3518–3526 *Nucleic Acids Research, 1999, Vol. 27, No. 17*

Characterization of the DNA binding and bending HMG domain of the yeast hypoxic repressor Rox1

Jutta Deckert, Roy A. Khalaf, Soo-Myung Hwang and Richard S. Zitomer*

Department of Biological Sciences, University at Albany/SUNY, Albany, NY 12222, USA

Received May 6, 1999; Revised and Accepted July 6, 1999

ABSTRACT

The yeast Rox1 hypoxic transcriptional repressor protein binds to and bends a specific DNA sequence through an HMG domain located at the N-terminus. To better understand the structure of Rox1 and how it interacts with DNA, 38 missense mutations in the HMG domain were isolated through a combination of random and site-directed mutageneses, the latter directed to two Ile residues that play an important role in DNA recognition and bending by HMG domains. The mutants were characterized in terms of their ability to repress the hypoxic gene *ANB1* and the auto-repressed *ROX1* gene *in vivo*. The mutant HMG domains were fused to maltose binding protein and expressed in and purified from *Escherichia coli* and their relative affinities for DNA and ability to bend DNA were determined. A model of the structure of the Rox1 HMG domain was derived using sequence similarities between Rox1 and the human protein SRY, the structure of which has been determined. The results of the mutational analysis are interpreted in terms of the model structure of Rox1.

INTRODUCTION

Rox1 is a DNA binding protein that represses expression of the hypoxic genes in the yeast *Saccharomyces cerevisiae* (1,2). It consists of 368 amino acids that are divided into two functional domains. The first third of Rox1 comprises a site-specific DNA binding domain with homology to the HMG motif (3). This domain directs binding to the regulatory regions of the hypoxic genes, each of which contains several Rox1 sites (3–5). The last two-thirds of the protein contains a repression domain that recruits the general repression complex Tup1/Ssn6 to the DNA (4). Analysis of the operator regions of the hypoxic gene *ANB1* indicated that at least two Rox1 binding sites are required for repression, but that Rox1 bound to each site independently (5). The two-site requirement was presumed to arise from the assembly of the higher order repression complex.

The HMG DNA binding motif is found in both site-specific and non-specific DNA binding proteins. The former class includes the human male-determining transcription factor SRY, the founding member of the SOX family of transcriptional activators, and the murine lymphocyte transcription activator LEF-1 (6–8). These proteins bind to DNA as monomers and bend DNA (6,8). Based upon primary sequence similarity, the SOX family has been grouped together as more closely related to each other than to the LEF-1 family and Rox1 shares greater similarity to the SOX family. The NMR-derived structures of the HMG domains of SRY and LEF-1 complexed with DNA have been determined (9,10). They both fold into L-shaped structures involving three α-helices. The DNA binds along the inside of the L, making extensive contacts with the protein and resulting in the induced bend in the DNA of 85° for the SRY–DNA complex and 130° for the LEF-1–DNA complex. The amino acid residues that contact the DNA are highly conserved within the SOX family and Rox1, as is the DNA binding site, which contains a core 6 bp consensus sequence (A/T)TTG(A/T)(A/T) (7). While the various proteins may have some sequence preferences within this consensus and extending beyond it, they appear to be able to bind to each others' sites well (11). Two particularly interesting interactions between DNA and SRY involve an Ile residue in the first helix (conserved in the SOX family and Rox1), which intercalates between the first two T/A residues of the core sequence to aid in the induced bend, and a second conserved Ile in the second helix, which may partially intercalate into the first A/T base pair of the core and the base pair immediately preceding.

Rox1 represents an excellent system for the study of the relationship between the structure and function of the HMG DNA binding motif. The yeast hypoxic gene regulatory system is genetically tractable so that both randomly induced and directed mutations can be generated and studied both *in vivo* and *in vitro*. A mutational analysis of the Rox1 binding site defined the effect of every possible single base pair substitution in the core sequence and the base pair immediately upstream (5). A preliminary mutational analysis of the Rox1 HMG domain identified eight missense mutations in seven residues, all conserved with SRY (4). This study and the observation that Rox1 bends DNA at an angle of 90°, similar to the bend induced by SRY, reinforces the proposition that Rox1 folds into an SRY-like structure.

In this study, the mutational analysis of the Rox1 HMG domain was extended. In the previous study, mutations were selected for constitutive expression of the hypoxic *ANB1* gene.

*To whom correspondence should be addressed. Tel: +1 518 442 4385; Fax: +1 518 442 4767; Email: rz144@cnsvax.albany.edu
Present address:
Jutta Deckert, Department of Biochemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA
Permanent address:
Soo-Myung Hwang, Department of Clinical Pathology, Jisan Junior College, Pugok-dong, Kumjung-Ku, Pusan, Korea

While the mutations obtained highlighted the similarity between Rox1 and other HMG proteins, all the mutant proteins were non-functional *in vivo* and *in vitro*, limiting the amount of structural information that could be derived. The severity of the mutations resulted because Rox1 is auto-repressed (12) and partially functional *rox1* mutations would result in increased Rox1 protein levels, which in turn would still repress *ANB1* expression. Thus, to obtain partially functional mutations, a new selection scheme was designed based upon loss of *ROX1* auto-repression. The new selection resulted in many mutations which retained some function. In addition, mutations were directed to the two Ile residues that specifically interact with DNA. The results of these studies, combined with computer modeling that indicates that Rox1 folds into a similar structure to that of SRY, provides some insight into the role of the different residues in the structure and DNA binding of the HMG motif.

MATERIALS AND METHODS

Strains, cell growth and transformations

The yeast strain RZ53-6Δ*rox1* has been described (12). MZ22-4Δ*r1*::*gK* (MAT**a**, *trp1*, *leu2*, *ura3*::*AZ4*, *his3*, *gal1*-Δ*152*, *rox1*::*R1*/*gK)* was constructed through standard yeast genetics (13) and deletion of the *ROX1* gene by integration of a *ROX1* promoter–*Escherichia coli galK* fusion from the plasmid YC(22)p*R1*::*gK* (described below). Yeast cells were grown in liquid at 30°C with vigorous shaking on either non-selective YPD or SC medium lacking the appropriate nutrients for selection for plasmid maintenance (13). For selection of *ROX1* mutants, cells were plated on SC medium lacking glucose and containing 2% galactose. Yeast transformation was carried out as described (14).

The bacterial strain HB101 was used for construction and maintenance of plasmids and growth and transformation were carried out as described (15).

Plasmids

The plasmids YCplac33, YCp(33)*ROX1e*, YCp(22)*ROX1/Z*, YCp(22)*R1*-2(9E10), YCp(33)*R1*Δ*100/145*, YCp(22)*R1*Δ*100/245*, YCp(33)*AZ*, p*MAL-ROX1*(HMG) and pCY4-R1Op have been described (4,16). All plasmid constructions were carried out using standard methods as described (15). The details and PCR primer sequences are available upon request.

YCp(22)*RI*/*gK* contains a replacement of the *ROX1* coding sequence with the 1.1 kb *galK* coding sequence from the *E.coli* galactokinase gene in plasmid YCp(22)*R1*-2(9E10). The fusion point is immediately after the *ROX1* translational initiation codon. Digestion of this plasmid with *Hin*dIII released the 2.8 kb fusion construct for integration into the yeast genome, creating a *ROX1* deletion allele and placing expression of *galK* under the control of the *ROX1* regulatory region.

YCp(33)*R1*Δ*100/145*-MluI contains a silent mutation in codon 16 that created a *Mlu*I site and a deletion of codons 100–145. YCp(22)*R1*Δ*100/245*-BstBI contains a silent mutation at codon 43 that created a *Bst*BI site and a deletion of codons 100–145.

YCp(33)*rox1-I18X* represent a set of plasmids where X represents mutations at codon 18. These mutations were constructed as follows. For random PCR mutagenesis at this codon, a set of oligonucleotides was synthesized that was identical to the *ROX1* coding sequence from codon 14 to 25 except for a silent substitution at codon 16 that created an *Mlu*I site and the substitution of all four bases in equal concentrations at each of the three positions at codon 18. This synthetic DNA was used as a primer to amplify the *ROX1* coding sequence, which was then ligated into YCp(33)*R1*Δ*100/245*-MluI to generate a set of plasmids containing all possible codons at position 18 (and restoring the wild-type sequences for codons 100–145). For mutations in which a single amino acid substitution in Rox1 was desired, the primer indicated above contained a specific codon.

YCp(22)*rox1-I45X* represent a set of plasmids where X represents mutations at codon 45. These mutations were generated as above except that the primer was complementary to codons 42–53 and carried a silent substitution at codon 43 that created a *Bst*BI site and all four bases at each of the three positions of codon 45. The PCR product was ligated into YCp(22)*R1*Δ*100/245*-BstBI to generate a set of plasmids containing all possible codons at position 45 (and restoring the wild-type sequences for codons 100–145). For mutations in which a single amino acid substitution in Rox1 was desired, the primer contained a specific codon.

The p*MAL-rox1-X*(HMG) plasmids, containing a maltose binding protein (MBP)–Rox1 HMG domain fusion with the various *ROX1* mutations, were constructed by PCR amplification of the mutant sequences and ligation into the p*MAL-ROX1*(HMG) vector (4). The sequence of all these constructs was confirmed.

Purification of MBP–Rox1(HMG)

The MBP–Rox1(HMG) fusion proteins were purified from *E.coli* extracts using amylose affinity chromatography as described (3). In all cases, the MBP–Rox1(HMG) protein was estimated to be >90% pure by Coomassie blue staining of SDS–polyacrylamide gels (15). Protein concentrations were determined using the Bio-Rad assay (Bio-Rad Laboratories, Hercules, CA).

β-Galactosidase assays, immunoblots and gel retardation experiments

Cells were grown and β-galactosidase assays were performed as described (12,13). Assays were carried out multiple times on at least two independent transformants.

Immunoblots were carried out using monoclonal antisera against the 9E10 c-myc epitope as recommended by the vendor (Oncogene Research Products, Cambridge, MA; 15).

Gel retardation experiments with either synthetic DNA or fragments from plasmid pCY4-R1Op were carried out as described (4) and the results were quantitated using a Betascope.

RESULTS

Selection of *rox1* mutants

Amino acid substitution mutations that result in only a partial loss of protein function can provide very useful information for structural analyses. Such mutant proteins still fold into an overall native-like structure and the role of the altered residue can be inferred more insightfully than in the case of a mutation that renders a protein devoid of any function. For the DNA binding domain of Rox1, the study of mutant proteins that still bind and

Table 1. Characterization of *rox1* mutant alleles

ROX1 allele	β-Galactosidase activity[a]		DNA binding affinity[b] (%)	DNA bending angle (°)
	ROX1/*lacZ*	*ANB1*/*lacZ*		
Wild-type	12	1.1	100	94
Δ*rox1*	166	71	—	
I18T	241	61	<0.1	
F20L	229	51	—	
H23P	204	73	<0.1	
S46F	240	58	- - -	
W64R	247	72	—	
L67P	235	77	1	91
P13L	128	36	<0.1	
W53R	122	62	—	
W64L	140	52	<0.1	
A68V	183	23	3	69
H75R	130	4.7	5	68
A68T	131	3.8	10	86
P86L	118	3.9	2	67
K61E	94	3.1	- - -	
I10M	81	1.4	30	92
S43P	68	4.6	40	92
E74G	54	1.8	20	90
K54E	50	1.6	- - -	
I45T	48	2.2	20	86
H25Y	35	1.4	20	87
K89E	33	1.3	- - -	

[a]Assays were performed with *MZ22-4*Δ*r1*::*gK* transformed with the indicated mutant for *ANB1*/*lacZ* or RZ53-6Δ*rox1* transformed with the YCp(22)*ROX1*/*lacZ* and the indicated mutant for *ROX1*/*lacZ*. The numbers given are the averages of more than three assays performed with at least two independent transformants.
[b]The symbols used are as follows: —, no DNA binding was observed; - - -, the protein bound DNA in a high molecular weight complex which probably represented non-specific binding.

bend DNA to some extent or bind to DNA with an altered specificity can provide a great deal of information about the role of specific residues in contacting DNA or in folding of the HMG domain. We had previously used a selection for mutations in the *ROX1* gene based upon constitutive expression of the strongly repressed hypoxic *ANB1* gene, but this selection resulted in mostly inactive protein (4). In hindsight, this result can be explained by the finding that Rox1 represses its own expression (12). Consequently, *rox1* mutations result in overexpression of the mutant gene and, for mutant proteins which retain some activity, these increased levels could still repress expression of the *ANB1* gene.

To circumvent this problem, we designed a selection based upon increased expression of *ROX1* itself. The *E.coli* galactokinase coding sequence, *galK*, was fused to the *ROX1* upstream region and this construct was integrated into and disrupted the *ROX1* locus of a *gal1*Δ yeast strain. The resulting strain was gal$^+$, able to grow on galactose, due to the high level of *galK* expression, but when a wild-type *ROX1* plasmid was transformed into cells, expression of the *ROX1*/*galK* fusion was repressed, resulting in a gal$^-$ phenotype. This strain also contained a fusion of the *E.coli lacZ* gene coding sequence to the *ANB1* regulatory sequences integrated into and disrupting the *URA3* gene so that repression of this hypoxic gene could be monitored on X-gal plates. To obtain *rox1* mutants, the *ROX1* sequence was mutagenized by PCR amplification, then ligated into a centromeric plasmid containing the *URA3* gene. The above strain was transformed with this mutagenized pool and plated onto galactose plates lacking uracil. Only plasmids carrying a *rox1* mutant allele could give transformants; those carrying a wild-type *ROX1* gene repressed expression of the *ROX1*/*galK* fusion, precluding growth on galactose plates.

Transformants were screened for expression of the *ANB1*/*lacZ* fusion and immunoblots were performed both to ascertain that a full-length Rox1 protein was present in the cells and that it was overexpressed. For each mutant overexpressing the full-length Rox1 protein, the plasmid was recovered and retransformed into fresh cells to confirm that the mutant phenotype

Table 2. Characterization of *rox1-I18* and *I45* mutant alleles

ROX1 allele	β-Galactosidase activity[a]		DNA binding affinity[b] (%)	DNA bending angle (°)
	ROX1/lacZ	*ANB1/lacZ*		
Wild-type	12	1.1	100	94
Δ*rox1*	166	71	ND	
I18L	76	2.4	10	76
I18V	118	3.0	10	78
I18M	35	1.1	20	92
I18F	65	2.6	- - -	
I18T[c]	241	61	<0.1	
I18G	201	40	ND	
I18K	134	41	ND	
I45L	73	11	15	93
I45V	46	8.4	50	92
I45M	134	30	ND	
I45F	59	0.4	40	93
I45Y	156	42	ND	
I45W	125	35	2	94
I45H	161	58	ND	
I45G	151	40	ND	
I45T[c]	48	2.2	20	86
I45S	76	8.4	3	79
I45R	50	5.4	30	84
I45K	190	42	5	74
I45D	235	96	ND	

[a]Assays were carried out as described in Table 1.
[b]The symbols used are as follows: - - -, protein could not be isolated from cells; ND, not determined.
[c]These mutant alleles were isolated through random mutagenesis and the data was taken from Table 1.

was plasmid borne. Finally, the *ROX1* coding sequence was determined to locate the mutant residue. The mutants obtained are shown in Table 1. Mutant alleles are designated using the single letter amino acid code with the wild-type residue given first, followed by the amino acid number, then the new residue.

Mutations targeted to Ile residues 18 and 45

Two residues of *ROX1* were specifically targeted for mutagenesis based upon the important role they appear to play in DNA binding. The Ile residue at position 18 is conserved in the SRY/SOX family and, based upon NMR structural analyses, has been proposed to intercalate between two A/T base pairs of the target DNA to promote bending. I45 is also highly conserved and is also proposed to contact adjacent base pairs, although it does not appear to intercalate, as does I18. To determine what substitutions might be tolerated, these residues were subjected to random mutageneses. In these cases, no selection or screen was applied in choosing which mutations to study; all those isolated were characterized. Unfortunately, the mutagenized plasmid pools did not yield a full set of all 19 possible substitutions at each residue and, consequently, site-directed mutagenesis was used to obtain other substitutions that were deemed to be potentially interesting. These site-directed mutations included conserved substitutions, such as I45L, and substitutions of aromatic or bulky hydrophobic groups to determine whether they affected DNA interactions, such as I18F, I18M, I45W, I45Y and I45H. In addition, the I45K substitution was generated because of the interesting features of the I45R mutation (see below). Combining both the random and directed mutations, six substitutions at I18 and 13 at I45 were generated (Table 2).

Effect of the *rox1* mutations on repression *in vivo*

The abilities of the mutant Rox1 proteins to repress expression of both the *ANB1* and *ROX1* genes were assessed using *lacZ* fusions and the results are presented in Tables 1 and 2. The mutants are divided into two groups in Table 1. The top group represent mutations, like F20L, that cause complete or near complete loss of repression activity as determined by complete derepression of both *ROX1* and *ANB1*. The second represents a range of partially functional mutations that retain some repression. Mutations like P13L resulted in complete derepression of *ROX1* but some level of repression of *ANB1*, indicating that the

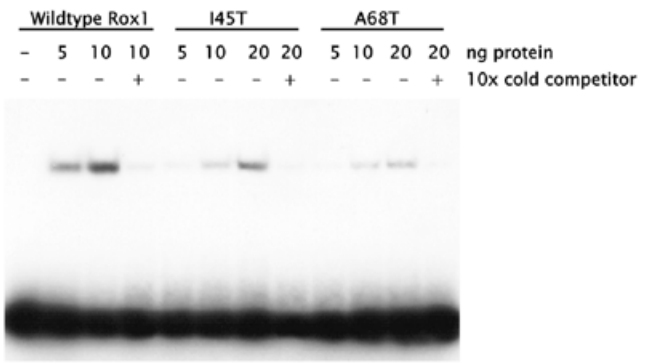

Figure 1. Comparison of wild-type and mutant protein binding to DNA. A gel retardation assay was carried out as described in Materials and Methods with the wild-type and the I45T and A68T mutant Rox1 HMG domains. The reactions contained 1 ng (15 000 c.p.m.) of the ^{32}P-labeled 32 bp synthetic DNA carrying a single Rox1 binding site and, where indicated by a + above the lane, 10 ng of the same DNA unlabeled. The amount of protein added to each reaction (in ng) is indicated above the lanes.

mutant proteins retained some level of activity. Others, like H25Y, retained substantial activity, as indicated by strong repression of *ANB1* and some level of repression of *ROX1*.

Table 2 shows the effects of various mutations at I18 and I45 on repression of the two reporter genes. The same regulatory classes can be found among these mutations; some, such as I18G and I45D, caused complete derepression of both genes, while others, like I18V and I45G, encoded proteins that retained some activity and others, like I18L and I45R, had substantial repression activity.

DNA binding activity of the mutant HMG domains

DNA binding studies were carried out to determine the relative affinity of the mutant Rox1 HMG domains for DNA. Codons 2–103 of each mutant were subcloned into a bacterial expression vector to generate a MBP fusion and these fusion proteins were purified to >90% homogeneity. We previously determined that the MBP–HMG fusions bound a Rox1 binding site with the same affinity and bent DNA at the same angle as the free HMG domain or the intact protein (4). The purified proteins were titrated with a 32 bp synthetic DNA containing a single Rox1 binding site; the wild-type protein binds to this fragment with a K_d of 20 nM (5). Since many of the mutant proteins bound DNA very poorly and showed substantial non-specific binding at high protein concentrations that interfered with the calculation of binding constants, we evaluated the effect of the mutations on DNA binding by comparing the amount of wild-type versus mutant protein that was required to give the same amount of complexed DNA. Each gel included a wild-type protein titration so that comparisons were only made within the same experiment.

A typical set of experiments is shown in Figure 1 for the wild-type and I45T and A68T mutant proteins. A lane containing unlabeled competitor DNA was included to ascertain that binding was specific. The percent wild-type binding for each mutant for which specific binding was detected is listed in Tables 1 and 2. As might be anticipated, there is an excellent correlation between the *in vivo* repression activity of the mutant Rox1 proteins and their abilities to bind DNA specifically *in vitro*. Mutants with no or weak repression activity showed either no DNA binding activity or, like L67P, very minimal binding, while mutants such as I10M and I45V that showed significant repression of both *ANB1* and *ROX1* bound DNA with ~30 and 50% of the affinity of the wild-type, respectively. There were a few problems. The mutants K54E, K61E and K89E were capable of repression *in vivo*, but in the gel retardation assays, these proteins generated a smear rather than a discrete band, suggesting that the mutations increased non-specific binding, precluding an assessment of their relative affinities for the specific binding site. It is possible that *in vivo* the higher order repression complex that forms at the operator sites allows a distinction between the specific and non-specific Rox1 binding that we cannot distinguish *in vitro* with purified protein. Also, one mutant, I18F, could not be expressed in the bacterial system.

DNA bending activity of the mutant HMG domains

We have used the circular permutation plasmid assay to measure a bending angle of ~90° for DNA bound by Rox1 (5). This assay is based upon the observation that bent DNA runs slower in a polyacrylamide gel than does linear DNA and that the rate of migration is inversely proportional to the bending angle. Furthermore, DNA molecules of equal length and bent at the same angle migrate differentially in accordance with the position of the bend; the closer the bend is to the end of the fragment, the faster the migration. Thus, placing a Rox1 binding site at different positions within a DNA fragment results in differential migration of these fragments in a gel retardation assay, as seen in Figure 2. The angle of the bend can be calculated from these data and for those Rox1 mutant HMG domains for which significant binding was measured, the bending angles were determined and are presented in Tables 1 and 2. None of the mutant proteins resulted in a drastic change in the calculated bending angle, as might be expected from the large surface over which HMG proteins contact DNA. However, we did observe some correlation between the binding affinity and the bending angle; the weaker binding mutants resulted in a lower calculated bending angle. We believe that it is possible that this general correlation may result from an artifact of how the binding affinity affects migration rate in this bending assay. Interestingly, there are two exceptions. The protein L67P was calculated to bend DNA at an angle of 91° despite a binding affinity reduced by ~100-fold. This correlation did not hold for I45 mutations; rather, there appears to be a direct correlation with the bulkiness of hydrophobic residues at this position and the bending angle. I45 sits near the end of the short arm of the L (see Discussion and Fig. 5C) and bulky groups may force a more severe bend. Future NMR studies of the protein–DNA complex are needed to determine the true bending angles.

Specificity of DNA binding

The Rox1 binding sites in the regulatory regions of the hypoxic genes consist of a central core sequence, 5'-ATTGTT at positions 4–9, designated as such because it is a common feature of many HMG protein binding sites, and less well conserved sequences 5' and 3' (7). A mutational analysis of the core sequence and the base pair immediately 5' to it, base pair 3, revealed that certain base pair substitutions were tolerated only in the first and last base pairs of the core sequence and

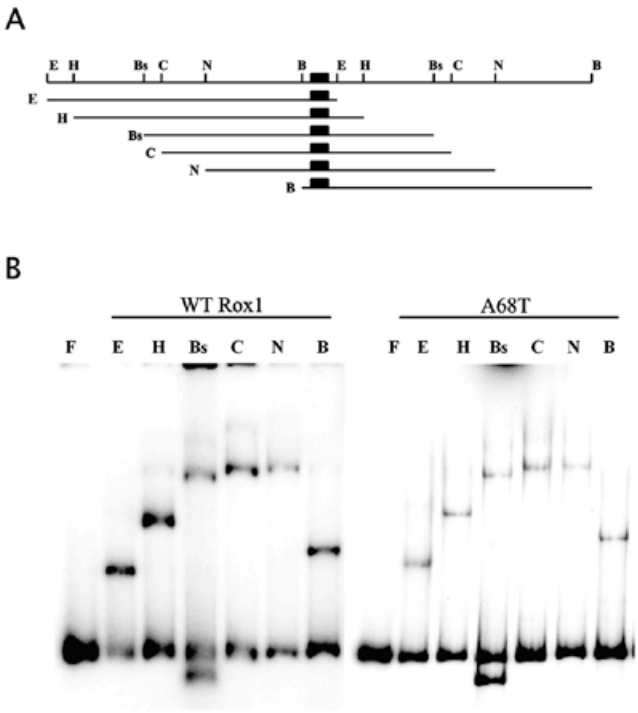

Figure 2. DNA bending by the wild-type and the A68T Rox1 HMG domains. A gel retardation assay to determine the extent of DNA bending was carried out with 10 ng/lane wild-type Rox1 and 50 ng/lane A68T mutant Rox1. (**A**) The DNA fragments used were generated by digestion of plasmid pCY4-R1OpWT containing a single Rox1 site between a direct repeat of the *Eco*RI–*Bam*HI segment of pBR322. The enzymes used were: E, *Eco*RI; H, *Hin*dIII; Bs, *Bst*NI; C, *Csp*61; N, *Nhe*I; B, *Bam*HI. Digestion of the plasmid with *Bst*NI generates a second fragment slightly smaller than that carrying the Rox1 binding site, which is responsible for the lowest band in that lane. The fragments were labeled with ^{32}P and ~1 ng (~15 000 c.p.m.) was used in each binding reaction. (**B**) Two different gels for the two different proteins.

either a C/G or T/A base pair was required at position 3. The *ANB1* regulatory region contains four Rox1 sites organized into two operator regions while the *ROX1* regulatory region contains two Rox1 sites. These sites are not identical, as seen in Figure 3A, and it was of interest to determine whether the mutant proteins with reduced repression activity affected both genes similarly. We plotted the fold repression for *ROX1* for each mutant against that for *ANB1* using the values taken from Tables 1 and 2 and a curve was fitted to the points, as seen in Figure 3B. The curve is not a straight line because the *ROX1* gene is not repressed as effectively as is *ANB1* and as *ROX1* is derepressed, more repressor is made and, therefore, *ANB1* is more efficiently repressed. Therefore, loss of *ANB1* repression is only seen with very weak repressors.

Interestingly, several points fall well off the curve and we believe this is a manifestation of an altered specificity of Rox1 for its binding site. Those mutant proteins that are represented by points which are above the curve repress *ANB1* relatively better than *ROX1* compared to the other mutant proteins and those represented by points that fall below the curve repress *ANB1* relatively less well compared to *ROX1*. The majority of the mutants that fall off the curve have substitutions at I45. By analogy to the SRY structure, this residue contacts the first base pair of the core sequence, which is conserved between *ANB1* and *ROX1*, and base pair 3 (see Fig. 5D); three of the four *ANB1* Rox1 sites have a C/G base pair while both *ROX1* sites have a T/A base pair at this position. We propose that the different substitutions at I45 alter the specificity of the protein for the DNA in a way that discriminates between C/G and T/A at position 3. Residue S43 is also proposed to contact base pair 3 (see Fig. 5D) and the mutant S43P displays differential repression of *ANB1* and *ROX1*. The effect of I10M, however, is unclear.

We attempted to investigate this proposed altered specificity by measuring the relative affinities of the mutant proteins for different synthetic DNAs, but we could not detect a difference, suggesting that the discrimination seen *in vivo* is either too small to measure with the sensitivity of the *in vitro* binding experiments or that other members of the repression complex, such as Tup1 and/or Ssn6 or the C-terminus of Rox1, enhance the differences *in vivo*.

DISCUSSION

We report here the isolation and characterization of over 36 amino acid substitution mutations in the HMG domain of the yeast transcriptional repressor Rox1. This collection includes both randomly generated and site-directed mutations and many resulted in proteins that retained some level of wild-type activity. These partially functional mutations hold a great deal of potential for obtaining information about the roles of specific residues in the functioning of the protein.

Computer model of Rox1 HMG domain structure

To complement this mutational analysis, a computer model of the structure of the Rox1 HMG domain was generated using the coordinates of the NMR-derived structure of the SRY HMG domain and the Insight II program (Biosym Technologies). We believe this model to be a close facsimile of the actual Rox1 HMG domain structure because: (i) Rox1 and SRY have extensive similarities in their amino acid sequences (Fig. 4); (ii) they bind to similar sites; (iii) they bend these sites at similar angles; (iv) of the 28 mutations isolated from the random mutageneses reported here and elsewhere, 19 were in residues identical and six were in residues conserved between Rox1 and SRY, as seen in Figure 4. Furthermore, a number of mutations have been identified in SRY that cause sex reversal (18–24). Those that lie within the HMG domain and strongly affect DNA binding highlight the similarity of the SRY and Rox1 structures. All the SRY mutations fall in residues that are either identical to those in Rox1 (mutations R7G, I13T, G40R, L46H, L51I, A58T and Y72C, numbered as in Fig. 4) or represent conservative substitutions between the two proteins (mutations V5L and M23T). In addition, we have isolated mutations in the Rox1 equivalent of four of these residues. Thus, we feel justified in interpreting the effects of the Rox1 mutations described here in terms of the model structure.

The model is shown in Figure 5. The backbone structures of SRY and Rox1 are superimposed in Figure 5A. There is only one major difference; the loop between the first and second α-helices of Rox1 has an additional five residues which causes it

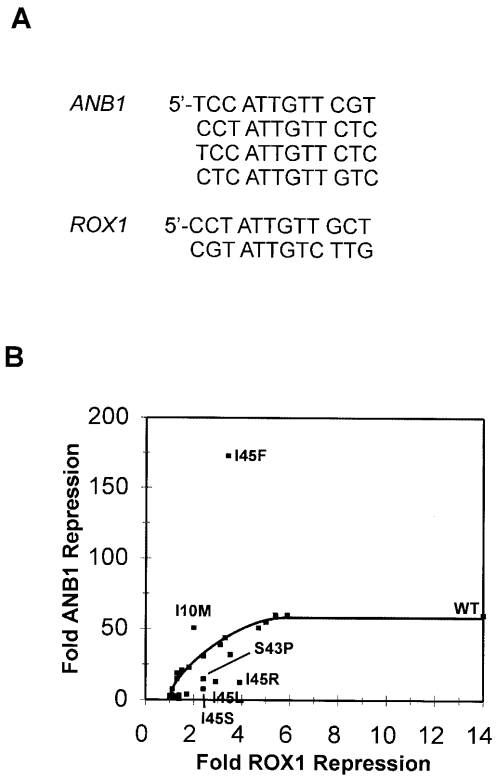

Figure 3. Differential repression of the *ROX1* and *ANB1* genes by mutant Rox1 proteins. (**A**) The Rox1 binding sites in the regulatory regions of *ANB1* and *ROX1* are presented with the six internal core base pairs offset. (**B**) Fold *ANB1* repression for each mutant (the β-galactosidase activity of the Δ*rox1* mutant divided by that of the mutant in question) was plotted against fold *ROX1* repression (calculated as above). The curve was drawn to fit the points and the mutants that fell off the curve are labeled.

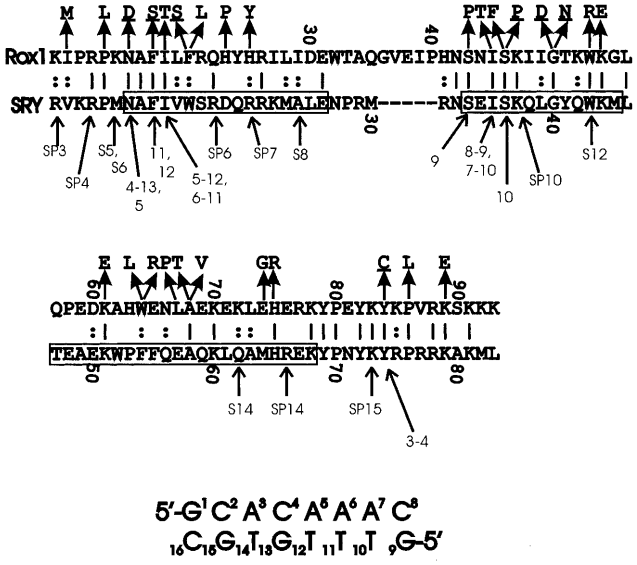

Figure 4. Comparison of the Rox1 and SRY HMG domain sequences. The sequences of the Rox1 and SRY HMG domains are aligned. The Rox1 protein is numbered from the initiation ATG, while the SRY protein numbering follows that in Werner *et al.* (9). Identical residues are indicated by a line between them; conservative substitutions are indicated by a colon. The boxed residues of SRY represent those involved in α-helices. The residues that contact DNA are indicated as follows: contacts with a specific base are indicated by the number of that base in the DNA sequence below; contacts with a sugar residue are indicated with an SX where X represents the number of the base to which the sugar is bonded; contacts with a phosphate residue are indicated as SPX where X represents the base 3' to the phosphate. The mutations in Rox1 are indicated above the sequence. The DNA sequence below the protein sequences is that for which the SRY protein–DNA structure was determined (9).

to protrude more. We did not dock the DNA to the Rox1 model, but in SRY the DNA binds to the inside of the L-shaped structure, as shown in Figure 5B.

We have used this model to interpret the effects that the Rox1 mutants would have on protein structure and function. The mutations mapped to residues throughout the HMG domain, as seen in Figures 4 and 5C. We have divided them into four categories: those that effect the structure of the protein; those that effect residues that interact with DNA; and those in the two targeted residues I18 and I45.

Structural mutations

A number of mutations were isolated in residues that do not contact DNA in the SRY structure and these are indicated in red in Figure 5C. There are three aromatic residues, F20, W53 and W64, that appear to interact with each other in the core of the protein, as highlighted in Figure 5C, where their side chains are presented. Each is part of one of the three α-helices that comprise the HMG domain and are conserved between SRY and Rox1, although, interestingly, the equivalent residues to F20 and W64 in SRY are reversed (Fig. 4). We obtained mutations in all three and in each case the phenotypes and *in vitro* activity were severely affected, as might be expected if these residues play a critical role in maintaining the helices in appropriate orientations. While the F20L and W64L substitutions retained the hydrophobic character of the residues, the non-planar Leu side chain probably interferes with the interactions of the aromatic residues. It should be noted that W53 is also proposed to make a contact with a phosphate residue of DNA, making the interpretation of W53R less clear. The same Trp→Arg substitution was reported in the analogous residue of another HMG domain protein, HMG1, where a similar loss of protein function was observed (17).

The mutations H23P and L67P introduced Pro residues into α-helical regions, which probably disrupted the backbone structure. However, the L67P protein bound DNA sufficiently to allow measurement of a bending angle of 91°, one of the exceptions to the general correlation between binding affinity and bending angle. It is possible that the kink introduced into the middle of the third helix actually collapsed the DNA binding pocket, resulting in a steeper bending angle.

Mutations in residues that interact with DNA

The model generated for Rox1 was not docked with DNA, but we assume that Rox1 makes similar DNA contacts as does SRY; the residues of SRY that interact with DNA are indicated

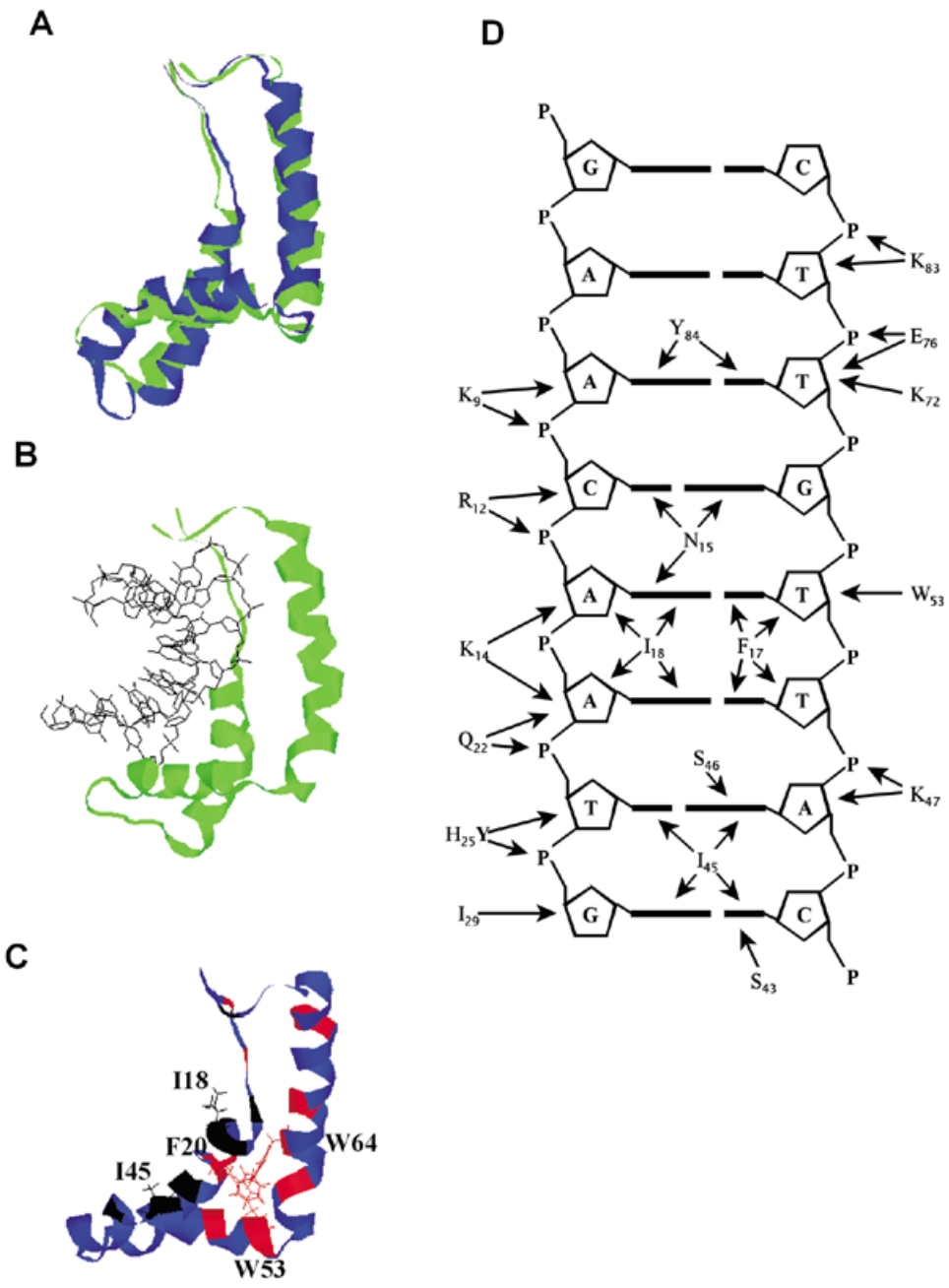

Figure 5. Rox1 model structure. (**A**) Superimposition of the Rox1 (blue) and SRY (green) backbones is presented. The coordinates for SRY were obtained from the Brookhaven Protein Data Base and those for Rox1 were derived from model building. (**B**) The backbone structure of the SRY–DNA complex is presented. The DNA wire frame is shown in black and the SRY ribbon in green. (**C**) The ribbon structure of Rox1 is shown with the side chains of residues I18 and I45 in black and the side chains of residues F20, W53 and W64 in red. Other residues which have been mutated are presented in the main chain as black for those that contact DNA and red for those that do not. (**D**) The Rox1 binding site is represented as a ladder with the base indicated within the sugar. The residues that are proposed to contact DNA, based upon similarity to the SRY structure, are indicated, with arrows indicating the contacts.

in Figure 4 and almost all are conserved in Rox1. In addition, the core of the DNA target sequence is the same. The equivalent residues in Rox1 and the contacts we propose they make with DNA are indicated in Figure 5D. We isolated a number of mutations in these residues and their positions in Rox1 are indicated in black in Figure 5C. We had previously reported the mutations N15D, F17S, S46P and Y84C, where only Y84C retained some repression activity. Among the mutations reported here, H25Y and S43P retained substantial repression and DNA binding activity. The analogous residue in SRY to

H25 is an Arg which contacts the sugar–phosphate backbone, as indicated in Figures 4 and 5D. The substitution of a Tyr for the His in Rox1 had a modest effect. S43 is at the beginning of the second helix and the corresponding Ser in SRY contacts the pyrimidine of base pair 3. It appears that the substitution of a Pro neither disrupts the overall structure nor affects the DNA binding activity dramatically. Two other mutations, S46F and W53R, caused complete loss of activity and DNA binding. The SRY Ser corresponding to S46 in the middle of the DNA binding site contacts base pair 4 and substitution of a bulky aromatic residue probably precludes DNA docking. W53 is also conserved in SRY and makes contact with a sugar. It may also stack with F17, which would be disrupted by the Arg substitution. The random mutagenesis also led to mutations in I18 and I45, but these will be discussed below.

Mutations in I18

The I18 residue in Rox1 corresponds to an Ile residue in SRY that intercalates between adjacent TA base pairs, which is presumably important in bending of the DNA. This residue was targeted for mutagenesis in this study. The most conservative substitutions of Val or Leu reduced DNA binding activity >10-fold, while, surprisingly, the substitution of a bulky Met side chain reduced binding only 5-fold. A Phe substitution resulted in a level of repression similar to that for the Leu and Val mutants, but the purified protein was inactive. Substitutions of Lys and Gly were isolated and, as might be expected, these mutations rendered the protein inactive. The results indicate that a wide range of aliphatic substitutions are permissible with some loss of binding, suggesting some flexibility in the requirements for intercalation.

Mutations in I45

The SRY Ile corresponding to the Rox1 I45 interacts with base pairs 3 and 4, but does not appear to intercalate, as does I18. Thirteen different substitutions were isolated at this residue. The most conservative changes, to Leu and Val, resulted in proteins with substantial *in vivo* activity and a 7- and only 2-fold reduction in binding activity respectively. The substitution of a Phe also resulted in a protein with substantial activity, with only a 2.5-fold reduction in binding. However, mutations that introduced the other aromatic amino acids, Tyr and Trp, had more severe effects. Of the other substitutions, I45R was the most intriguing; this mutant retained substantial repression and DNA binding activity. It is unlikely that the Arg residue functions in the same fashion as the Ile, but perhaps new contacts are generated. The Lys substitution had much weaker activity. I45T also retained substantial activity, with a 5-fold reduction in DNA binding, while a Ser substitution had a more severe effect. The other mutations were relatively inactive.

The mutations described here provide some powerful tools for a structure–function analysis of the HMG DNA binding motif. While the model structure allows some insight, it is obvious that future studies will require direct analysis of the wild-type and mutant structures complexed with DNA to derive the full benefit of these genetic studies.

ACKNOWLEDGEMENT

This work was supported by NIH grant GM26061.

REFERENCES

1. Zitomer,R.S. and Lowry,C.V. (1992) *Microbiol. Rev.*, **56**, 1–11.
2. Zitomer,R.S., Carrico,P. and Deckert,J. (1997) *Kidney Int.*, **51**, 507–513.
3. Balasubramanian,B., Lowry,C.V. and Zitomer,R.S. (1993) *Mol. Cell. Biol.*, **13**, 6071–6078.
4. Deckert,J., Rodriguez-Torres,A.M., Simon,J.T. and Zitomer,R.S. (1995) *Mol. Cell. Biol.*, **15**, 6109–6117.
5. Deckert,J., Rodriguez-Torres,A.M., Hwang,S.M., Kastaniotis,A.J. and Zitomer,R.S. (1998) *Genetics*, **150**, 1429–1441.
6. Grosschedl,R., Giese,K. and Pagel,J. (1994) *Trends Genet.*, **10**, 94–100.
7. Pevny,L.H. and Lovell-Badge,R. (1997) *Curr. Opin. Genet. Dev.*, **7**, 338–344.
8. Wolffe,A. (1994) *Science*, **264**, 1100–1101.
9. Werner,M.H., Huth,J.R., Gronenborn,A.M. and Clore,G.M. (1995) *Cell*, **81**, 705–714.
10. Love,J.J., Li,X., Case,D.A., Giese,K., Grosschedl,R. and Wright,P.E. (1995) *Nature*, **376**, 791–795.
11. Komachi,Y., Cheah,K.S.E. and Kondoh,H. (1999) *Mol. Cell. Biol.*, **19**, 107–120.
12. Deckert,J., Perini,R., Balasubramanian,B. and Zitomer,R.S. (1995) *Genetics*, **139**, 1149–1158.
13. Rose,M.D., Winston,F. and Hieter,P. (1990) *Methods in Yeast Genetics, A Laboratory Course Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
14. Chen,D.-C., Yang,B.-C. and Kuo,T.-T. (1992) *Curr. Genet.*, **21**, 83–84.
15. Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.M., Seidman,J.G., Smith,J.A. and Struhl,K. (1994) *Current Protocols in Molecular Biology*. John Wiley & Sons, New York, NY.
16. Gietz,R.D. and Sugino,A. (1988) *Gene*, **74**, 527–534.
17. Teo,S.-H., Grasser,K.D., Hardman,C.H., Broadhurst,R.W., Laue,E.D. and Thomas,J.O. (1995) *EMBO J.*, **14**, 3844–3853.
18. Berta,P., Hawkins,J.R., Sinclair,A.H., Taylor,A., Griffiths,B., Goodfellow,P.N. and Fellous,M. (1990) *Nature*, **348**, 448–450.
19. Hawkins,J.R., Taylor,A., Goodfellow,P.N., Migeon,C.J., Smith,K.D. and Berkovitz,G.D. (1992) *Am. J. Hum. Genet.*, **51**, 979–984.
20. Hawkins,J.R., Taylor,A., Berta,P., Levilliers,J., Van der Auwera,B. and Goodfellow,P.N. (1992) *Hum. Genet.*, **88**, 471–474.
21. Affara,N.A., Chalmers,I.J. and Ferguson-Smith,M.A. (1993) *Hum. Mol. Genet.*, **2**, 785–789.
22. Braun,A., Kammerer,S., Cleve,H., Lohrs,U., Schwarz,H.P. and Kuhnle,U. (1993) *Am. J. Hum. Genet.*, **52**, 578–585.
23. Goodfellow,P.N. and Lovell-Badge,R. (1993) *Annu. Rev. Genet.*, **27**, 71–92.
24. Pontiggia,A., Rimini,A., Harley,V.R., Goodfellow,P.N., Lovell-Badge,R. and Bianchi,M.E. (1994) *EMBO J.*, **13**, 6115–6124.