chimeric proteins to a typical UaY site, the one present in the *uaZ* promoter (Suárez *et al*., 1995; G. Cecchetto, C. Scazzocchio and N. Oestreicher, unpublished), was also tested. Only the UaY(1–153) protein binds to this site (not shown; concentrations tested from 2.9 to 9 μM). The UaY–PrnA chimera has low affinity for all the PrnA sites. No binding is seen in Fig. 6C–E (concentration of protein 2 μM) but, varying the concentration of protein between 3 and 14 μM, some binding is seen with the three oligonucleotides (not shown).

The most striking result of these experiments is that the PrnA–UaY chimera binds strongly to CCGG-N-CCGG in the inverted repeat mode (Figs 6B and C and 7). It is, however, clear that the Zn binuclear cluster and

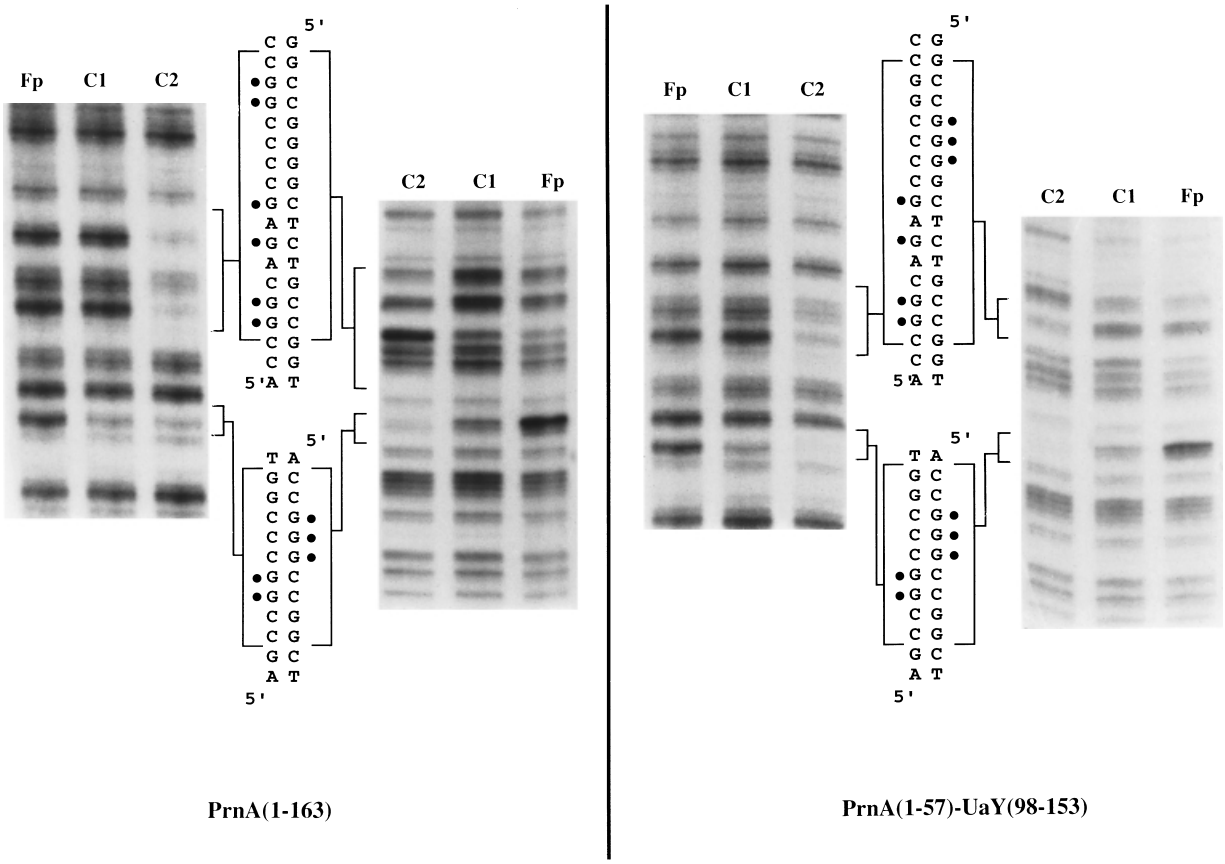

Fig. 7. Methylation interference patterns obtained with the chimeric protein PrnA(1–57)–UaY(98–153) (right). This is compared with the interference pattern obtained with the PrnA(1–163) protein (left). Symbols as in Fig. 2. The probe used is the same as in Fig. 6B. C1 and C2 correspond to the high- and low-mobility complexes seen with this protein in Fig. 6B. For PrnA, the high-mobility complex reveals site 4 (CCGG-N-CCGG inverted repeat), and the low-mobility complex reveals sites 4 and 5 (CCGG-7N-CCGG direct repeat). For the PrnA–UaY chimeric protein, the high-mobility complex reveals site 4, and the low-mobility complex reveals site 4 and a CCGG-7N-CCGG inverted repeat, which contains an aberrant UaY binding site (see text). These overlapping sites are identical to those contained in oligonucleotide stp22. Notice that the interference pattern is different from the symmetrical pattern published previously for a UaY protein (Suárez *et al*., 1995) and also found for all other UaY binding sites characterized with their cognate protein in our laboratory (G. Cecchetto, C. Scazzocchio and N. Oestreicher, unpublished).

sequences amino-terminal to it are not, on their own, sufficient for binding, as a PrnA(1–107) construct fails to bind any PrnA site (shown for site 4 in Fig. 5A). Thus, dimerization is essential for the recognition of the CCGG-N-CCGG inverted repeat, but the specificity of recognition of this configuration does not reside in the linker and dimerization elements but in the PrnA Zn binuclear cluster and/or in sequences amino-terminal to it. The methylation interference results (Fig. 7) are only consistent with the binding of the PrnA–UaY chimera to stp20 and stp22 occurring in an inverted repeat mode, characteristic of UaY. The UaY linker and dimerization elements do not allow the recognition of sequences in the direct repeat mode.

In vivo *detection of PrnA binding*

We have studied the attachment of PrnA to sites 2, 3 and 4 by *in vivo* methylation protection. Work to be published elsewhere shows that these sites are of major physiological importance. These experiments are shown in Fig. 8. In each case, we see clearly the protection of one G, the one that shows strongest interference in the *in vitro* experiments. For site 4, we have carried out protection experiments on both strands, and the symmetrical pattern of binding is confirmed *in vivo*. It is striking that, for the three sites, the protection of this G is completely dependent on induction. In time course experiments, we see protection of site 2 between 5 and 15 min after induction, whereas protection of site 3 is only visible 30 min after the addition of proline (not shown). The expression of the *prn* genes is strongly repressed only by the simultaneous presence of glucose and ammonium (Arst and MacDonald, 1975; Gonzalez *et al*., 1997). Under repressed conditions, PrnA is not detached from its cognate sites (shown for sites 2 and 5). Thus, repression does not act directly on PrnA binding.

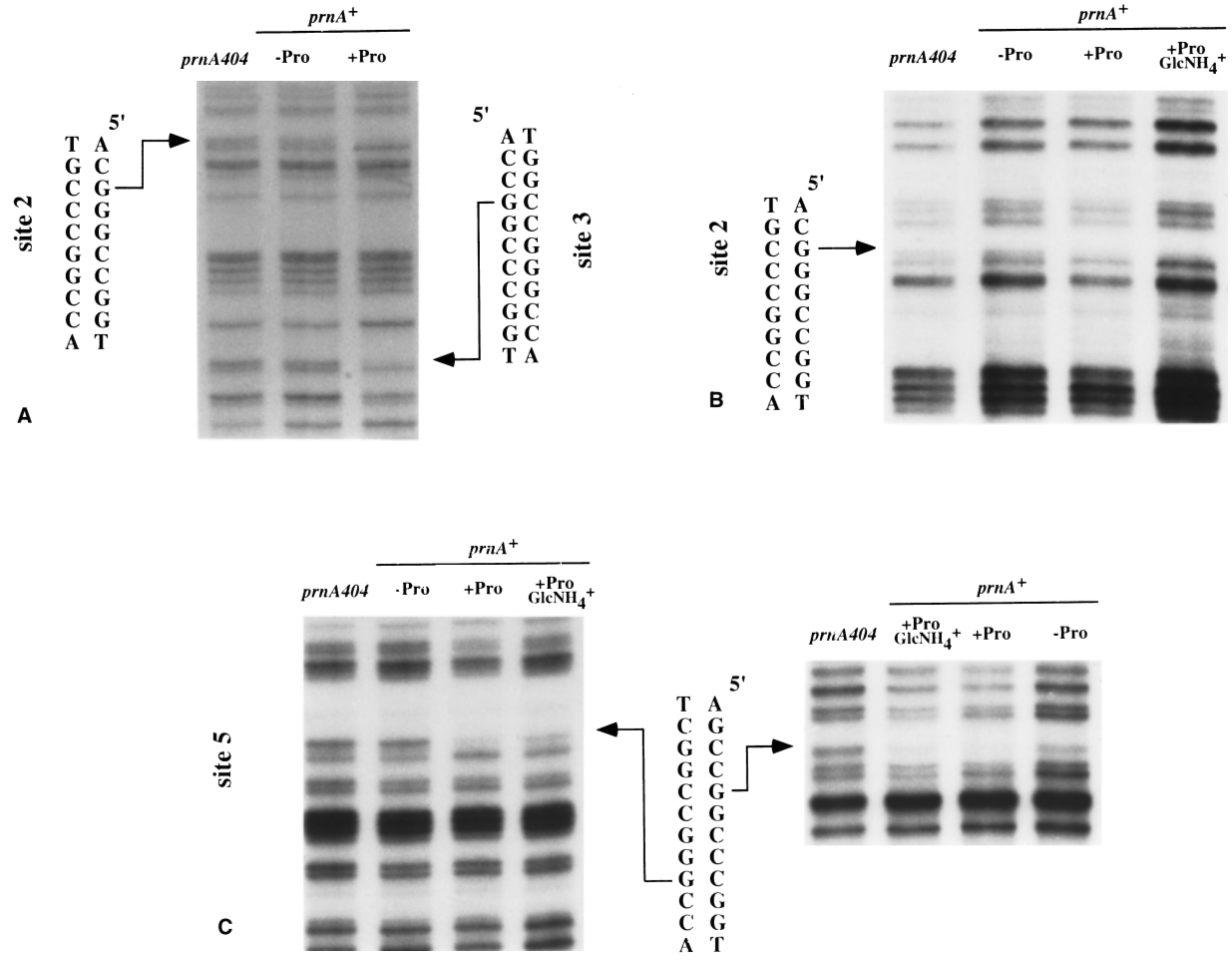

Fig. 8. *In vivo* footprints of sites of the form CCGG-N-CCGG.
A. Methylation protection of sites 2 and 3 in the *prnD–prnB* intergenic region.
B. Methylation protection of site 2 under non-induced, induced and induced/repressed conditions (see below and text).
C. Methylation protection of site *5* (*prnB–prnC* intergenic region), shown on both strands and in all three conditions.
We have included the methylation protection pattern found in a strain carrying a deletion of the PrnA gene (*prnA404*; Pokorska *et al*., 1998) grown under induced conditions. *prnA*$^+$, wild type; *prnA404*, deletion strain; –Pro, non-induced (5 mM urea, 0.1% fructose), +Pro, induced (20 mM proline, 5 mM urea, 0.1% fructose); + ProGlcNH4$^+$, induced in the presence of both glucose and ammonia; inducing–repressing conditions (20 mM proline, 20 mM ammonium D(+)tartrate, 1% glucose).

Discussion

The unique mode of binding of PrnA

Three parameters characterize a DNA binding site containing two repeats that are recognized by a dimeric protein. One is the basepair sequence recognized, the second is the orientation of the repeats (direct, inverted or everted), and the third is the distance between repeats. The standard model is that the Zn cluster is responsible for basepair recognition and that the dimerization and linker elements are responsible for the pattern of recognition (orientation of and distance separating the repeats) (Reece and Ptashne, 1993; Schwabe and Rhodes, 1997).

This model, generalizing the binding pattern of Gal4p-like proteins, is inadequate to account for other existing modes of binding (Strauss *et al*., 1998; King *et al*., 1999; Nikolaev *et al*., 1999; Cerdan *et al*., 2000; Cahuzac *et al*., 2001). Here, we describe a protein that, although binding preferentially to CCGG-N-CCGG inverted repeats, shows striking plasticity. We have not sought all possible binding sequences but, within the *prn* cluster, two other classes of binding sequences can be recognized. One of these, represented by site 1, can be interpreted as an incomplete inverted repeat. On the other hand, sites 5 and 6 are CCGG-6/7N-CCGG direct repeats and necessitate an entirely different mode of binding. As PrnA always binds as a dimer, this implies a striking flexibility of the linker element around a possibly rigid dimerization domain. A 1 bp variation in the distance between the repeats has a drastic effect for UaY or PPR1 (Liang *et al*., 1996) binding, whereas it has no drastic effect on PrnA binding to the

direct or the inverted repeat sequences. This result underscores the flexibility of the PrnA linker domain.

In line with the above, the approximate mapping of the dimerization element places it at a greater distance from the Zn cluster than in any other protein of this class characterized experimentally. The Zn cluster terminates at residue 56. An earlier *in silico* prediction placed the PrnA dimerization element between amino acids 121 and 133 (Pokorska *et al*., 2000). Figure 5 shows that a sufficient dimerization element cannot be contained within the first 107 residues, but must be contained within the first 126 residues. In fact, from residue 108 onwards, there is a clear putative leucine zipper.

The results with chimeric proteins permit the correlation of each of the parameters described above with specific domains of the peptidic sequence. The fact that a UaY–PrnA chimeric protein shows very low affinity to all PrnA sites implies that binding specificity depends crucially on the recognition residues of the Zn cluster (and/or amino-terminal to it). Conversely, the PrnA–UaY chimera does not recognize the *uaZ* UaY canonical binding site. The PrnA Zn cluster recognizes a CCGG module, whereas UaY recognizes a CGG module. The *uaZ* binding site has a TCGG-6N-CCGA structure, whereas the oligomers (stp20 and stp22) used in Fig. 6 have a CCGG-6/7N-CCGG structure. The differences in specificity of the PrnA and UaY Zn clusters are also revealed by the differences in affinity between the UaY and PrnA–UaY chimeric proteins for oligonucleotides stp20 and stp22 in Fig. 6 (bottom). The PrnA–UaY chimeric protein binds about equally to both probes, whereas UaY binds very well to the oligonucleotide containing a canonical UaY site (stp20, 6N) and very badly to the one containing the aberrant site (stp22, 7N). Methylation interference patterns are also different between the two proteins (compare Figs 2 and 7 of this article with Fig. 6 in Suárez *et al*. (1995); see legend to Fig. 7).

What is unique for PrnA is that both the dimerization and the linker element and sequences in the Zn cluster and/or amino-terminal to it determine the PrnA pattern (orientation and distance of repeats; see above) of specific recognition. The role of the former elements is shown by the fact that the PrnA–UaY chimera, which carries the rigid (Suárez *et al*., 1995) UaY (PPR1-like) linker and dimerization elements, does not recognize PrnA direct repeats. The role of the latter elements is strikingly demonstrated by the fact that this same chimeric protein binds strongly to CCGG-N-CCGG in the inverted repeat mode.

The results presented can be rationalized as follows: the dimerization element located between residues 108 and 126 ensures, independently of DNA binding, the formation of a flexible dimer that is able to recognize CCGG-6/7N-CCGG direct repeats. When two inverted repeats occur at a suitable distance (1 bp), a second interaction occurs, which fixes the two Zn clusters in a pseudo-symmetrical configuration around the middle G/C base-pair. This second interaction involves, of necessity, the Zn cluster and/or elements amino-terminal to it. The sequence of the recognition loop of the PrnA Zn cluster uniquely includes acidic residues alternating with the usual basic residues. This may suggest a direct recognition loop–recognition loop interaction. Zn finger motifs can act as elements of protein–protein interaction (Crossley *et al*., 1995; Ravagnani *et al*., 1997; Feng and Marzluf, 1998; reviewed by Mackay and Crossley, 1998). The Zn cluster HAP1 protein recognizes a CGG-6N-CGG direct repeat. In this dimeric protein, Zn cluster interactions dictate the asymmetric mode of binding (Zhang and Guarente, 1996). The structure of the HAP1–DNA complex has revealed interactions between the two Zn clusters and between one of the Zn clusters and the linker element of the second subunit (King *et al*., 1999).

PrnA binding elements compared with its closest homologues

It is interesting to compare PrnA with Put3p, the isofunctional protein of *S. cerevisiæ*, and with Thi1p of *S. pombe*, its closest homologue. All elements needed for Put3p binding lie within 35 residues of the binuclear Zn cluster. For PrnA, there are essential elements beyond 41 residues downstream of the binuclear Zn cluster. No binding studies have been carried out with Thi1p, but computer predictions place its putative dimerization element even further away towards the carboxy-terminus than the one identified experimentally for PrnA (Cazelle *et al*., 1998). There are other similarities between PrnA and Thi1p that are not shared by Put3p. Amino-terminal to the Zn binuclear cluster, both proteins share two basic sequences. These are absent in Put3p. These sequences have been shown to be essential for nuclear targeting of PrnA (Pokorska *et al*., 2000). This does not preclude the possibility that one or both may also be included in the specific determinant of recognition that the present work has shown must be contained in or be amino-terminal to the Zn binuclear cluster.

PrnA in vivo *binding requires the effector molecule proline*

We have shown that the whole PrnA protein, driven by its own weak physiological promoter, is localized in the nucleus even in the absence of inducer. In fact, a striking sublocalization in one distinct intranuclear dot is seen under both induced and non-induced conditions (Pokorska *et al*., 2000). Thus, induction does not act by promoting nuclear internalization but by specifically

eliciting DNA binding. Further work will be necessary to determine whether proline binds directly to PrnA or whether a more complex induction pathway is involved. The dependence of *in vivo* binding upon induction constitutes a functional difference from the Put3p protein of *S. cerevisiæ*. Similarly to Gal4p (Selleck and Majors, 1987), Put3p is always bound to its cognate DNA sites (Axelrod *et al*., 1991). On the other hand, Hap1p, another Zn binuclear cluster protein of *S. cerevisiæ*, only binds with high affinity to its cognate DNA targets *in vitro* in the presence of inducer (Hon *et al*., 1999). In the presence of the inducer, a complex comprising Hap1p and Hsp90 is disrupted, and Hap1p becomes activated and available for DNA binding (Zhang *et al*., 1998; Lee *et al*. 2002). J. Strauss and colleagues have shown recently that the NirA protein of *A. nidulans*, the transcriptional activator of the nitrate assimilation pathway, behaves like PrnA in that *in vivo* binding depends specifically on nitrate induction (Narendja *et al*., 2002). There is, however, an important difference; NirA binding, besides demanding the induction signal nitrate, cannot occur under conditions of nitrogen metabolite repression (ammonium repression; Muro-Pastor *et al*., 1999; Narendja *et al*., 2002), whereas PrnA binding is indifferent to repression and responds only to the specific induction signal.

Experimental procedures

Plasmid constructions

Plasmids carrying *prnA*(1–107), *prnA*(1–126), *prnA*(1–144), *prnA*(1–163) and *prnA*(1–818) were constructed by polymerase chain reaction (PCR) amplification from the entire cDNA of the *prnA* gene. The chimeric proteins PrnA(1–57)–UaY(98–153) and UaY(1–96)–PrnA(59–163) were constructed by ligation of two PCR-amplified DNA fragments coding for suitable peptides. In all cases, an *Nco*I site was introduced at the ATG codon and an *Eco*RI site in the 3′ end. The plasmid carrying *uaY*(1–153) was constructed by PCR amplification of the *uaY* genomic clone (Suárez *et al*., 1995). A *Sca*I site was introduced at the ATG codon and a *Not*I site at the 3′ end. The *prnA* and chimeric DNA fragments were cloned into the *Nco*I site of the pET-22b(+) expression vector (Novagen). The *uaY* fragment was cloned in the *Nde*I site.

Protein purification

The *prnA*(1–163)-, *uaY*(1–153)-, *prnA*(1–57)–*uaY*(98–153)- and *uaY*(1–96)–*prnA*(59–163)-carrying plasmids were introduced into *E. coli* strain BL21-(DE3) pLysS (Invitrogen), and cells were grown at 37°C in the presence of 50 μg ml^{-1} ampicillin until an absorbance of 0.6–0.8 at 600 nm was reached. After 4 h of induction with 1 mM IPTG, the cells were harvested by centrifugation and resuspended in 50 mM sodium phosphate buffer (pH 7.9) containing 0.3 M NaCl, 5 mM β-mercaptoethanol. After sonication, the proteins were purified on a Ni^{2+}/nitricotriacetic acid–agarose column according to the recommendations of the supplier (Qiagen), using a stepwise gradient of imidazole.

In vitro *transcription–translation*

Proteins PrnA(1–107), PrnA(1–126), PrnA(1–144) PrnA(1–163) and PrnA(1–818) were expressed in a transcription–translation system (Promega) according to the recommendations of the supplier by adding 1 μg of each plasmid either alone or in an appropriate mixture to the reaction system. Expression of the proteins was monitored by SDS–PAGE followed by autoradiography.

Electrophoretic mobility shift assays (EMSAs)

Binding assays were performed as described previously (Cubero and Scazzocchio, 1994). Radiolabelled double-stranded oligonucleotides for mobility shift assays were prepared as follows: equimolar amounts of single-stranded oligonucleotides were annealed in annealing buffer (100 mM Tris-HCl, pH 8, 150 mM NaCl) by heating the mixture at 95°C for 2 min and cooling it to 40°C. Then, the mixtures were chilled on ice. The receding ends of double-stranded oligonucleotides were filled in with Sequenase version 2.0 (Amersham) and [α^{32}P]-dCTP as specified by the manufacturer.

Estimation of relative affinities (apparent K_{diss}*)*

The apparent K_{diss} for the PrnA(1–163) protein and a number of oligonucleotides were calculated from gel shift reactions. At equilibrium, the concentration of free protein when half the probe is complexed is equivalent to the apparent K_{diss} (Riggs *et al*., 1970; Choo and Klug, 1993). The probes were used in every assay at a concentration of 5 nM, and protein concentrations varied from 0.1 to 200 nM. Thus, the apparent K_{diss} are given by the concentration of total protein at half saturation of the probe minus 2.5 nM, which is within the experimental error of the total concentration of the protein affording half saturation of the probe for probes showing low affinity. Quantification was performed by Phosphorimager analysis (Molecular Dynamics; IMAGEQUANT).

Methylation interference assays

Methylation assays were performed as described previously (Cubero and Scazzocchio, 1994).

In vivo *footprints*

Strains were grown at 37°C for 7–7.5 h in media containing 0.1% fructose (as sole carbon source) and 5 mM urea (as sole nitrogen source) with appropriate supplements. A *pabaA1* (*para*-aminobenzoic acid-requiring) strain was used as the wild type (*prnA*$^+$). As a null mutant for the *prnA* gene, we used strain *pabaA1 prnA404.* This is a deletion extending from nucleotide 68 to nucleotide 1510 of the published

prnA sequence. The deleted region includes the DNA-binding domain and, moreover, strains carrying this deletion have no *prnA* mRNA (Cazelle *et al*., 1998; 1999). The mycelia were then harvested by filtering through sterile blutex tissue, washing with sterile water and transfer to a supplemented minimal medium without any nitrogen or carbon source. The concentration was adjusted to 25 mg wet weight ml^{-1}, and 18 ml of this suspension was aliquoted into 100 ml Erlenmeyer flasks. For non-induced and induced conditions, urea and fructose were added to final concentrations of 5 mM and 0.1% respectively. For the induced condition, proline was added to a final concentration of 20 mM. For the induced–repressed condition, 1% glucose (repressing carbon source), 20 mM ammonium D(+)tartrate (nitrogen-repressing source) and 20 mM proline (final concentrations) were added simultaneously to the media. Samples were incubated in a shaking bath at 37°C for 30 min. Mycelial suspensions of each sample were treated with dimethyl sulphate (DMS) as described elsewhere (Wolschek *et al*., 1998). DMS-treated samples were processed as described by Mueller and Wold (1989) and Garrity and Wold (1992) as modified for use in filamentous fungi (Wolschek *et al*., 1998).

Acknowledgements

We thank colleagues of this and Joseph Strauss' laboratory for allowing us to quote their data before publication. Gianna Cecchetto thanks Matilde Soubes for constant encouragement and helpful discussion. We thank Stéphane Demais for checking a number of sequences. Dennis Gómez was supported by a predoctoral scholarship from the French Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche, a postdoctoral fellowship from ARC and European Union EUROFUNG contract no. BIO4-CT96-0535. Beatriz Cubero was supported by INRA and European Union contract BIO2CT930. Gianna Cecchetto was supported by CSSIC and Pedeciba Química. This work was supported by the CNRS, the Université Paris-Sud, the Institut Universitaire de France and European Union EUROFUNG contract (BIO4-CT96-0535).

References

Arst, H.N., Jr, and Cove, D.J. (1973) Nitrogen metabolite repression in *Aspergillus nidulans*. *Mol Gen Genet* **126:** 111–141.

Arst, H.N., Jr, and MacDonald, D.W. (1975) A gene cluster in *Aspergillus nidulans* with an internally located *cis*-acting regulatory region. *Nature* **254:** 26–31.

Arst, H.N., Jr, Jones, S.A., and Bailey, C.R. (1981) A method for the selection of deletion mutations in the 1-proline catabolism gene cluster of *Aspergillus nidulans*. *Genet Res* **38:** 171–195.

Axelrod, J.D., Majors, J., and Brandriss, M.C. (1991) Proline-independent binding of PUT3 transcriptional activator protein detected by footprinting *in vivo*. *Mol Cell Biol* **11:** 564–567.

Bailey, C., and Arst, H.N., Jr (1975) Carbon catabolite repression in *Aspergillus nidulans*. *Eur J Biochem* **51:** 573–577.

Battaglia, P.A., Longo, F., Ciotta, C., Del Grosso, M.F., Ambrosini, E., and Giliani, F. (1994) Genetic test to reveal Tat homodimer formation an select Tat homodimer inhibitor. *Biochem Biophys Res Commun* **201:** 795–802.

Brandriss, M.C., and Magasanik, B. (1979) Genetics and physiology of proline utilisation in *Saccharomyces cerevisiæ*: enzyme induction by proline. *J Bacteriol* **140:** 1898–1953.

Cahuzac, B., Cerdan, R., Felenbok, B., and Guittet, E. (2001) Variation along a structural theme. The solution structure of an AlcR–DNA complex sheds light onto the unique tight binding and monomeric DNA-binding of a Zn_2-Cys_6 protein. *Structure* **9:** 827–836.

Carey, M., Kakidani, H., Leatherwood, J., Mostashari, F., and Ptashne, M. (1989) An amino-terminal fragment of GAL4 binds DNA as a dimer. *J Mol Biol* **209:** 423–432.

Cazelle, B., Pokorska, A., Hull, E., Green, P.M., Stanway, G., and Scazzocchio, C. (1998) Sequence, exon-intron organization, transcription and *mutational* analysis of *prnA*, the gene encoding the transcriptional activator of the *prn* gene cluster in *Aspergillus nidulans*. *Mol Microbiol* **28:** 355–370 [Erratum: *Mol Microbiol* **31:** 1283].

Cerdan, R., Cahuzac, B., Felenbok, B., and Guittet, E. (2000) NMR solution structure of AlcR (1–60) provides insight in the unusual DNA binding properties of this zinc binuclear cluster protein. *J Mol Biol* **295:** 729–736.

Choo, Y., and Klug, A. (1993) A role for the linker sequences of the first three zinc fingers of TFIIIA. *Nucleic Acids Res* **21:** 3341–3346.

Crossley, M., Merika, M., and Orkin, S.H. (1995) Self association of the erythroid transcription factor GATA-1 mediated by its zinc finger domains. *Mol Cell Biol* **15:** 2448–2456.

Cubero, B., and Scazzocchio, C. (1994) Two different, adjacent and divergent zinc finger binding sites are necessary for CREA-mediated carbon catabolite repression in the proline gene cluster of *Aspergillus nidulans*. *EMBO J* **13:** 407–415.

Daugherty, J.R., Rai, R., El Berry, H.M., and Cooper, T.G. (1993) Regulatory circuit for responses of nitrogen catabolite expression of the GLN3 and DAL80 proteins and nitrogen catabolite repression in *Saccharomyces cerevisiæ*. *J Bacteriol* **175:** 64–73.

Dowzer, C.E., and Kelly, J.M. (1991) Analysis of the *creA* gene, a regulator of carbon catabolite repression in *Aspergillus nidulans*. *Mol Cell Biol* **11:** 5701–5709.

Feng, B., and Marzluf, G.A. (1998) Interaction between major nitrogen regulatory protein NIT2 and pathway-specific regulatory factor NIT4 is required for their synergistic activation of gene expression in *Neurospora crassa*. *Mol Cell Biol* **18:** 3983–3990.

Garrity, P.A., and Wold, B.J. (1992) Effects of different DNA polymerases in ligation-mediated PCR: enhanced genomic sequencing and *in vivo* footprinting. *Proc Natl Acad Sci USA* **89:** 1021–1025.

Gonzalez, R., Gavrias, V., Gomez, D., Scazzocchio, C., and Cubero, B. (1997) The integration of nitrogen and carbon catabolite repression in *Aspergillus nidulans* requires the GATA factor AreA and an additional positive-acting element, ADA. *EMBO J* **16:** 2937–2944.

Hellauer, K., Rochon, M.H., and Turcotte, B. (1996) A novel DNA binding motif for yeast zinc cluster proteins: the Leu3p and Pdr3p transcriptional activators recognize everted repeats. *Mol Cell Biol* **16:** 6096–6102.

Hon, T., Hach, A., Tamalis, D., Zhu, Y., and Zhang, L. (1999) The yeast heme-responsive transcriptional activator Hap1 is a preexisting dimer in the absence of heme. *J Biol Chem* **274:** 22770–22774.

Hu, J.C., O'Shea, E.K., Kim, P.S., and Sauer, R.T. (1990) Sequence requirements for coiled coils: analysis with 1 repressor-GCN4 leucine zipper fusions. *Science* **250:** 1400–1403.

Hull, E.P., Green, P.M., Arst, H.N., Jr, and Scazzocchio, C. (1989) Cloning and physical characterization of the 1-proline catabolism gene cluster of *Aspergillus nidulans*. *Mol Microbiol* **3:** 553–559.

King, D.A., Zhang, L., Guarente, L., and Marmorstein, R. (1999) Structure of a HAP1–DNA complex reveals dramatically asymmetric DNA binding by a homodimeric protein. *Nature Struct Biol* **6:** 64–71.

Kudla, B., Caddick, M.X., Langdon, T., Martinez-Rossi, N.M., Bennett, C.F., Sibley, S., *et al.* (1990) The regulatory gene *areA* mediating nitrogen metabolite repression in *Aspergillus nidulans*. Mutations affecting specificity of gene activation alter a loop residue of a putative zinc finger. *EMBO J* **9:** 1355–1364.

Lee, H.C., Hon. T., and Zhang, L. (2002) The molecular chaperone Hsp90 mediates heme transcriptional activation of the yeast transcriptional activator Hap1. *J Biol Chem* **277:** 7430–7437.

Liang, S.D., Marmorstein, R., Harrison, S.C., and Ptashne, M. (1996) DNA sequence preferences of GAL4 and PPR1: how a subset of Zn2Cys2 binuclear cluster proteins recognizes DNA. *Mol Cell Biol* **16:** 3773–3780.

Mackay, J.P., and Crossley, M. (1998) Zinc fingers are sticking together. *Trends Biochem Sci* **23:** 1–4.

Marmorstein, R., and Harrison, S.C. (1994) Crystal structure of a PPR1–DNA complex: DNA recognition by proteins containing a Zn_2Cys_6 binuclear cluster. *Genes Dev* **8:** 2504–2512.

Marmorstein, R., Carey, M., Ptashne, M., and Harrison, S.C. (1992) DNA recognition by GAL4: structure of a protein–DNA complex. *Nature* **356:** 408–414.

Mueller, P.R., and Wold, B. (1989) *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* **246:** 780–786.

Muro-Pastor, M.I., Gonzalez, R., Strauss, J., Narendja, F., and Scazzocchio, C. (1999) The GATA factor AreA is essential for chromatin remodelling in a eukaryotic bidirectional promoter. *EMBO J* **18:** 1584–1597.

Narendja, F., Goller, S.P., Wolscheck, M.F., and Strauss, J. (2002) Nitrate and the GATA factor AreA are necessary for *in vivo* binding of NirA, the pathway-specific transcriptional activator of *Aspergillus nidulans*. *Mol Microbiol* **44:** 573–583.

Nikolaev, I., Cochet, M.F., Lenouvel, F., and Felenbok, B. (1999) A single amino acid, outside the AlcR zinc binuclear cluster, is involved in DNA binding and in transcriptional regulation of the *alc* genes in *Aspergillus nidulans*. *Mol Microbiol* **31:** 1115–1124.

Pokorska, A., Drevet, C., and Scazzocchio, C. (2000) The analysis of the transcriptional activator PrnA reveals a tripartite nuclear localisation sequence. *J Mol Biol* **298:** 585–596.

Ravagnani, A., Gorfinkiel, L., Langdon, T., Diallinas, G., Adjadj, E., Demais, S., *et al.* (1997) Subtle hydrophobic interactions between the seventh residue of the zinc finger loop and the first base of an HGATAR sequence determine promoter-specific recognition by the *Aspergillus nidulans* GATA factor AreA. *EMBO J* **16:** 3974–3986.

Reece, R.J., and Ptashne, M. (1993) Determinants of binding-site specificity among yeast C6 zinc cluster proteins. *Science* **261:** 909–911.

Riggs, A.A., Suzuki, H., and Bourgois, S. (1970) *lac* repressor–operator interaction. I. Equilibrium studies. *J Mol Biol* **48:** 67–83.

Scazzocchio, C. (2000) The fungal GATA factors. *Curr Opin Microbiol* **3:** 126–131.

Schwabe, J.W.R., and Rhodes, D. (1997) Linkers made to measure. *Nature Struct Biol* **4:** 680–683.

Selleck, S.B., and Majors, J.E. (1987) *In vivo* DNA-binding properties of a yeast transcription activator protein. *Mol Cell Biol* **7:** 3260–3267.

Sharma, K.K., and Arst, H.N., Jr (1985) The product of the regulatory gene of the proline catabolism gene cluster of *Aspergillus nidulans* is a positive-acting protein. *Curr Genet* **9:** 299–304.

Sophianopoulou, V., and Scazzocchio, C. (1989) The proline transport protein of *Aspergillus nidulans* is very similar to amino acid transporters of *Saccharomyces cerevisiae*. *Mol Microbiol* **3:** 705–714.

Strauss, J., Muro-Pastor, M.I., and Scazzocchio, C. (1998) The regulator of nitrate assimilation in ascomycetes is a dimer which binds a nonrepeated, asymmetrical. *Mol Cell Biol* **18:** 1339–1348.

Suárez, T., de Queiroz, M.V., Oestreicher, N., and Scazzocchio, C. (1995) The sequence and binding specificity of UaY, the specific regulator of the purine utilization pathway in *Aspergillus nidulans*, suggest an evolutionary relationship with the PPR1 protein of *Saccharomyces cerevisiae*. *EMBO J* **14:** 1453–1467.

Swaminathan, K., Flynn, P., Reece, R.J., and Marmorstein, R. (1997) Crystal structure of a PUT3-DNA complex reveals a novel mechanism for DNA recognition by a protein containing a Zn_2Cys_6 binuclear cluster. *Nature Struct Biol* **4:** 751–759.

Todd, R.B., Andrianopoulos, A., Davis, M.A., and Hynes, M.J. (1998) FacB, the *Aspergillus nidulans* activator of acetate utilization genes, binds dissimilar DNA sequences. *EMBO J* **17:** 2042–2054.

Wolschek, M.F., Narendja, F., Karlseder, J., Kubicek, C.P., Scazzocchio, C., and Strauss, J. (1998) *In situ* detection of protein–DNA interactions in filamentous fungi by *in vivo* footprinting. *Nucleic Acids Res* **26:** 3862–3864.

Xu, S., Falvey, D.A., and Brandriss, M.C. (1995) Roles of URE2 and GLN3 in the proline utilization pathway in *Saccharomyces cerevisiæ*. *Mol Cell Biol* **15:** 2321–2330.

Zhang, L., and Guarente, L. (1994) The yeast activator HAP1 – a GAL4 family member – binds DNA in a directly repeated orientation. *Genes Dev* **8:** 2110–2119.

Zhang, L., and Guarente, L. (1996) The C6 zinc cluster dictates asymmetric binding by HAP1. *EMBO J* **15:** 4676–4681.

Zhang, L., Hach, H., and Wang, C. (1998) Molecular mechanism governing heme signalling in yeast: a higher-order complex mediates heme regulation of transcriptional activator HAP1. *Mol Cell Biol* **18:** 3819–3828.