© 1996 Oxford University Press

Nucleic Acids Research, 1996, Vol. 24, No. 8 **1453–1459**

Mutations in target DNA elements of yeast HAP1 modulate its transcriptional activity without affecting DNA binding

Nhuan Ha, Karen Hellauer and Bernard Turcotte*

Department of Medicine, McGill University, Royal Victoria Hospital, 687 Pine Avenue West, Montréal, Québec H3A 1A1, Canada

Received January 17, 1996; Accepted February 13, 1996

ABSTRACT

The yeast zinc cluster protein HAP1, a member of the GAL4 family, is a transcriptional activator that binds as a homodimer to target DNA sequences. These targets include the upstream activating sequences of the *CYC1* and *CYC7* genes, which have no obvious sequence similarity. Even though both sites have the same affinity for HAP1, activation differs at these two sites, even when the sequences are placed in an identical promoter context. In addition, mutants of HAP1 that can bind to both sites but are specifically transcriptionally inactive at *CYC7* have been previously isolated. In order to identify nucleotides that are responsible for this differential activity, we have performed random and site-directed mutagenesis of these target sites and assayed their binding to HAP1 *in vitro* and their activity *in vivo* in reporter plasmids. Our results show that HAP1 binding sites are degenerate forms of the direct repeat CGG N_3 TA N CGG N_3 TA. Moreover, we show that activity of HAP1 mutants defective for activation of the *CYC7* gene is restored by specific mutations in the *CYC7* binding site. Conversely, other mutations of the target sites prevent activation by HAP1, without interfering with DNA binding. The results suggest that the sequence of the target sites influences the conformation and, hence, the activity of DNA-bound HAP1.

INTRODUCTION

Many fungal transcriptional regulators contain a zinc finger known as a binuclear zinc cluster. The consensus sequence of the cluster is $CX_2CX_6CX_{5-9}CX_2CX_{6-8}C$, where the cysteines bind two zinc atoms which coordinate folding of the domain. A well-characterized member of this family is GAL4, which binds as a homodimer to palindromic DNA sequences (CGG N_{11} CCG), with each zinc cluster recognizing a CGG triplet, as shown by X-ray crystallography (1 and references therein). The zinc cluster domain is followed by a short linker region and a dimerization domain. While the spacing between the CGG triplets is 11 base pairs (bp) for GAL4, other zinc cluster proteins also recognize the same palindromic triplets, but with a different spacing. For example, the PUT3 and PPR1 activators bind CGG triplets spaced by 10 and 6 bp respectively (2–3). Moreover, the geometry of the zinc cluster controls binding specificity, as changing the spacing of the GAL4 site from 11 to 10 bp greatly reduces GAL4 binding *in vitro* (4). Construction of GAL4, PUT3 and PPR1 chimeric proteins has shown that the linker region is responsible for binding to a site of a given spacing (5–7).

HAP1, another member of this family of zinc cluster proteins (8–10), possesses an acidic activation domain at its C-terminus (residues 1307–1483) and an N-terminal DNA binding domain (residues 1–174) (10,11). The zinc cluster of HAP1 (residues 64–93) is homologous to that of GAL4 (43% identity) and the coiled coil sequence of the dimerization domain is similar to the domain identified in GAL4 (11). The middle region of HAP1 is required for regulation of DNA binding through interaction with heme. In the absence of heme HAP1 is present in a high molecular weight complex (12,13). Binding of heme to HAP1 apparently stimulates dissociation of HAP1 from the complex and subsequent homodimerization, which is a prerequisite for specific DNA binding.

Upstream activating sequences (UASs) for HAP1 (see Fig. 2 for DNA sequences) have been identified in the *CYC1* and *CYC7* genes, which encode isoforms of cytochrome c (14–16). In addition, other binding sites for HAP1 have been identified in the *CTT1* (catalase T), the *CYB2* (cytochrome b_2) and the *CYT1* (cytochrome c_1) genes (17–19). These sites share little homology. This is particularly evident when comparing the UAS of *CYC1* with that of *CYC7*, which has led to the suggestion that HAP1 binds to unrelated DNA sequences (14).

Dissimilarity between these sites results in different levels of HAP1-dependent activation. Activity at the *CYC1* promoter is higher than at *CYC7*, even if the UASs are placed in a similar promoter context. This different transcriptional activity is not a result of different affinity of HAP1 for *CYC1* and *CYC7* (20). Furthermore, a single amino acid change in HAP1 of Ser63 to Arg, immediately N-terminal of the first cysteine of the zinc cluster, prevents binding to the UAS of *CYC1* and results in greatly increased activation (10- to 100-fold) at *CYC7*, even though the affinity of the mutant HAP1-18 for the UAS of *CYC7*

* To whom correspondence should be addressed

is similar to wild-type HAP1 (20). Other HAP1 mutants, termed positive control (PC) mutants, with mutations that flank the zinc cluster show wild-type binding and activation at *CYC1* but are transcriptionally inactive at *CYC7*, even though they show wild-type or increased binding at that site (21).

We wished to learn more about the nucleotides responsible for this differential activity of HAP1 (and its mutants) at *CYC1* and *CYC7*. We first performed saturation mutagenesis of the *CYC1* and *CYC7* UASs (as well as site-directed mutagenesis at other UASs of HAP1) to identify key nucleotides responsible for binding of HAP1. Our results show that all UASs of HAP1 are imperfect versions of the direct repeat CGG N_3 TA N CGG N_3 TA. Moreover, at a similar position the *CYC7* site has two CGC triplets instead of the two CGGs. Changing these triplets to CGG restores the activity of HAP1 PC mutants, while the activity of HAP1-18 is greatly decreased.

MATERIALS AND METHODS

Strain and media

Strain THl (Mat**a** *ura3-52 his4-519 ade1-100 leu2-2 Δhap1::hisG*) is a derivative of LPY22 (21). Most of the HAP1 coding region has been deleted in that strain. Rich (YPD) and synthetic (SD) media were prepared as described (25).

Plasmids

HAP1 expression vectors. A *HIS4* version of SD5-HAP1 (20) was constructed by deleting the *URA3* gene of SD5-HAP1 by cutting with *Stu*I, adding *Not*I linkers and ligating a *Not*I fragment (containing the *HIS4* gene) from CYC7-5lacZ-HIS4 (21). SD5-HAP1-PC1-HIS, SD5-HAP1-PC2-HIS and SD5-HAP1-18-HIS were constructed by introducing *Dra*III fragments (containing the HAP1 mutations) from *URA3* marked expression vectors (21) into *Dra*III-cut SD5-HAP1-HIS4. HAP1 expression vector deleted of its activation domain (SD5-HAP1ΔKp-HIS) was constructed by linearizing SD5-HAP1-HIS4 with *Kpn*I, treating with T4 DNA polymerase and inserting a *Xba*I linker (New England Biolabs) containing nonsense codons in the three reading frames.

Reporter plasmid. Plasmid pLG178-M was constructed by destroying the unique *Mun*I site in the 2µ origin of replication of pSLFΔ178K (26) by cutting with *Mun*I, filling in with Klenow fragment and ligating. pLG178-M was linearized with *Xho*I and an oligonucleotide (5′-TCGAGAGATCTAAAAAACAATTGC-3′) and its complement were inserted at that site to give p178MB. This plasmid has unique *Mun*I and *Bgl*II sites (flanked by *Xho*I sites) in front of a minimal *CYC1* promoter driving *lacZ* transcription (see Fig. 1).

HAP1 UAS mutants. Spiked oligonucleotides (level of contamination 3% for each of the three other nucleotides) containing the UAS of *CYC1* (5′-CGCGGATCCTCATCGTCCGTAAACCCCGGCC-ACTGTAGGAATTCGGA-3′) or the UAS of *CYC7* (5′-CGCGG-ATCCGCTAATAGCGATAATAGCGAGGGCTGTAGGAATTC GGA-3′) were synthesized. They were then made double-stranded by hybridizing them with a second oligonucleotide (5′-TCCGAAT-TCCTACAG-3′) and filling in with Klenow fragment. The DNA was then cut with *Eco*RI and *Bam*HI, gel purified on a 4% polyacrylamide gel and subcloned into p178MB cut at the compatible sites with *Mun*I and *Bgl*II. A similar strategy, or PCR, was used to introduce specific nucleotide alterations, except that the oligonucleotides contained a *Bgl*II site instead of a *Bam*HI site. Mutations were identified by sequencing the double-stranded DNA using a kit from Pharmacia. UASs of *CTT1* and *CYB2* (and mutants) were generated as described above with oligonucleotides having the DNA sequences shown in Figure 2.

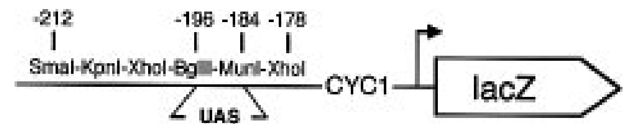

Figure 1. Schematic representation of the plasmid reporter (p178MB) used in this study. Unique sites located upstream of a minimal (UAS-less) *CYC1* promoter are shown. The arrow indicates the major transcriptional initiation site of the promoter. Single copies of UASs were inserted upstream of the minimal *CYC1* promoter using the *Bgl*II and *Mun*I restriction sites (see Materials and Methods).

```
CYC1              TGGC CGG GGTTTA CGG ACGATGA
CYC7              CCCT CGC TATTAT CGC TATTAGC
CTT1              GGAA TGG AGATAA CGG AGGTTCT
CYB2              GGCA AGG AGATAT CGG CAGGCTT
CYT1              CCGC CGG AAATAC CGG CCGCCCA
CYT1(reverse)     CGGC CGG TATTTC CGG CGGCCAA
'OPTIMAL'              CGG NNNTAN CGG NNNTA
```

Figure 2. Alignment of the known UASs of HAP1 (14–19). Two possible orientations for the UAS of *CYT1* are shown. Brackets show the nucleotides protected from DNase I by HAP1. Bold characters indicate nucleotides that match the CGG triplets of the 'optimal' sequence. Underlined characters indicate nucleotides that match the TA repeats of the 'optimal' sequence.

β-Galactosidase assays

Cells were grown to saturation in YEP containing 2% raffinose. They were then diluted into minimal medium containing 1% glucose and 1% galactose and grown for ~12–18 h (OD_{600} 0.6–1.0) before assaying for β-galactosidase activity. β-Galactosidase assays were performed with permeabilized cells (27).

Preparation of extracts and DNA binding assays

Extracts were prepared and DNA binding assays were carried out as described (21,28) using a HAP1 expression vector under the control of UAS_{GAL} (SD5-HAP1; 20). Probes were generated by PCR amplification using the reporter plasmids as templates and purified with G50 spin columns. Quantitative data were obtained by measuring the amount of radioactivity present in the retarded bands using a phosphorimager (Fuji).

RESULTS

Strategy used to generate and characterize HAP1 UAS mutants

Reporter plasmids with mutated *CYC1* and *CYC7* UASs were generated using 'spiked' oligonucleotides (level of contamination 3% of each of the three other nucleotides), flanked by appropriate restriction sites (see Materials and Methods).

Table 1. *In vitro* binding and *in vivo* activity of the UAS of *CYC7* and mutants

MUTANT		IN VIVO ACTIVITY (β-GAL UNITS)	IN VITRO BINDING (% WILD TYPE)
1-WT	TGGC CGG GGTTTA CGG ACGATGA	6.0	100
1-1	TGGG CGG GGGTTA TGG ACGATGA	0.6	1
1-2	TGGC CAC GGTTTA CGG ACGCTGA	0.3	5
1-3	TTGC CGG GGTTTG CGG TCTATAA	0.4	120
1-5	AGGC CGG GGTTTA CGG ACGATTA	2.4	20
1-6	TGGC CGC GGTATA CGG ACCATGA	0.5	5
1-7	TGGC CGG GGTATA CAT ACGATAA	0.3	5
1-8	TGGC CGG GGTTTA CGG ACGAAGA	17.5	120
1-9	TGGC CGA GGTTTA CGG ACGATGA	0.6	20
1-10	TGAC CGG GGTTTA CGG ACGATGA	2.7	115
1-14	TGGC CGG GGTTTA CTG ACGATGA	0.3	15
1-16	TAGT CGG AGTTTA CGG ACGAAGA	17.0	190
1-18	TGGA CTG GGTATA CGG TAGAAGA	0.3	5
1-22	TCGC GGG GGTATA CGG ACGATTA	0.5	10
1-26	TGGC CGG GGGTTA CGG ACGATGA	0.8	55
1-28	TGGT GGG GGTTTA CGG ACGATGA	0.5	30
1-32	AGGC CGG GGTTTA CGG ACGGTTA	3.4	125
1-40	TGTC CGG GGTTTA CGG ACGATGA	2.4	130
1-41	GGGC AGC GGTTTA CGG ACGATGA	0.1	15
1-42	TGTC CGG GGTTTA AGG AAGATGA	0.2	15
1-43	TGAT CGG GGTTTA CGG ACGATGA	2.0	120

Nucleotide changes are underlined and in bold characters. Binding of HAP1 to mutant UASs is given relative to the wild-type UAS of *CYC1* (100%).

Mutagenized UASs were then subcloned into a reporter plasmid containing a minimal *CYC1* promoter driving *lacZ* transcription (Fig. 1). Mutations in the UASs were sequenced and the activity of the reporters was determined *in vivo* by transforming them, along with a HAP1 expression vector, into a hap1$^-$ strain. In the presence of HAP1, activity at *CYC1* and *CYC7* was 6 and 2.7 U respectively (Tables 1 and 2). Moreover, when a HAP1 mutant deleted of its activation domain was used (HAP1ΔKp) only background activity was measured (0.1–0.3 U β-galactosidase activity; data not shown), even though this HAP1 mutant showed wild-type DNA binding *in vitro* (21). This indicates that activity of these reporters was dependent on the activation domain of HAP1. Mutants were also tested for their *in vitro* DNA binding by the electrophoretic mobility shift assay (EMSA) using an extract prepared from a strain that contained a HAP1 overexpression vector under the control of UAS_{GAL4}.

Mutations that disrupt activity of HAP1 at *CYC1*

Many mutants generated by random mutagenesis of the *CYC1* UAS displayed reduced binding and *in* vivo activity, like the single point mutants 1-9 and 1-14 (Table 1). Activation from these sites is 10- to 20-fold lower than from the wild-type UAS of *CYC1*. Strikingly, these mutations are located in CGG triplets, motifs recognized by the GAL4, PUT3 and PPR1 activators. Moreover, all other mutants with alterations in either triplet showed reduced *in vitro* binding and *in vivo* activation. For example, mutant 1-2, with the first CGG triplet changed to CAC, and mutant 1-28, with a change to GGG, showed reduced *in vitro* binding and *in vivo* activation. Similar results were observed with mutants 1-1, 1-2, 1-6, 1-7, 1-18, 1-22, 1-41 and 1-42, containing nucleotide alterations in either CGG triplet (Table 1). These results suggest that two CGG triplets are important for binding of HAP1 to the UAS of *CYC1*.

Interestingly, another class of mutants was isolated with mutations outside the CGG triplets. In the case of mutant 1-3, which carries five mutations, binding was not affected, but activation was 15 times lower than wild-type *CYC1*. Similar results were seen with mutant 1-26. The properties observed at this UAS are analogous to those seen for HAP1 positive control mutants at the wild-type UAS of *CYC7*, i.e. wild-type *in vitro* binding but no transcriptional activation (21).

Other mutants (1-10, 1-32, 1-40 and 1-43) with changes outside the triplets showed wild-type binding properties, as well as significant β-galactosidase activity, suggesting that many nucleotides outside the CGG triplets are not important for binding of HAP1.

Mutations that enhance activity of HAP1 at *CYC1*

Two mutants (1-8 and 1-16; Table 1) of the UAS of *CYC1* showed increased transcriptional activity (3-fold). Mutant 1-8 carries a single nucleotide change 5 nt downstream of the second CGG triplet and mutant 1-16 has 4 nt alterations, including that found in 1-8. In order to rule out the possibility that these UASs could be bound by an activator other than HAP1 we measured their activity with a HAP1 mutant that is transcriptionally inactive (SD5-HAP1ΔKp-HIS; see Materials and Methods). Activity of the mutants was reduced to background levels (0.2 U β-galactosidase activity; data not shown), indicating that their activity was dependent on HAP1.

Mutations at the UAS of *CYC7*

The UAS of *CYC7* has two CGC triplets instead of the two CGGs found at a similar position in *CYC1*. Deletion of one G in one triplet (mutant 7-25) had drastic effects on HAP1 activity (8-fold reduction in activity; Table 2). Some other mutants with changes in these triplets also had decreased binding *in vitro* and activation *in vivo*. For example, mutant 7-20, with a change of the first CGC triplet to CTC, had 6-fold less β-galactosidase activity (Table 2). Similarly, 7-30, which had reduced activity, carries a mutation in the second CGC triplet (along with two other changes). Interestingly, changing one CGC to CGG (7-24) increased the *in vivo* activity ~4-fold. A single nucleotide change between the two CGC triplets (7-5) reduced *in vitro* binding ~10-fold. Many mutations only minimally affected binding and activation at *CYC7*. For instance, mutant 7-1 with 3 nt changes before the first CGC triplet, had wild-type activity. Similar results were obtained with mutants 7-2, 7-3, 7-8 and 7-34. All of these alterations fall outside the CGC repeats.

Affinity of mutants for HAP1

We then tested the affinity of various key mutants by competition assays, as shown in Figure 3. As expected from previous results (20), the UASs of *CYC1* and *CYC7* had a similar affinity for HAP1. Higher *in vivo* activity of mutants 1-8 and 7-24 was correlated with an increased affinity of HAP1 for 7-24 and, to a lesser extent, for 1-8, as compared with the UAS of *CYC1*. Conversely, 1-9, which had reduced *in vitro* binding, did not compete with the wild-type probe *CYC1*, even at a 200 molar excess (Fig. 3). Similarly, 7-5 showed reduced activity and affinity. On the other hand, mutants 1-3, 1-26 and 7I (Table 3) showed increased affinity for HAP1 as compared with the wild-type UAS of *CYC1*, but were transcriptionally inactive.

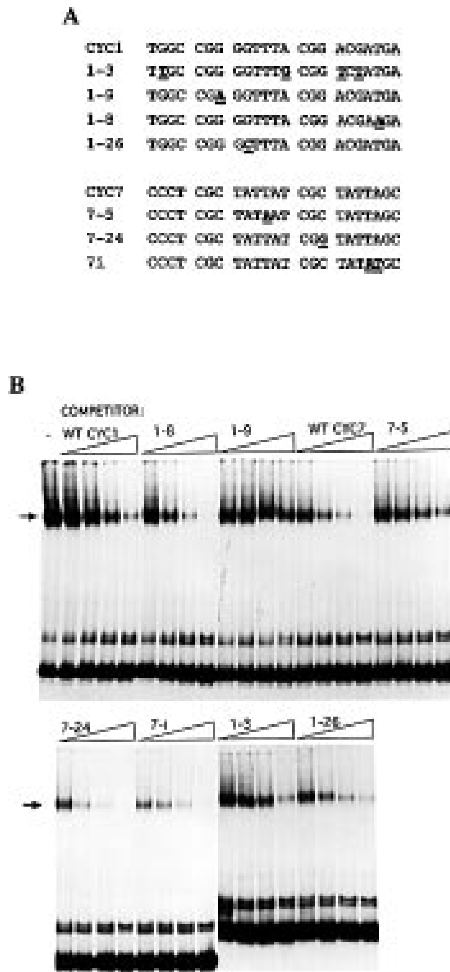

Figure 3. Affinity of HAP1 for selected UAS mutants. (**A**) Nucleotide sequence of the UASs used for competition assays. (**B**) Eletrophoretic mobility shift assay. HAP1 extract and the *CYC1* probes were prepared as described in Materials and Methods. Lane – contains no competitor. Triangles represent increasing concentrations of cold competitors used in the assay: 4×, 20×, 50× and 200× molar excess over the labeled *CYC1* probe. The arrow indicates the HAP1–DNA complex. Competition with 1-3 and 1-26 was performed on a separate gel.

Overall, data from random mutagenesis indicate that either CGG or CGC are important nucleotides for binding of HAP1. Changes at these triplets often resulted in lowered binding as exemplified by mutants 1-9, 1-14, 7-20 and 7-25. In addition, changing the second CGC triplet in *CYC7* to CGG resulted in increased activity, suggesting that a CGG triplet is more favorable for binding of HAP1. Other key nucleotides are located 4 bp downstream of the first CGG/C) triplet and 5 bp downstream of the second CG(G/C) triplet, as shown by the single point mutants 7-5 and 1-8. Finally, many mutants (1-3, 1-10, 1-26, 1-32, 1-40, 1-43, 7-1, 7-2, 7-3 and 7-8) that carry mutations outside these above-mentioned nucleotides still showed strong binding to HAP1.

Site-directed mutagenesis of the UASs of *CYC1* and *CYC7*

Based on the above results and comparison with other UASs of HAP1 (see Fig. 2), we concentrated on specific bases within the CGG/C motifs, as well as the downstream TA motifs, and made

Table 2. *In vitro* binding and *in vivo* activity of the UAS of CYC7 and mutants

MUTANT		IN VIVO ACTIVITY (β-GAL UNITS)	IN VITRO BINDING (% WILD TYPE)
7-WT	CCCT CGC TATTAT CGC TATTAGC	2.7	100
7-1	ACTG CGC TATTAT CGC TATTAGC	3.2	45
7-2	ATCT CGC TATTAT CGC GATTAGC	2.7	50
7-3	CCCT CGC TATTAT CGC TATTATC	2.8	70
7-5	CCCT CGC AATAAT CGC TATTAGC	0.2	10
7-8	CCCT CGC TATTAT CGC TATTAGG	2.8	65
7-9	CCCT GGC TATTAT CGC TATTAAC	3.9	55
7-11x	ACCT GGC TATTAT CGC CTATTAGC	<0.1	35
7-19	CCCT AGC TATGAT AGC TATTAGC	<0.1	<1
7-20	CCCT CTC TATTAT CGC TAATAGC	0.4	35
7-21	TCCT CGA TATTAT CGC TATAAGC	0.1	30
7-24	CCCT CGC TATTAT CGG TATTAGC	12.5	120
7-25	CCCT C_C TATTAT CGC TATTAGC	0.3	25
7-26	GCCG CGT TATTAT GGC AATTAGC	0.4	15
7-27	CCCT CGT TATTAT CGC TATTAGC	0.9	35
7-28	CACT CGC TAATAT CGT TATTAGC	0.9	50
7-29	CCCT CGA TGCTAT AGC TAATAGC	<0.1	10
7-30	CCCT CGC TATGAT AGC TAATAGC	0.4	5
7-32	CCCT GGC CATTAT CGC TGTTAGC	0.6	30
7-33	CCCT CGG TATTAT AGC TATTAGC	0.3	20
7-34	CCCT CGC TATTAT CGC CTTTAGC	5.6	40
7-44	CCCG CGC TATTAA CGG ACTATGA	2.5	25
7-45	CCCT CGC TATTAA CGG ACGAGGA	2.1	40

Nucleotide changes are underlined and in bold characters. Binding of HAP1 to mutant UASs is given to relative to wild type UAS of CYC7 (100%)

other alterations using site-directed mutagenesis, as shown in Table 3. First, we changed either (or both) CGG triplet(s) in *CYC1* to the sequence found at similar positions in *CYC7*: CGC. This resulted in a reduction in binding and activation (10-fold) (mutants 1A, 1B and 1C; Table 3). Conversely, changing CGC to CGG in *CYC7* resulted in increased activity (mutants 7D, 7E and 7F; Table 3). In addition, any alteration of the TA nucleotides in *CYC7* reduced binding and activation. For example, changing 1 nt between the two CGC triplets (mutant 7K; Table 3) had drastic effects on binding of HAP1 *in vitro* and activation. Similar results were obtained with mutants 7H, 7I and 7J. At similar positions the *CYC1* site has TT and AT sequences (Table 3). Mutant 1G (TT→TA) showed increased activity, as did 1F (AT→AA). However, an AT→TT change had a negative effect (mutant 1E). Finally, mutating the *CYC1* site to match the CGC triplets and the TA repeats found in *CYC7* (mutant 1D; Table 3) gave levels of activation that were 30% of the UAS of *CYC7*, indicating a secondary role for flanking sequences. From these data it appears that HAP1 can recognize a CGG repeat and also a CGC repeat, although less efficiently. In addition, TA repeats are also important for HAP1 binding. Their presence (and probably that of some other nucleotides, as suggested by mutant 1D) seems to be more important when HAP1 is bound to the weaker CGC motif of *CYC7*. We propose that the 'optimal' UAS for HAP1 is the direct repeat CGG N_3 TA N CGG N_3 TA.

Activity of HAP1 mutants

Since a major difference between the *CYC1* and *CYC7* sites is the sequence of the two triplets, we tested the possibility that they are be responsible for the hyperactivity of HAP1-18 and the PC phenotype of HAP1 mutants, i.e. a defect in activation but not binding at *CYC7*. As expected (21), HAP1-PC1 had wild-type activity at *CYC1* and gave background activity at *CYC7* (Table 3),

Table 3. *In vitro* binding and *in vivo* activity of UASs of *CYC1*, *CYC7* and mutants using expression vectors for HAP1 and mutants

		IN VITRO BINDING (WT HAP1)	IN VIVO ACTIVITY (β-GAL UNITS)		
			HAP1	PC1	HAP1-18
CYC1	–CGG—TT–CGG—AT–	++	6.0	0.1	0.4
1A	–CGC—TT–CGG—AT–	+	0.6	ND	ND
1B	–CGC—TT–CGC—AT–	+/–	0.5	0.3	0.9
1C	–CGG—TT–CGC—AT–	+	0.8	ND	ND
1D	–CGC—TA–CGC—TA–	+	0.7	0.4	9.1
1E	–CGG—TT–CGG—TT–	+	2.4	ND	ND
1F	–CGG—TT–CGG—AA–	++	33.7	ND	ND
1G	–CGG—TA–CGG—AT–	ND	17.6	ND	ND
CYC7	–CGC—TA–CGC—TA–	++	2.5	0.3	48.6
7D	–CGG—TA–CGC—TA–	++	10.5	4.9	7.8
7E	–CGG—TA–CGG—TA–	++	14.5	15.9	0.5
7F	–CGC—TA–CGG—TA–	++	12.8	19.5	0.5
7H	–CGC—TA–CGC—AA–	+	0.4	ND	ND
7I	–CGC—TA–CGC—AT–	++	0.2	ND	ND
7J	–CGC—TT–CGC—TA–	+/–	0.4	ND	ND
7K	–CGC—AA–CGC—TA–	+/–	0.2	ND	ND

Only the targeted areas for site-directed mutagenesis of the UASs are shown. For complete sequence of the UASs refer to Figure 2. Nucleotide changes are underlined and in bold. +/–, <30% binding relative to wild-type HAP1; +, >30% and <50%; ++, >50%. ND, not determined.

Table 4. *In vivo* activity of UASs of *CTT1*, *CYB2* and mutants

MUTANT		IN VIVO ACTIVITY (β-GAL UNITS)
CTT1	–TGG—TA–CGG—TT–	7.5
CTT1-A	–CGG—TA–CGG—TT–	50.7
CYB2	–AGG—TA–CGG—GC–	7.1
CYB2-A	–AGG—TA–CGG—GA–	15.1
CYB2-B	–AGG—TA–CGG—TA–	15.9
CYB2-C	–AGG—TA–CGG—TC–	14.3
CYB2-D	–CGG—TA–CGG—GC–	41.6

Only the targeted areas for site-directed mutagenesis are shown. For complete sequence of the UASs refer to Figure 2. Nucleotide changes are underlined and in bold.

while HAP1-18 shows the opposite pattern of activity (Table 3). Changing either CGC triplet (or both) of *CYC7* to CGG resulted in increased activity of HAP1-PC1 and reduced activity of HAP1-18 (mutants 7D, 7E and 7F; Table 3). Similar results were observed with another PC mutant, HAP1-PC2 (data not shown). In addition, activation by HAP1-PC1 was greater when the second CGC triplet was changed to CGG as compared with the first one. The opposite pattern was seen with HAP1-18, where some activity was retained with a mutation in the first CGC triplet (mutant 7D), while HAP1-18 was transcriptionally inactive with a mutation in the second CGG triplet (7F). The increased activity (>50-fold) of HAP1-PC1 at 7E and 7F (relative to wild-type *CYC7*) cannot be explained by greater binding, since only a modest increase in binding (<2-fold) was observed with HAP1-PC1 tested with 7E and 7F probes in EMSA (data not shown). These results show that the effect of HAP1 PC mutations can be suppressed by mutating 1 or 2 nt of the *CYC7* site. Conversely, activity of HAP1-PC1 was abolished when the CGG triplets of *CYC1* were mutated to CGC (mutants 1B and 1D). It is possible that mutants of HAP1 adopt a different conformation according to the site they are bound to; this would result in different transcriptional activity (see Discussion).

Site-directed mutagenesis at other UASs of HAP1

Comparisons with other known UASs of HAP1 are shown in Figure 2 (see also Table 4). All these sites have, in the middle of the UAS, the sequence 'TA N CGG', which matches the 'optimal' sequence. However, instead of CGG as the first triplet, the sequences TGG and AGG are found in *CTT1* and *CYB2* respectively. In addition, the sequences TT and GC are found downstream of the second CGG triplet for *CTT1* and *CYB2* respectively, as compared with TA for the 'optimal' sequence. If HAP1 binds similarly to these sites, then changes that would allow a better fit with the 'optimal' sequence should result in higher activity. Mutating the first triplet of the UAS of *CTT1* from TGG to CGG resulted in 7-fold higher activity (Table 4). Similarly, changing AGG of *CYB2* to CGG also increased the β-galactosidase activity >5-fold. Increases were also observed when nucleotides were mutated at a position equivalent to the second TA of *CYC7*. However, these changes resulted in a more modest increase (2-fold) in activity (Table 4). Comparison of the UAS of *CYT1* with the optimal sequence shows that this site has the consensus sequence except for the second TA, which is CC. The second TA appears to be less important if other nucleotides match the 'optimal sequence'. Taken together the data show that all HAP1 binding sites are related sequences, i.e. they are imperfect versions of an 'optimal site' that is a direct repeat.

DISCUSSION

HAP1 DNA targets are related imperfect direct repeats

We have performed random mutagenesis of the HAP1 binding sites found in the *CYC1* and *CYC7* genes. The results show that HAP1 binds to a direct repeat with the 'optimal' sequence CGG N_3 TA N CGG N_3 TA. This is in contrast to other zinc binuclear cluster proteins, such as GAL4, PPR1 and PUT3, which recognize palindromic sequences containing inverted CGG triplets (2,3,29). HAP1 can accommodate some changes at these triplets. For instance, the UASs of *CTT1* and that of *CYB2* have the sequence TGG N_6 CGG and AGG N_6 CGG, respectively (Fig. 2). However, activation at these sites is increased 5- to 7-fold when TGG or AGG is mutated to CGG. Variation of the CGG triplets is also seen for other zinc cluster proteins. For instance, one UAS_{GAL4} has an AGG triplet instead of CGG (30).

Divergence from the 'optimal' site is even greater in *CYC7*, where two CGC triplets are found instead of the two CGGs. Again, mutating these triplets to CGG results in greater activation. This is in agreement with studies performed on the native *CYC7* promoter (31). In addition, our mutational analysis shows that the 2 nt (TA) located 4 bp downstream of the two CGG/C triplets are important for binding of HAP1. Increased activity at the UAS of *CYC1* was observed when some nucleotides were mutated at positions equivalent to the first or the second TA of *CYC7* (Table 3). However, similar changes resulted in a more modest increase in activity of *CYB2* and *CTT1* (Table 4). Comparison of the UAS of *CYT1* (Fig. 2) shows that this site has the consensus sequence except for the second TA, which is CC. The second TA appears to be less critical if other nucleotides match the 'optimal' sequence. More drastic effects are seen when mutating the second TA in *CYC7*. Since two CGC triplets instead of CGG are present at that site, it is likely that the second TA sequence helps to stabilize interaction of HAP1 with that UAS. Moreover, the *CYC7* site, as opposed to other HAP1 UASs, is an

almost perfect direct repeat, but is the only site that does not have at least one CGG triplet. Other nucleotides must also be important for binding of HAP1 to that site, as suggested by a mutant of the UAS of *CYC1* (mutant 1D; Table 3). This mutant has the repeat CGC N_3 TA N CGC N_3 TA found in *CYC7*, but shows reduced binding and activation as compared with the wild-type UAS of *CYC7*. Therefore, the presence of two unfavorable CGC triplets in *CYC7* is compensated for by other nucleotides that form an almost perfect direct repeat. Moreover, mutant 7E, which matches the 'optimal' sequence, shows a lower activity than mutant CTT1A, in agreement with a secondary role for other nucleotides located outside the 'optimal' sequence.

Taken together these results show that HAP1 binds to related DNA sequences. Our mutational analysis is in agreement with recently published experiments (32) where random site selection was used to identify HAP1 binding sites. However, in that study the second TA repeat was not identified as being important for HAP1 binding. Our results show that this second repeat is important for increased activity at *CYC1*, *CTT1* and *CYB2*, and essential for activity at *CYC7*.

Discrimination between palindromic and directly repeated sequences

All of the known target sites for GAL4 or the *Kluyveromyces lactis* homolog LAC9 are palindromic (see 4,29,33 for a compilation of the UASs). Conversely, all the known binding sites for HAP1 are imperfect direct repeats (14–19). In addition, no binding of HAP1 could be detected with various palindromic sequences derived from the consensus UAS_{GAL4} with spacing between the two triplets varying from 1 to 14 bp (unpublished results). This is in agreement with the random site selection, where only sequences with direct repeats were recovered (32). In addition, changing the orientation of the second CGG triplet to generate a palindromic sequence prevents binding of HAP1 (32). These observations suggest that there are constraints that prevent HAP1 from binding to a palindromic sequence and GAL4 from binding to a direct repeat. It has been suggested that HAP1 binds to a direct repeat through swiveling of one DNA binding domain relative to the dimerization domain (32). However, if the DNA binding domain of HAP1 shows such flexibility, HAP1 should also be able to bind to a palindromic sequence. One possibility is that the relatively long N-terminal segment of HAP1 that precedes its zinc finger (63 amino acids as compared with 10 for GAL4) prevents HAP1 from binding to a palindromic sequence. Orientation and spacing of the CGG triplets appear to be major determinants for the binding specificity of a given zinc cluster protein. This contrasts with nuclear receptors, which show more flexibility for their binding sites (34–38). Thus our results show that all HAP1 sites are related and are imperfect direct repeats with the optimal sequence CGG N_3 TA N CGG N_3 TA.

Correlation between *in vitro* DNA binding and *in vivo* transactivation

Some mutants do not show a correlation between *in vitro* binding and *in vivo* activation. For instance, in the case of mutant 1-3 (Table 1), which carries five mutations, binding is not affected, but activation is 15 times lower than wild-type *CYC1*. Similarly, mutant 1-26, with one alteration 3 nt downstream of the first CGG triplet, shows greatly reduced activation. Affinity of these UASs for HAP1 was shown to be equivalent to that of *CYC1* by competition assay (Fig. 3). Similar results were observed for nuclear receptors, where a heterodimer formed of LXR/RXR binds to various sites but activates only a certain subset. (39).

In addition, the properties observed at these UASs are analogous to those seen for HAP1 PC mutants (which carry amino acid changes in the DNA binding domain), i.e. wild-type *in vitro* binding but no transcriptional activation (21; Table 3). For instance, HAP1-PC1 shows background activity at *CYC7*, while binding of HAP1-PC1 to *CYC7* was shown to be stronger than to wild-type *CYC7* (21). However, when either CGC triplet is changed to CGG increased activity was observed (Table 3). Binding of HAP1-PC1 to *CYC7* was shown to be stronger than wild-type *CYC7* (21). In contrast to PC mutants, HAP1-18 shows a dramatic increase in activation at the UAS of *CYC7* (~16 times more than wild-type HAP1), even though it has a similar affinity for the UAS of *CYC7* as compared with wild-type HAP1 (20). Introduction of CGG triplets decreases the activity of HAP1-18 as opposed to the effect seen for PC mutants, as suggested from studies with the intact *CYC7* promoter (31).

We proposed (21) that the phenotype of the PC mutants could be explained by the fact that these mutants would prevent interaction of the DNA binding domain of HAP1 with a cofactor protein that would normally act in synergy with the activation domain of HAP1. The model was based on genetic evidence that suggested that a mutant of GAL11, GAL11P, increases the activity of a GAL4 mutant by interacting with its DNA binding domain (40), which contains a zinc finger homologous to that of HAP1. However, more recent data showed that wild-type GAL11, unlike GAL11P, does not interact with GAL4 and is, rather, a component of the RNA polymerase II holoenzyme (41–42). Therefore, HAP1-PC mutants may simply have an altered conformation that may inhibit the activation domain. Changing the CGC triplets to CGG would alter the conformation of PC mutants, enabling them to be active, as seen when bound to the *CYC1* site, which contains two CGG triplets. Conversely, HAP1-18 could have an alternate conformation that would allow very efficient interaction with some components of the basic transcriptional machinery. A similar model has been suggested to explain the phenotype of PC mutants of the glucocorticoid receptor (43), where it is proposed that DNA can act as an allosteric effector.

In conclusion, two different types of mutations modulate the activity of HAP1 without changing its affinity for target sites. Firstly, mutations in the DNA binding domain of HAP1 can lead to increased (HAP1-18) or decreased (PC mutants) activity at *CYC7*. Secondly, mutations in the target sites for HAP1 also affect activation of HAP1 (or mutants). It will be interesting to determine if these two types of mutations have an allosteric effect on the activation domain of HAP1.

ACKNOWLEDGEMENTS

We are grateful to Andrew Nice for help in sequencing mutants. We thank Caterina Russo for synthesis of the spiked oligonucleotides. We are grateful to Drs John White and Hans Zingg for critical reading of the manuscript. We also thank members of the Laboratory of Molecular Endocrinology for helpful suggestions. This work was supported by grants from the Medical Research Council of Canada and the Fonds de la Recherche en Santé du Québec. NH is the recipient of a FCAR (Quebec government) studentship. BT is a scholar of the Medical Research Council of Canada.

REFERENCES

1 Marmorstein,R., Carey,M., Ptashne,M. and Harrison,S.C. (1992) *Nature*, **356**, 408–414.
2 Siddiqui,A.H. and Brandriss,M.C. (1989) *Mol. Cell. Biol.*, **9**, 4706–4712.
3 Roy,A., Exinger,F. and Losson,R. (1990) *Mol. Cell. Biol.*, **10**, 5257–5270.
4 Vashee,S., Xu,H., Johnston,S.A. and Kodadek,T. (1993) *J. Biol. Chem.*, **268**, 24699–24706.
5 Reece,R.J. and Ptashne,M. (1993) *Science*, **261**, 909–911.
6 Corton,J.C. and Johnston,S.A. (1989) *Nature*, **340**, 724–727.
7 Witte,M.M. and Dickson,R.C. (1990) *Mol. Cell. Biol.*, **10**, 5128–5137.
8 Timmerman,J.E., Guiard,B., Shechter,E., Delsuc,M.A., Lallemand,J.Y. and Gervais,M. (1994) *Eur. J. Biochem.*, **225**, 593–599.
9 Creusot,F., Verdiere,J., Gaisne,M. and Slonimski,P.P. (1988) *J. Mol. Biol.*, **204**, 263–276.
10 Pfeifer,K., Kim,K.S., Kogan,S. and Guarente,L. (1989) *Cell*, **56**, 291–301.
11 Zhang,L., Bermingham-McDonogh,O., Turcotte,B. and Guarente,L. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 2851–2855.
12 Fytlovich,S., Gervais,M., Agrimonti,C. and Guiard,B. (1993) *EMBO J.*, **12**, 1209–1218.
13 Zhang,L. and Guarente,L. (1994) *J. Biol. Chem.*, **269**, 14643–14647.
14 Pfeifer,K., Prezant,T. and Guarente,L. (1987) *Cell*, **49**, 19–27.
15 Prezant,T., Pfeifer,K. and Guarente,L. (1987) *Mol. Cell. Biol.*, **7**, 3252–3259.
16 Zitomer,R.S., Sellers,J.W., McCarter,D.W., Hastings,G.A., Wick,P. and Lowry,C.V. (1987) *Mol. Cell. Biol.*, **7**, 2212–2220.
17 Winkler,H., Adam,G., Mattes,E., Schanz,M., Hartig,A. and Ruis,H. (1988) *EMBO J.*, **7**, 1799–1804.
18 Lodi,T. and Guiard,B. (1991) *Mol. Cell. Biol.*, **11**, 3762–3772.
19 Schneider,J.C. and Guarente,L. (1991) *Mol. Cell. Biol.*, **11**, 4934–4942.
20 Kim,K.S. and Guarente,L. (1989) *Nature*, **342**, 200–203.
21 Turcotte,B. and Guarente,L. (1992) *Genes Dev.*, **6**, 2001–2009.
22 Guarente,L., Nye,J.S., Hochschild,A. and Ptashne,M. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 2236–2239.
23 Hochschild,A., Irwin,N. and Ptashne,M. (1983) *Cell*, **32**, 319–325.
24 Schena,M., Freedman,L.P. and Yamamoto,K.R. (1989) *Genes Dev.*, **3**, 1590–1601.
25 Sherman,F., Fink,G.R. and Hicks,J.B. (1986) *Methods in Yeast Genetics.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
26 Forsburg,S.L. and Guarente,L. (1988) *Mol. Cell. Biol.*, **8**, 647–654.
27 Guarente,L. and Mason,T. (1983) *Cell*, **32**, 1279–1286.
28 Pfeifer,K., Arcangioli,B. and Guarente,L. (1987) *Cell*, **49**, 9–18.
29 Bram,R.J., Lue,N.F. and Kornberg,R.D. (1986) *EMBO J.*, **5**, 603–608.
30 Johnston,M. and Davis,R.W. (1984) *Mol. Cell. Biol.*, **4**, 1440–1448.
31 Cerdan,M.E. and Zitomer,R.S. (1988) *Mol. Cell. Biol.*, **8**, 2275–2279.
32 Zhang,L. and Guarente,L. (1994) *Genes Dev.*, **8**, 2110–2119.
33 Halvorsen,Y.C., Nandablan,K. and Dickson,R.C. (1991) *Mol. Cell. Biol.*, **11**, 1777–1784.
34 Baniahmad,A., Steiner,C., Kohne,A.C. and Renkawitz,R. (1990) *Cell*, **61**, 505–514.
35 Kato,S., Tora,L., Yamauchi,J., Masushige,S., Bellard,M. and Chambon,P. (1992) *Cell*, **68**, 731–742.
36 Park,H.Y., Davidson,D., Raaka,B.M. and Samuels,H.H. (1993) *Mol. Endocrinol.*, **7**, 319–330.
37 Tini,M., Otulakowski,G., Breitman,M.L., Tsui,L.-C. and Giguère,V. (1993) *Genes Dev.*, **7**, 295–307.
38 Dana,L.S., Hoener,P.A., Wheeler,D.A., Lawrence,C.B. and McDonnell,D.P. (1994) *Mol. Endocrinol.*, **8**, 1193–1207.
39 Willy,P.J., Umesono,K., Ong,.S., Evans,R.M., Heyman,R.A. and Mangelsdorf,D.J. (1995) *Genes Dev.*, **9**, 1033–1045.
40 Himmelfarb,H.J., Pearlberg,J., Last,D.H. and Ptashne,M. (1990) *Cell*, **63**, 1299–1309.
41 Kim,Y.J., Bjorklund,S., Li,Y., Sayre,M.H. and Kornberg,R.D. (1994) *Cell*, **77**, 599–608.
42 Barberis, A., Pearlberg,J., Simkovich,N., Farrell,S., Reinagel,P., Bamdad,C., Sigal,G. and Ptashne,M. (1995) *Cell*, **81**, 359–368.
43 Lefstin,J.A., Thomas,J.R. and Yamamoto,K.R. (1994) *Genes Dev.*, **8**, 2842–2856.