Genetic Analysis of Phage Mu Mor Protein Amino Acids Involved in DNA Minor Groove Binding and Conformational Changes*

Received for publication, June 8, 2011, and in revised form, August 14, 2011 Published, JBC Papers in Press, August 22, 2011, DOI 10.1074/jbc.M111.269860

Muthiah Kumaraswami1, Lakshmi Avanigadda1, Rajendra Rai1, Hee-Won Park1, and Martha M. Howe2

From the 1Department of Microbiology, Immunology, and Biochemistry, University of Tennessee Health Science Center, Memphis, Tennessee 38163 and the 2Department of Pharmacology, Structural Genomics Consortium, University of Toronto, Toronto, Ontario M5G1L7 Canada

Gene expression during lytic development of bacteriophage Mu occurs in three phases: early, middle, and late. Transcription from the middle promoter, Pm, requires the phage-encoded activator protein Mor and the bacterial RNA polymerase. The middle promoter has a −10 hexamer, but no −35 hexamer. Instead Pm has a hyphenated inverted repeat that serves as the Mor binding site overlapping the position of the missing −35 element. Mor binds to this site as a dimer and activates transcription by recruiting RNA polymerase. The crystal structure of the His-Mor dimer revealed three structural elements: an N-terminal dimerization domain, a C-terminal helix-turn-helix DNA-binding domain, and a β-strand linker between the two domains. We predicted that the highly conserved residues in and flanking the β-strand would be essential for the conformational flexibility and DNA minor groove binding by Mor. To test this hypothesis, we carried out single codon-specific mutagenesis with degenerate oligonucleotides. The amino acid substitutions were identified by DNA sequencing. The mutant proteins were characterized for their overexpression, solubility, DNA binding, and transcription activation. This analysis revealed that the Gly-Gly motif formed by Gly-65 and Gly-66 and the β-strand side chain of Tyr-70 are crucial for DNA binding by His-tagged Mor. Mutant proteins with substitutions at Gly-74 retained partial activity. Treatment with the minor groove- and GC-specific chemical chromomycin A3 demonstrated that chromomycin prevented His-Mor binding but could not disrupt a pre-formed His-Mor-DNA complex, consistent with the prediction that Mor interacts with the minor groove of the GC-rich spacer in the Mor binding site.

Bacteriophage Mu, the prototype for a large family of transposable phages, has two alternate life styles, lytic and lysogenic (1). Gene expression during the lytic cycle occurs in three phases: early, middle, and late (2, 3). The early promoter P0 has the characteristics of a typical bacterial promoter with both recognizable −10 and −35 hexamers, the promoter recognition elements for the Escherichia coli RNA polymerase (RNAP)3 with the “housekeeping” sigma factor, σ70 (4, 5). Pm possesses a −10 hexamer but lacks similarity to the consensus −35 element TTTGACA (Fig. 1A). As a result, gene expression from Pm requires a phage-encoded transcription activator protein, the middle operon regulator, Mor (6). Likewise, transcription from the Mu late promoters requires activation by the Mu C protein (7), which shares significant sequence similarity with Mor (6–9) (Fig. 1B).

Mor and C are the most well studied members of this Mor/C family of transcription factors. They are characterized by an acidic N terminus and a basic C terminus that contains a helix-turn-helix (HTH) DNA-binding motif (6). Mor is a homodimeric, sequence-specific DNA-binding protein with a monomer length of 129 amino acids (Fig. 1B) (6, 10). Consistent with its role as a transcription activator, Mor binds to the inverted repeats of a pseudo-palindromic DNA element located between −51 and −36 base pairs upstream of the transcription start site +1 (Fig. 1A) and overlapping the region normally occupied by the missing −35 recognition element (Fig. 1A) (11). In addition to Mor binding, activation of the middle promoter requires the C-terminal domains of the α and σ subunits of the E. coli RNA polymerase (12–14). Thus, the working model for middle promoter activation involves binding of Mor

1 Present address: Center for Molecular and Translational Human Infectious Diseases Research, The Methodist Hospital Research Institute, Houston, TX 77030.
2 To whom correspondence should be addressed: 858 Madison Ave., Memphis, TN 38163. Tel: 901-448-8215; Fax: 901-448-8462; E-mail: mhowe@uthsc.edu.

* This work was supported, in whole or in part, by National Institutes of Health Grant GM-35642, the Harriett S. Van Vleet Chair of Excellence in Virology, Harriett S. Van Vleet Chair of Excellence in Microbiology and Immunology, the College of Medicine at the University of Tennessee Health Science Center, and National Science Foundation Grants MCB-0318108 and MCB-0418108.

3 The abbreviations used are: RNAP, RNA polymerase; Pm, middle promoter; P0, early promoter; PlaccUV5, one late promoter; PlacUV5, a mutant promoter for lacZ; PlacSY5g6, a modified PlacUV5; Pm-lacZ, the middle promoter fused to the lacZ gene encoding β-galactosidase; Mor, middle operon regulator protein; His-Mor, the histidine-tagged Mor protein; his-mor, the gene encoding His-Mor; HTH, helix-turn-helix DNA binding motif; α1–α5, the α-helical portions of Mor protein numbered from the N terminus to the C terminus; ΔD69 and ΔDE3, two A phages with specific genes, ΔDE3 was derived from ΔD69 and carries the gene for phage T7 RNA polymerase under PlacUV5 control; IPTG, isopropyl-β-D-thiogalactopyranoside.
Conformational Changes and DNA Binding by Mu Mor Protein

The high-resolution crystal structure of a histidine-tagged Mor (His-Mor) dimer showed the locations of Mor amino acids 27 to 120; the His tag and Mor amino acids 1–26 and 121–129 were not visible (15). The dimer structure has three domains: an N-terminal dimerization domain and two C-terminal DNA-binding domains, each with a classical HTH motif (Fig. 2A) (15).

The N-terminal /H<sub>9251</sub>-helices, /H<sub>9251</sub>1 and /H<sub>9251</sub>2, of each His-Mor subunit intertwine with each other to form a four-helix bundle, creating a single central dimerization domain (Fig. 2A). The C-terminal three-helix bundle in each monomer forms the canonical HTH DNA-binding motif in which helix /H<sub>9251</sub>3 forms the “scaffolding helix,” and helices /H<sub>9251</sub>4 and /H<sub>9251</sub>5 together with the intervening loop form the HTH motif (Fig. 2A) (16, 17). The DNA-binding domains of the His-Mor dimer flank the dimerization domain, putting them at opposite ends of the structure (Fig. 2A).

Comparisons of the HTH domain structure with other proteins identified the trp repressor TrpR and region 4.0 of the /H<sub>9268</sub>-subunit of Thermus aquaticus RNAP as close structural neighbors (15, 18, 19). These proteins use a perpendicular “ends-on” mode of DNA binding, inserting their recognition helices into the major groove using only the first two turns of the recognition helix for DNA contacts. This led us to propose that His-Mor might bind similarly (15); however, the recognition helices of the His-Mor dimer are located too far apart (63 Å) to interact with two adjacent major grooves of B-DNA, which reach a maximum distance at the outside edges of 54 Å, indicating that conformational changes in both His-Mor and P<sub>m</sub> are probably required to

FIGURE 1. Promoter and protein sequences and plasmids. A, the sequence of the Mu middle promoter P<sub>m</sub> was annotated with the locations of the −10 hexamer (box), the transcription initiation site at +1 (bent arrow), the region protected from DNase I digestion by Mor binding (bar), and the imperfect dyad-symmetry element recognized by Mor (inverted arrows). Dots mark 10-bp intervals numbered by their positions in the promoter relative to +1. B, the Mor and C amino acid sequences are aligned with each other and with the secondary structure elements from the His-Mor crystal structure (15). The amino acids are identified by the single letter code. Gray shading indicates amino acids that are chemically similar at that position among the 15 total Mor/C family members previously identified (15). Black boxes indicate identical amino acids in those homologs (15). Dots indicate 10-amino acid intervals in Mor. C, primary sequence alignment of the inter-domain β-strand linker region of Mor and its previously identified homologs (15). The invariant residues are boxed in black and the chemically conserved amino acids are shaded gray. The numbers above the alignment indicate the positions of those amino acids in Mor. D, the two plasmids that comprise the in vivo His-Mor-P<sub>m</sub> activation assay are shown. Addition of IPTG leads to expression of His-Mor, which activates P<sub>m</sub>, whose activity is assayed by measuring β-galactosidase produced by the P<sub>m</sub>−lacZ fusion. Panel C was excerpted from Fig. 1D originally published in 2004 (15).
achieve DNA binding by His-Mor (15). Consistent with this prediction, circular permutation gel-shift assays with His-Mor bound to \(P_m\) sequences revealed DNA curvature with a \(45^\circ\) bending angle. Earlier we proposed that conformational changes in His-Mor would involve movement of the DNA-binding domains up and away from the dimerization domain to contact the DNA (Fig. 2, C and D).

The structure of the His-Mor dimer revealed an additional novel secondary structure element that might play a role in DNA binding (15). The anti-parallel \(\beta\)-strands at the top of the molecule (Fig. 2, A and B) form a \(\beta\)-ribbon in a 12-amino acid linker that connects the dimerization and DNA-binding domains. The residues in the \(\beta\)-strand and flanking loops of this linker are highly conserved, with 7 of 12 amino acids identical between Mor and C (Fig. 1, B and C). The \(\beta\)-strand amino acid side chains of Gln-68 and Tyr-70 (numbered as in native Mor) extend away from the protein (Fig. 2B). When the two HTH motifs make contacts in the major groove of the inverted repeats in \(P_m\) (Fig. 1A), the surface-exposed side chains of Gln-68 and Tyr-70 are ideally positioned to interact with the...
Conformational Changes and DNA Binding by Mu Mor Protein

minor groove of the GC-rich spacer between the inverted repeats (Fig. 1A).

In the crystal structure the side chains of residues Val-69, Ile-71, and Pro-72 are buried and involved in hydrophobic interactions in the dimer interface; they form a "cap-like" structure that seals the hydrophobic interior from the solvent (15).

This linker region contains three highly conserved glycines (Fig. 1C) that are predicted to form hinges that contribute to the conformational flexibility of Mor (15) (Fig. 2, B-D). The N-terminal amino acids of this hinge region form a "Gly-Gly-Gly" motif in which two of the glycines, Gly-65 and Gly-66, are highly conserved among Mor/C family members (Fig. 1C); whereas the C-terminal amino acids include only one conserved glycine, Gly-74 (Fig. 1C). Given their location between the two domains and the tolerance of glycine to extreme bond angles, we proposed that glycine hinges might act as pivot points for the conformational changes in Mor required for it to interact with the DNA (15).

In this study, the roles of the potentially key linker amino acids in His-Mor function were investigated by codon-specific site-directed mutagenesis of a his-tagged mor gene (his-mor). A two-plasmid transcription-activation phenotypic screening assay (Fig. 1D) for Pm–lacZ function was employed to identify candidate mutants with different levels of Pm activity, and their his-mor genes were sequenced. The mutant His-Mor proteins were then characterized for their overexpression, solubility, DNA binding, and transcription activation. The results presented here clearly demonstrate that multiple conserved amino acids in the β-strand and flanking loops of this linker are crucial for DNA binding by His-Mor.

EXPERIMENTAL PROCEDURES

Media, Chemicals, and Enzymes—Protein overexpression and routine cell growth were done in LB medium (20), whereas cultures for β-galactosidase assays were grown in minimal medium with casamino acids (M9CA) (7). MacConkey lactose plates with 25 g/liter of MacConkey agar and 25 g/liter of MacConkey agar base (Difco) and thus only half the normal amount of lactose was used for plate phenotyping. When necessary, chloramphenicol and ampicillin were used at 25 and 40 µg/ml, respectively. Oligonucleotides for mutagenesis, sequencing, and probe preparation were purchased from Integrated DNA Technologies, Inc.; their sequences are given in Table 1. Chromomycin A₃ was obtained from AG Scientific Inc. Sources for the remaining chemicals and enzymes are given in previous publications (10, 12, 15).

Bacterial Strains and Plasmids—The host strain background for most of the plasmid constructions and in vivo assays was E. coli K-12 strain MH13312 (mcrA Δpro-lac thY gyrA96 endA1 hsdR17 relA1 supE44 recA1 F' pro+ lacQF ΔlacZY), a derivative of JM109 (21) carrying an F' plasmid deleted for both lacZ and lacY and expressing higher than normal levels of Lac repressor (11). Strain MH13435 was made by transforming MH13312 with the Pm–lacZ fusion plasmid pIA14; it was used as a host for phenotypic assays of the ability of mutant Mor proteins to activate transcription from Pm. Strain MH13422 is a derivative of MH13312 containing both pIA14 and pIA69 and was used as the positive control in the MacConkey plate and β-galactosidase assays. Strain MH18211 was derived from MH13435 by transformation with the Mor deletion plasmid pMut10; it was used as the negative control in the MacConkey plate and β-galactosidase assays. Strain MH13355 was derived from JM109(DE3) and contains the same F' plasmid as MH13312. The ΔDE3 in MH13355 is a derivative of ΔDE9 containing the T7 RNA polymerase gene under control of the IPTG-inducible PlacUV5; it also carries imm21 and nin5 mutations (21). This strain was used for protein overexpression and detection of wild-type and mutant proteins.

Plasmid pIA14 contains wild-type Pm sequences from −62 to +10 (the base at −62 is provided by the vector) cloned between the EcoRI and BamHI sites in the lacZ reporter plasmid (Fig. 1D) such that β-galactosidase levels provide an indicator of Pm activity (11). Plasmid pIA69 contains sequences encoding His-Mor with introduced silent restriction sites; relevant sites are shown in Fig. 1D. His-Mor expression in pIA69 is under the control of both a T7 promoter and PmSyn (Fig. 1D), a slightly altered form of PlacUV5 (15). Amino acid changes were made in this his-mor gene; then the ability of the mutant protein to activate Pm was assayed by using the Pm–lacZ fusion in pIA14. The His-Mor deletion plasmid, pMut10, is a derivative of pIA69 deleted from PstI to SphI, missing Mor amino acids 8–116, and in the wrong reading frame after 116.

| Name       | Sequencea | Comments                                      |
|------------|-----------|-----------------------------------------------|
| MLK7       | cctgtgcgcggcgtcattcctgctgcacag             | Bottom strand primer, downstream of BamHI site of pIA14, for probe preparation |
| MUT215     | caggtctacattcgcctgacctgaccgccg            | Bottom strand primer to change Gln-68 to Asn or Lys |
| MUT216     | gtaaagcgcgtggaatgcttacactttagctt          | Top strand primer to change Gln-68 to Asn or Lys |
| MUT36      | gtcaagcgttacggtgactgtagctgtagctagctccta   | Bottom strand primer to mutate Gln-68 |
| MUT37      | gtcaagcgttacggtgactgtagctgtagctgtagctgtagctgtagctt | Bottom strand primer to mutate Gly-67 |
| MUT38      | gtcaagcgttacggtgactgtagctgtagctgtagctgtagctgtagctt | Bottom strand primer to mutate Gly-66 |
| MUT39      | gtcaagcgttacggtgactgtagctgtagctgtagctgtagctgtagctt | Bottom strand primer to mutate Gly-65 |
| MUT213     | ggcgtcagctgGNNatattccagcggtcga           | Top strand primer to mutate Tyr-70 |
| MUT214     | acgcgtggaacctGNNagctgacccggcgg           | Bottom strand primer to mutate Tyr-70 |
| MUT215     | caggtctacattcgcctgacctgaccgccg            | Bottom strand primer to mutate Gly-74 |
| MUT216     | gtaaagcgcgtggaatgcttacactttagctt          | Bottom strand primer to change Gln-68 to Asn or Lys |
| MUT680     | gtcaagcgttacggtgactgtagctgtagctgtagctgtagctt | Top strand sequencing primer for pIA69 |
| MUT215     | caggtctacattcgcctgacctgaccgccg            | Bottom strand sequencing primer for pIA69 |
| MUT216     | gtaaagcgcgtggaatgcttacactttagctt          | Top strand primer, upstream of EcoRI site in pIA14, for probe preparation |
| MUT36      | gtcaagcgttacggtgactgtagctgtagctgtagctgtagctgtagctt | Bottom strand primer, downstream of BamHI site of pIA14, for probe preparation |

a The oligonucleotide sequences are given from 5' to 3'. The letter S designates a mixture of G and C; the letter N indicates a mixture of A, T, G, and C. The upper case letters identify the codon being mutated.
Conformational Changes and DNA Binding by Mu Mor Protein

Site-directed Mutagenesis—Plasmid pIA69 was used as the template for the PCR-based mutagenesis of his-mor and as the vector for cloning the PCR products. Degenerate mutagenic primers (Table I) were designed to introduce base substitutions at all three positions of a single codon by using an equimolar mixture of bases (NNN or NNG + C) at the targeted codon (22). In most cases the degenerate primers were used for PCR with wild-type primers MUT13 or ZAO3 flanking the his-mor gene to create a library of mutant DNA fragments containing part of the his-mor gene with single codon mutations. The libraries were combined with overlapping wild-type PCR products containing the remaining part of the his-mor gene and were used as templates for overlapping PCR. In two cases both strands were mutagenized in separate PCR reactions and then combined and used as templates for overlapping PCR. After gel purification, the resulting mutagenized cassettes were cloned (using standard procedures; 23) into pIA69 between the PsiI and either HindIII or BamHI sites, in and following his-mor, respectively. The ligation mixtures were transformed as described previously (15) into strain MH13435 containing the Pm-lacZ fusion reporter plasmid, pIA14 (Fig. 1D) (11). In early constructions, the transformants were spread on LB plates with ampicillin (40 μg/ml) and chloramphenicol (25 μg/ml), and the resulting colonies were screened for Pm activity by picking and stabbing 48 to 96 colonies onto MacConkey lactose indicator plates with ampicillin (40 μg/ml) and chloramphenicol, and different concentrations of IPTG (10, 50, and 100 μM) to induce His-Mor expression. In later experiments transformants were plated directly onto the MacConkey lactose plates containing 50 μM IPTG. The direct plating gave much better resolution of distinct single colony phenotypes. Representative mutant plasmids were chosen based on their plate phenotypes, and the mutations were identified by automated DNA sequencing of their his-mor genes by the UTHSC Molecular Resource Center. For protein overexpression, mutant plasmids were transformed into MH13355 for His-Mor protein overproduction and purification.

Summary of Assays Used to Characterize Mutant Proteins—Overexpression of His-Mor proteins was achieved by using the T7 promoter in pIA69 and the T7 RNA polymerase provided by the MH13355 host. The expression levels and solubility of the mutant proteins were assessed by SDS-PAGE of supernatants from crude cell extracts that were made from those cells. Small scale His-Mor protein purification was accomplished using His tag affinity chromatography of the above supernatants. The ability of the mutant proteins to activate transcription was based on in vivo assays of β-galactosidase produced from a Pm-lacZ fusion plasmid. Detailed descriptions of these assays can be found in Ref. 15.

Probe Preparation and Gel Retardation Assay—The DNA-binding ability of the mutant His-Mor proteins was analyzed by using a gel retardation assay. The DNA probe was a PCR product made with 32P-labeled oligonucleotide MK7 and unlabeled oligonucleotide IR12 using pIA14 (containing Pm sequence – 62 to +10) as template. After purification with a QiaQuick PCR purification kit (Qiagen), the probe concentration was estimated on a 2% agarose gel by comparison with a low mass DNA ladder (Invitrogen). A 20-μl reaction volume containing 400 pg of probe, 50 ng of calf thymus DNA, and 200, 400, or 800 ng of His-Mor protein in binding buffer (20 mM Tris-HCl, pH 7.9, 50 mM NaCl, 5% glycerol and 1 mM DTT) was incubated at room temperature for 20 min, then resolved on a 10% nondenaturing acrylamide gel containing 0.5 × TBE and run in 0.5 × TBE buffer at 260 V for 3 h at 4°C. Initial exposure of the gels to X-Omat Bio-Max MR film was done without drying.

Chromomycin Assays—Chromomycin interference and disruption assays were performed by using the probe preparation described above and chromomycin A3 from A.G. Scientific Inc. For the interference assay, different concentrations of chromomycin (final concentrations 0.5–50 μM) were incubated with the probe (~400 pg) in a 20-μl reaction volume at room temperature for 15 min followed by addition of 800 ng of purified His-Mor (1.18 μM dimer) and incubation at room temperature for an additional 20 min. For the disruption assay, the effect of chromomycin on preformed Mor-DNA complexes was assayed similarly except that the order of Mor and chromomycin addition was reversed. The reaction mixtures were resolved on a 10% native polyacrylamide gel as described above, and the gels were exposed to X-Omat-MR film without drying but with an intensifier screen for 16 h.

RESULTS

Rationale—To demonstrate that a particular amino acid is essential for protein function, one could construct a set of mutants, one for each of the other 19 amino acid substitutions, assay them, and show that none retain protein function. The mutants could be made by brute force, one mutant at a time, or more efficiently by making a library of substitutions at a single codon by using a primer with degeneracy at all three positions of the codon, and then screening clones until the 19 mutants have been found and tested for function. An even more efficient approach is to use an indicator plate assay for screening the library, where functional and nonfunctional phenotypes are easily discriminated. Sequencing of the target gene from phenotypically functional colonies will identify substitutions that retain protein function. If the library is diverse and all the phenotypically functional colonies have the wild-type protein sequence, then one can conclude that the amino acid at that position is essential. To demonstrate that the library is sufficiently diverse to contain all 19 substitutions, the genes from seven to 10 nonfunctional colonies are sequenced; they should all contain different single amino acid substitutions.

Mutant Isolation—To determine the importance and roles of the inter-domain linker amino acids in His-Mor function, we carried out site-directed mutagenesis by PCR with oligonucleotide primers containing degeneracy at all three positions of a single codon, i.e. oligonucleotides designed to introduce all possible amino acid substitutions at a given position (numbered as in native Mor). The mutated DNA fragments were cloned into the His-Mor expression plasmid pIA69 (Fig. 1D) and transformed into strain MH13435 containing the reporter Pm-lacZ fusion plasmid pIA14 (Fig. 1D). The resulting libraries of transformants containing mutant His-Mor proteins were screened on MacConkey lactose plates containing small amounts (50 μM) of IPTG to induce low levels of His-Mor expression. The
TABLE 2

Mutant colony phenotypes
Transformants containing the pIA69-derived plasmids with codon-specific amino acid (aa) substitutions and the P<sub>mw</sub>-lacZ reporter plasmid were plated directly or streaked for single colonies on MacConkey lactose plates, incubated overnight, and then scored for their colony colors. The pIA69-derived plasmids were sequenced to determine their amino acid substitutions, which are indicated below.

| WT Mor amino acid | Frequency (%) of colonies with mutant phenotypes | Substitutions making white or pale pink colonies | Substitutions making white colonies with red centers | Substitutions making red colonies |
|-------------------|---------------------------------------------|-----------------------------------------------|-----------------------------------------------|--------------------------------|
| Gln-68            | 63                                          | Ile, Phe, Gly, Leu, Pro, Arg, Lys             | None found                                   | No aa change                   |
| Tye-70            | 71                                          | Ile, Val, Ser, Gly, Arg                       | None found                                   | No aa change                   |
| Gly-65            | 91                                          | Arg, Met, Asn, Val, Thr                       | None found                                   | No aa change                   |
| Gly-66            | 94                                          | Pro, His, Leu, Ala, Arg                       | None found                                   | No aa change                   |
| Gly-74            | 56                                          | Arg, Thr<sup>a</sup>, Tyr<sup>a</sup>, Ser<sup>a</sup>, Ile<sup>a</sup> | Trp                                           | Phe                            |

<sup>a</sup> If incubated longer, these mutant colonies became white with red centers.

Conformational Changes and DNA Binding by Mu Mor Protein

The colony color then reflected the ability of its mutant His-Mor protein to activate transcription from P<sub>mw</sub>. Transformants were initially scored in a “pick and stab” assay as defective (white or very light pink) or functional (red) by comparison to the phenotypes of strain MH18211 containing a deletion in His-Mor (Δhis-mor) and strain MH13422 with wild-type His-Mor, respectively. Subsequent re-testing by streaking for single colonies, and by direct plating of the libraries on the MacConkey plates allowed us to identify two additional intermediate phenotypes, diffuse pale pink colonies and white colonies with red centers. The phenotypes being reported here (Table 2) were derived from this isolated colony assay. Libraries with substitutions in the most highly conserved positions Gly-65, Gly-66, and Tye-70 produced 91, 94, and 71% white colonies, respectively. DNA sequencing showed that the plasmids from these colonies contained his-mor genes encoding single amino acid substitutions at the targeted positions. When the his-mor genes in the red colonies were sequenced, they had no mutations. In contrast, libraries with substitutions in the less highly conserved positions Gln-68 and Gly-74 had only 63 and 56% mutant phenotypes and exhibited a range of mutant colony colors including white, pale pink, and white with red centers. For these positions we chose representatives from each phenotype for sequencing. The Gly-74 mutants were unusual in that one red colony and one white colony with a red center each contained single amino acid substitutions, demonstrating that Gly-74 is not essential. In addition, multiple mutants from the Gly-74 library made white colonies, but those colonies acquired red centers upon continued incubation; these colonies made His-Mor mutant proteins that retained partial function. The amino acid substitutions and colony phenotypes for the above mutants are given in Table 2. Finally, to validate this genetic approach we also made libraries for the nonconserved position Gly-67. For these libraries only 8% of the colonies had a mutant phenotype. Beyond sequencing to demonstrate the presence of mutations, analysis of these mutants was not pursued.

**Protein Overexpression**—Amino acid substitutions in a protein can cause misfolding and lead to aggregation, precipitation, and increased sensitivity to proteolytic degradation. These consequences should be detectable by a reduction in levels of His-Mor protein in supernatants from crude cell extracts, as observed previously for several mutant proteins with substitutions in the dimerization domain (15). To ensure that the severe defects of these inter-domain linker mutants were not caused by precipitation or degradation of the mutant proteins, we assayed for the presence of overexpressed mutant protein in the supernatants from sonicated cell extracts by SDS-polyacrylamide gel electrophoresis. The gels in Fig. 3 show that the mutant proteins were overexpressed and present at levels comparable with that of wild-type His-Mor. Given the chemical diversity of the substitutions (Table 2) and their negligible effects on protein solubility and degradation (Fig. 3), we favor the interpretation that the defective phenotypes caused by the majority of these mutations are due to their interference with protein function.

**DNA Binding**—Representative mutant His-Mor proteins were purified by His tag affinity chromatography and tested for their ability to bind to a wild-type middle promoter DNA fragment by using a gel retardation assay (Fig. 4). Substitutions at the most highly conserved positions Gly-65, Gly-66, and Tye-70 abolished DNA binding (Fig. 4), making it most likely that their defects in transcription activation are due to their inability to bind to promoter DNA. Substitutions in Gln-68 also prevented DNA binding with the exception of Q68R, which retained some ability to bind DNA but failed to activate transcription. Mutations in the less conserved Gly-74 position conferred less defective phenotypes, and almost half of their mutant His-Mor proteins retained at least partial function for DNA binding and for transcription activation as observed in the colony color assay.
Intriguingly, two substitutions at Gly-74 that exhibited significant DNA-binding activity were aromatic substitutions, G74W and G74F, whereas those with other substitutions, G74Y and G74S, failed to bind DNA.

Transcription Activation—To quantify the transcription activation ability of the mutant His-Mor proteins in vivo, we performed liquid β-galactosidase assays with the Pm-lacZ fusion as the reporter. The results are shown in Table 3. Consistent with their DNA-binding defects, mutants with substitutions at the most conserved positions Gly-65, Gly-66, and Tyr-70 were extremely defective, with β-galactosidase levels comparable with those of the Mor deletion (<10 units). The

### Table 3

Summary of β-galactosidase and DNA binding by mutant proteins

The wild-type His-Mor amino acids predicted to be involved in conformational change and DNA minor groove binding were subjected to codon-specific mutagenesis, and the mutants were assayed for the effects of the mutations on DNA binding and transcription activation at Pm as described earlier. This summary presents the location and identity of the key wild-type amino acids in His-Mor in the top row. The amino acid changes that were recovered and characterized for each wild-type position are listed under column A. The β-galactosidase activities produced by transcription activation of the Pm-lacZ fusion by the mutant protein are given under column B, and are expressed relative to the wild-type level assayed in parallel and set to 1000 Miller units (53); they represent the average of results from either two (Gln-68 and Tyr-70) or three (all three glycines) independent experiments. The DNA-binding activity for each mutant is given in column D.

| WT Mor Gln-68 | WT Mor Tyr-70 | WT Mor Gly-65 | WT Mor Gly-66 | WT Mor Gly-74 |
|---------------|---------------|---------------|---------------|---------------|
| A | B | D | A | B | D | A | B | D | A | B | D |
| Phe | <10 | -- | Arg | <10 | -- | Pro | <10 | ND | Tyr | 46 | -- |
| Gln | <10 | -- | Val | <10 | -- | Met | <10 | ND | His | <10 | -- |
| Arg | <10 | -- | Arg | <10 | -- | Leu | <10 | ND | Leu | <10 | ND |
| Leu | <10 | -- | Ser | <10 | -- | Asn | <10 | ND | Ala | <10 | -- |
| Phe | <10 | -- | Gly | <10 | -- | Val | <10 | ND | Trp | 715 | ++ |
| Ile | <10 | ND | Arg | ND | -- | Thr | <10 | -- | Tyr | 46 | -- |
| Val | <10 | ND | Arg | ND | -- | Phe | <10 | ND | Ser | 68 | -- |
| Leu | <10 | ND | Arg | ND | -- | Phe | 592 | + | Ser | 68 | -- |
| Thr | 17 | ND | Thr | 17 | ND | Phe | 592 | + | Ser | 68 | -- |

* A minus sign (−) indicates that no DNA binding was detected in gel retardation experiments such as those shown in Fig. 4.

* ND indicates that DNA binding was not tested for that particular mutant.

* For mutants that gave detectable DNA binding, the degree of binding was indicated by one or more "+" signs, with "++++" set as the wild-type level, "++" indicating strong but definitely reduced binding at the lowest protein concentration, "+" with DNA binding only at the two highest protein concentrations, and "+" significantly reduced binding even at the two highest protein concentrations. Mutant Y70Q is not included in this summary because it was recently found to contain a second mutation.

Conformational Changes and DNA Binding by Mu Mor Protein

**FIGURE 4. Gel retardation assays testing for DNA binding.** Labeled probe containing Pm sequences −62 to +10 was made by PCR. Each gel shows the DNA binding results for wild-type (WT) and three mutant His-Mor proteins, with their identities indicated above the brackets over the three lanes for each mutant protein. Lanes marked P show the probe-alone control without any His-Mor protein. The amounts of purified protein used in each lane are indicated by "+" signs, with (+) representing 200 ng, (++) 400 ng, and (+++) 800 ng. The arrows indicate the migration positions for DNA with bound His-Mor (B) and free probe (F), respectively.
TABLE 4
Tally of amino acid substitutions
The Mor amino acid (WT Mort) sequence was subjected to a BLAST (24) similarity search, identifying a total of 90 proteins (including Mor), which were then directly aligned on the NCBI server using COBALT (26). For Mort amino acids Gly-65 through Ala-76, which comprise the interdomain linker, the frequencies of each amino acid at each position were tallied. The identity of each amino acid substitution and the number of times that amino acid was present at that position in the 90 proteins are listed below each Mor amino acid. The amino acids shown in italics were found among the highly defective mutants in Table 2.

| Mor amino acids | Gly-65 | Gly-66 | Gly-67 | Gln-68 | Val-69 | Tyr-70 | Ile-71 | Pro-72 | Arg-73 | Gly-74 | Gln-75 | Ala-76 |
|----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Gly (85)       | Gly (89)| Gln (36)| Asn (27)| Val (38)| Tyr (90)| Phe (40)| Pro (89)| Lys (25)| Gly (75)| Asp (18)| Ala (28) |
| Asn (3)        | Asp (1) | Arg (19)| Gln (24)| Ile (23)| Ile (25)| Val (1) | Met (24)| Asp (4) | Ile (13)| Arg (13) |
| Ala (1)        | Leu (8) | Ser (14)| Phe (15)| Met (2) | Leu (22)| Asp (9) | Ala (3) | Met (12)| Gly (7) |
| Arg (1)        | Gly (6) | Asp (6) | Tyr (9) | Val (1) | Asn (9) | Ala (3) | Val (7) | Met (12)| Gly (7) |
| Tyr (6)        | Ala (4) | Leu (3) | Met (1) | Val (1) | Ala (5) | Asp (2)| Leu (9)| Lys (7) |
| Lys (6)        | Leu (4) | Met (1) | Gly (3) | Val (3) | Thr (3) | Tyr (1)| Val (4) | Ala (5) |
| Glu (3)        | Val (3) | Gln (1) | His (1) | Ile (2) | Leu (1)| Ser (1) | Ser (2)| Pro (2) |
| Met (2)        | Gly (3) | His (1) | Thr (1) | Thr (1) | Ala (1)| Glu (1)| Thr (1)| Arg (1) |
| His (1)        | Ile (2) | Ala (1)| Phe (1)| His (1) | Arg (1)| Arg (1)| Ile (1)| Glu (2) |
| Thr (1)        | Thr (1) | Ala (1)| Phe (1)| His (1) | Arg (1)| Arg (1)| Ile (1)| Glu (2) |
| Ala (1)        | Glu (1) | Ala (1)| Phe (1)| His (1) | Arg (1)| Arg (1)| Ile (1)| Glu (2) |
| Phe (1)        | His (1) | Ala (1)| Phe (1)| His (1) | Arg (1)| Arg (1)| Ile (1)| Glu (2) |

majority of Gln-68 mutants were similarly defective. Interestingly, for position Gly-74, the same two mutant proteins with aromatic substitutions, G74W and G74F, that retained some DNA-binding activity also exhibited more than 50% of wild-type levels of transcription activation. A summary of the β-galactosidase and DNA binding properties for the mutant proteins is given in Table 3.

Natural Amino Acid Substitutions—A second complementary approach to identifying amino acids that are likely to be important for protein function is to examine the degree to which they are conserved in members of a protein family. In 2003 a BLAST similarity search (24) of the GenBank protein data base (25) identified 15 Mor/C family members (Fig. 1C, 15). In late 2010 a similar single-round BLAST search identified 90 family members that were then directly aligned on the NCBI server by using COBALT (26). The majority of the family members were located in complete prophage or prophage remnants in bacterial genome sequences. Although it is not feasible for us to determine which of these proteins function, it is possible to tally the amino acids at each position and identify those that occur frequently and others that are relatively rare. The rationale for this analysis is that amino acids that occur frequently are reasonably likely to retain function; whereas amino acids present rarely would more likely occur in mutant proteins that do not function.

Examination of the tally results shown in Table 4 reveals that the amino acids that were highly conserved in the earlier 15-member family (Fig. 1C) are extremely highly conserved in the current 90-member family (Table 4). They are Gly-65, Gly-66, Tyr-70, and Pro-72; at these positions there are only 3, 1, 0, and 1 nonconserved amino acids, respectively. For Gly-65 two of the three exceptions were found in the initial mutant isolation and produced white colonies and <10 units of β-galactosidase, supporting our hypothesis that rare amino acid changes result in nonfunctional mutant proteins. These exceptions also lead us to suspect that amino acids present three or fewer times are likely to be in defective proteins. The BLAST results for Gly-74 show that it is highly conserved, with Gly in 75 of 90 members, and the other 15 distributed among 7 different amino acids, with a maximum of four times for a single amino acid.

These data alone would have led us to predict that these other 15 proteins would be entirely nonfunctional, but the colony colors in Table 2 and β-galactosidase values in Table 3 indicate that most of them retain partial Mort function. For Gln-68, Asn is the most frequent amino acid, being present 27 times, with Gln found 24 times and Ser found 14 times; the remainder of the substitutions were present from 1 to 6 times and contain Leu, Gly, and Ile substitutions, which were also found among the highly defective mutants (Table 2). Because Asn occurred so frequently, we constructed an Asn substitution mutant at position 68 of His-Mor; in contrast with the other mutations it made white colonies with red centers, retaining partial His-Mor function. Positions Val-69, Ile-71, and Pro-72 were not mutated because of their role in the hydrophobic cap on the dimerization domain. The BLAST results show that other hydrophobic amino acids are frequently present at positions Val-69 and Ile-71, consistent with that role.

Effects of Chromomycin—Modeling studies demonstrated that the side chains of β-strand amino acids Gln-68 and Tyr-70 are favorably placed to interact with the minor groove of the GC-rich spacer between the inverted repeats in the Mort binding site. To investigate the significance of the proposed minor groove contacts to Mort-Pm interactions, we used gel mobility shift assays to test whether His-Mor could bind to the middle promoter in the presence of the GC-specific minor groove binding ligand, chromomycin A₃, at concentrations known to inhibit binding of other proteins (27, 28). In the first experiment chromomycin was incubated with the promoter DNA and then followed by addition of His-Mor. Chromomycin at a concentration as low as 2 μM completely inhibited His-Mor binding to the promoter DNA and reduced Mor binding at 0.5 and 1 μM; only the chromomycin-DNA complex was present at and above, 2 μM (Fig. 5A).

To ask whether chromomycin would be able to access the minor groove and displace His-Mor from a His-Mor-DNA complex, as observed previously for an EGR1-DNA complex (28), we conducted a second experiment in which the order of addition of His-Mor and chromomycin was reversed. When chromomycin was added to preformed His-Mor-DNA complexes, it was unable to completely disrupt the complexes even

Conformational Changes and DNA Binding by Mu Mor Protein
Conformational Changes and DNA Binding by Mu Mor Protein

FIGURE 5. Inhibition and disruption of His-Mor-DNA complexes by chromomycin A₃. Labeled probe (400 pg) containing Pₖ sequences −62 to +10 was used with 800 ng of His-Mor dimer (1.18 μM) and the micromolar concentrations of chromomycin indicated above each lane. In panel A chromomycin was incubated with the DNA probe prior to addition of His-Mor. In panel B His-Mor was incubated with the DNA probe before adding chromomycin. Lane P contains DNA probe alone. Arrows labeled B designate the migration position of the bound His-Mor-DNA complex; arrows labeled C indicate the position of the chromomycin-DNA complex; and arrows labeled F indicate the position of free probe.

at concentrations as high as 50 μM (Fig. 5B), suggesting that binding of chromomycin and His-Mor is mutually exclusive, as it would be if they both interact with the minor groove. Similar results were obtained when the same experiments were performed with short oligo duplexes containing Pₖ sequences from −33 to −54 only, and thus including only the single minor groove and two flanking major grooves of the His-Mor binding site (data not shown). These results are consistent with our hypothesis that His-Mor interacts with the intervening GC-rich minor groove.

DISCUSSION

The crystal structure of His-Mor revealed a novel structural element, a 12-amino acid inter-domain linker containing a β-strand, which we proposed serves as a hinge needed for DNA binding of His-Mor (15). The results from this genetic analysis demonstrate that the highly conserved Tyr-70 residue in the β-strand and the Gly-65, Gly-66, and Gly-74 residues in the flanking loops play a significant role in DNA binding and, thus, transcription activation by His-Mor.

The Roles of Gln-68 and Tyr-70—Docking of the crystal structure of His-Mor onto a 16-bp long B-DNA, permitting the HTH motifs to contact the DNA major grooves, revealed that the side chains of β-strand residues Gln-68 and Tyr-70 would be brought into close proximity to the minor groove and, therefore, might participate in DNA binding by His-Mor. The severe DNA-binding and transcription-activation defects of the mutant proteins with most substitutions at Gln-68 and Tyr-70 are consistent with this hypothesis. One possibility is that the Tyr-70 side chain intercalates into the DNA, causing the observed DNA bend. Tyrosine intercalation is found in multiple DNA-binding proteins, including bovine pancreatic DNase I (29), human 3-methyladenine DNA glycosylase (30), and MutY of Bacillus stearothermophilus (31) leading to DNA bends of 20, 22, and 55 degrees, respectively. It seems more likely that the amide side chain of Gln-68, and its chemically conserved substitution Asn-68, play a stabilizing role in interactions with the negatively charged sugar-phosphate backbone of the DNA.

The Role of Flexible Glycines in DNA Binding by His-Mor—In the DNA-free crystal structure of His-Mor, the tips of the predicted recognition helices of the HTH motif are unfavorably positioned to interact with two adjacent major grooves of DNA (15) (Fig. 2C). In particular, they are 63 Å apart, which is ~29 Å farther than the distance between the centers, and ~9 Å farther than the outermost edges, of two adjacent major grooves of B-form DNA. Structural changes that would allow His-Mor to assume an altered DNA-bound conformation were predicted to stem from the flexibility provided by the three highly conserved glycines flanking the β-strands (15). The mutational analysis presented here demonstrates that these glycines, Gly-65, Gly-66, and Gly-74, are important for His-Mor function. It also revealed the unexpected finding that some substitutions at Gly-74 retain partial His-Mor function. Given the lack of side chains in glycines and their known roles in the flexibility of other proteins (32–35), it is more likely that they play a role in the conformational change in His-Mor to allow it to bind to DNA than in making direct DNA contacts.

Minor Groove DNA Binding by His-Mor—Sequence-specific DNA-binding proteins generally make base-specific contacts in the major groove due to its better accessibility and greater information content (36–38). Other proteins such as TATA-binding proteins and proteins with high mobility group domains, e.g. sex-determining region Y and Lef-1, exclusively contact bases in the minor groove (39–41). Finally, some transcription regulators make simultaneous contacts with both the major and minor grooves (17, 42–44). For example, proteins containing winged HTH motifs use their HTH motifs for major groove binding, and their “wings” to probe the flanking minor grooves (45–48). Similarly, transcription regulators in the LacI family utilize an HTH motif for major groove interactions and a hinge region connecting the domains to contact the intervening minor groove (43, 49). Despite the variation in the secondary structure elements making the minor groove contacts, these proteins induce DNA unwinding and bending toward the major groove by inserting amino acid side chains into the minor groove. In the case of DNase I (29) and human 3-methyladenine DNA glycosylase (30), it is tyrosine intercalation into the minor groove that results in widening of the minor groove, narrowing of the major groove and bending of the DNA toward the major groove.

Based on the crystal structure of the DNA-free His-Mor dimer, we proposed that His-Mor also uses two elements for DNA binding: the HTH motifs for major groove interactions and the inter-domain linker for binding to the intervening minor groove on the same face of the DNA. To address this possibility we assayed the effect of the GC-specific minor groove binding drug chromomycin A₃ on His-Mor-DNA inter-
actions. Consistent with our prediction, chromomycin was able to prevent His-Mor from binding to the DNA, but it is not clear that this occurred due to minor groove binding by His-Mor or narrowing of the adjacent major grooves as a consequence of chromomycin binding. Chromomycin did not inhibit DNA binding by the HSV ICP4 (infected-cell polypeptide 4) (earlier called IE175) major transcriptional regulator (50, 51), which from bioinformatic analyses is predicted to bind in the major groove via a helix-turn-helix DNA binding motif (52). However, chromomycin did prevent binding of proteins EGR1 and WT1 (Wilms tumor 1) known to bind in the major groove (28). It is not clear whether this is due to groove width changes from the widened minor groove causing narrowing of the major groove or because of structural changes it causes in the DNA immediately flanking its binding site (27, 53).

When we tested whether preformed His-Mor-DNA complexes could be disrupted by chromomycin, they were not, i.e. chromomycin was unable to access the GC-rich minor groove in preformed His-Mor-DNA complexes even at very high concentrations (50 μM). Just the opposite was observed for the zinc finger, major groove binding, transcription factor EGR1 (28). When chromomycin was added to preformed EGR1-DNA complexes, it was able to access the minor groove and disrupt the preformed complexes, even at the very low chromomycin concentrations (~1 μM) that blocked binding of EGR1 to its DNA-binding site (28). We suggest that the inability of chromomycin to disrupt the preformed His-Mor-DNA complexes indicates that chromomycin cannot access the minor groove when His-Mor is bound and, thus supports the hypothesis that Mor interacts with the minor groove as well as the two flanking major grooves. Providing additional support, binding of the Mu C protein to the late promoter, P\textsubscript{trans}, blocked the minor groove accessibility to the minor groove-specific chemical nuclease, 1,10-phenanthroline (54). Taken together, these results support our hypothesis that His-Mor interactions with the minor groove of the GC-rich spacer in the Mor-binding site are essential for His-Mor-DNA binding and that the minor groove becomes inaccessible upon His-Mor binding.

A Role for Pro-72 in His-Mor Function—In the crystal structure of the His-Mor dimer the hydrophobic residues Val-69, Ile-71, and Pro-72 were found in a “cap-like” structure at the top of the dimerization interface with their side chains extending down into the hydrophobic core (15). We thought that these residues would serve the same role in DNA-bound His-Mor, so we did not include them in the mutational analysis. The natural hydrophobic substitutions observed in the BLAST results for Val-69 and Ile-71 are consistent with this hypothesis. The essential nature of Pro-72, however, forces us to consider whether Pro-72 plays a different role in the DNA-bound form of His-Mor. Proline is unique among naturally occurring amino acids in the cyclization of its side chain to the backbone amide and its ability to adopt two conformations, \textit{cis} and \textit{trans}. It is known to function as a molecular hinge (55), mediating conformational transitions in a number of proteins (55–58). In His-Mor, Pro-72 is located at the C-terminal end of the β-strand (15), and its isomerization could act as a molecular switch inducing structural changes in His-Mor to confer, in combination with the conserved glycines, the DNA-bound form of His-Mor. Alternatively, the above structural changes in His-Mor could be independent of Pro-72 but its side chain might participate directly in His-Mor-DNA interactions, either by intercalation between bases or by interactions with the sugar-phosphate DNA backbone. Proline intercalation and the resulting ~90° DNA bend has been well documented in the bacterial histone-like proteins, integration host factor and HU, and also in eukaryotic high mobility group proteins (36, 59, 60). The smaller, ~45°, bend observed for His-Mor might occur if there is only partial intercalation. On the other hand, interactions between the proline side chain and sugar phosphate backbone does occur, for example, in DNA binding by the \textit{E. coli} transcription regulator PutA (61). Thus, similar interactions between the side chain of Pro-72 and the sugar-phosphate DNA backbone might participate in His-Mor-DNA interactions and play a role in stabilizing the bending angle generated by His-Mor binding. Finally, it is also possible that Tyr-70 and Pro-72 both interact within the minor groove. There is precedence for such interactions, for example, in the 3-methyladenine DNA glycosylase (30) Tyr-162 intercalates between the bases, resulting in widening of the minor groove, which is then filled by Met-164 and Tyr-165 (30).

Conclusion—Results from this mutational and homolog identification analysis favor the model for middle promoter binding by His-Mor in which the conformational changes in His-Mor stemming from the inter-domain linker region and the structural changes in DNA due to His-Mor-minor groove interactions are crucial for DNA binding and transcription activation by His-Mor. Given the high degree of sequence conservation in the β-strand linker region among the Mor/C family members, this mode of DNA binding is likely to be common in this family.

Acknowledgment—We acknowledge the Molecular Resource Center of the University of Tennessee Health Science Center for determining the DNA sequences reported herein.

REFERENCES

1. Paolozzi, L., and Ghelardini, P. (2006) in \textit{The Bacteriophages} (Calendar, R., ed) pp. 469–496, Oxford University Press, Oxford
2. Stoddard, S. F., and Howe, M. M. (1989) \textit{J. Bacteriol.} 171, 3440–3448
3. Margolin, W., and Howe, M. M. (1990) \textit{Virology} 174, 192–203
4. Krause, H. M., and Higgins, N. P. (1986) \textit{J. Biol. Chem.} 261, 3744–3752
5. Hawley, D. K., and McClure, W. R. (1983) \textit{Nucleic Acids Res.} 11, 2237–2255
6. Mathee, K., and Howe, M. M. (1990) \textit{J. Bacteriol.} 172, 6641–6650
7. Margolin, W., Rao, G., and Howe, M. M. (1989) \textit{J. Bacteriol.} 171, 2003–2018
8. Margolin, W., and Howe, M. M. (1986) \textit{Nucleic Acids Res.} 14, 4881–4897
9. Margolin, W., and Howe, M. M. (1990) \textit{J. Bacteriol.} 172, 1424–1429
10. Kahmeyer-Gabbe, M., and Howe, M. M. (1996) \textit{J. Bacteriol.} 178, 1585–1592
11. Artsimovitch, I., and Howe, M. M. (1996) \textit{Nucleic Acids Res.} 24, 450–457
12. Ma, J., and Howe, M. M. (2004) \textit{J. Bacteriol.} 186, 7858–7864
13. Artisimovitch, I., Murakami, K., Ishihama, A., and Howe, M. M. (1996) \textit{J. Biol. Chem.} 271, 32343–32348
14. Artisimovitch, I., Kahmeyer-Gabbe, M., and Howe, M. M. (1996) \textit{Proc. Natl. Acad. Sci. U.S.A.} 93, 9408–9413
15. Kumaraswami, M., Howe, M. M., and Park, H. W. (2004) \textit{J. Biol. Chem.} 279, 16581–16590
16. Brennan, R. G., and Matthews, B. W. (1989) \textit{J. Biol. Chem.} 264, 1903–1906
