Potential for H-DNA in the Human MUC1 Mucin Gene Promoter*

(Received for publication, February 20, 1996, and in revised form, May 9, 1996)

Kristie L. Nelson, Nicole A. Becker, Gurcharan S. Pahwa, Michael A. Hollingsworth, and L. James Maher, III§

From the Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, Omaha, Nebraska 68198-6805 and the Department of Biochemistry and Molecular Biology, Mayo Foundation, Rochester, Minnesota 55905.

Similar imperfect purine/pyrimidine mirror repeat (PMR) elements have previously been identified upstream of the human MUC1 mucin and CFTR genes. These elements confer S1 nuclease sensitivity on isolated plasmid DNA at low pH. We now present a detailed characterization of the non-B DNA structure responsible for S1 nuclease sensitivity upstream of the MUC1 gene. A 90-base pair (bp) DNA fragment containing a 32-bp M-PMR element termed M-PMR3 was subcloned into a recombinant vector. This fragment conferred S1 nuclease sensitivity on the resulting supercoiled plasmid. High resolution mapping of sites reactive to S1 and P1 nuclease demonstrates that cleavage occurs within the M-PMR3 element. High resolution mapping with chemical agents selective for non-B DNA provides evidence that M-PMR3 adopts an H-DNA structure (intramolecular triple helix) in the less common H-y5 isomer at low pH. This result is observed in the presence or absence of Mg2++. Mutation of the native M-PMR3 element to create perfect homopurine/homopyrimidine mirror symmetry alters the preferred folding to the more common H-y3 tripleplex DNA isomer. These results demonstrate that imperfections in mirror symmetry can alter the relative stabilities of different H-DNA isomers.

The human MUC1 gene encodes the core protein of a mucin-like molecule that is constitutively expressed by differentiated secretory epithelial cells of the breast, pancreas, prostate, lungs, bladder, vas deferens, and other organs (1). The MUC1 gene is overexpressed (relative to the normal epithelial cells) by many adenocarcinomas that arise in these organ sites (1–3). Although their patterns of expression are not identical, MUC1 is expressed by several cell types that also express the cystic fibrosis transmembrane conductance regulator; bp, base pair(s); CAA, chloroacetaldehyde; DEP, diethylpyrocarbonate; DMS, dimethyl sulfate; MOPS, 4-morpholinopropanesulfonic acid; M-PMR3, 32-bp PMR element centered at 118 bp relative to the transcription start point of the human MUC1 gene; OT, osmium tetroxide; PMR, purine/pyrimidine mirror repeat; R, A or G; Y, C or T.

*This work was supported by the Mayo Foundation and by Grants DK44762, CA57362, GM47814, and CA36727 from the National Institutes of Health, by Grant SIG 16 from the American Cancer Society, by Grant 5-P30-CA36727-08 from the National Cancer Institute, and a Junior Faculty Research award (to L. J. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Dept. of Biochemistry and Molecular Biology, Mayo Foundation, 200 First St. SW, Rochester, MN 55905. Tel.: 507-284-9041; Fax: 507-285-2053; E-mail: maher@mayo.edu.

‡ The abbreviations used are: CFTR, cystic fibrosis transmembrane conductance regulator; bp, base pair(s); CAA, chloroacetaldehyde; DEP, diethylpyrocarbonate; DMS, dimethyl sulfate; MOPS, 4-morpholinopropanesulfonic acid; M-PMR3, 32-bp PMR element centered at 118 bp relative to the transcription start point of the human MUC1 gene; OT, osmium tetroxide; PMR, purine/pyrimidine mirror repeat; R, A or G; Y, C or T.

epithelial cell types in which MUC1 is expressed may be of use in the design of vectors for human gene therapy of adenocarcinomas, cystic fibrosis, and other diseases of secretory epithelia. Some basal elements that regulate transcription of MUC1 have been identified (6–8); however, the elements responsible for tissue-specific expression have not been clearly defined. Upon scanning the proximal promoter regions of the human CFTR and MUC1 genes for shared sequence elements, several regions of purine/pyrimidine bias were noted. Some of these homopurine/homopyrimidine elements displayed imperfect mirror symmetry and are here termed purine/pyrimidine mirror repeats (PMRs). Previous studies of the region upstream of the MUC1 transcription start site identified three imperfect PMRs, M-PMR1 (27 bp, centered at –627), M-PMR2 (17 bp, centered at –244), and M-PMR3 (32 bp, centered at –118). Two imperfect PMRs in the CFTR promoter are similar to those in the MUC1 promoter, C-PMR1 (36 bp, centered at –170) and C-PMR2 (18 bp, centered at –468).

The results of S1 nuclease hypersensitivity studies showed that non-B DNA structures arose within C-PMR1 and M-PMR3 under conditions of supercoiling and acidic pH (9). S1 nuclease hypersensitivity mapping to a PMR is strong evidence for the formation of H-DNA (reviewed in Ref. 10). H-DNA refers to a family of structures characterized by an intramolecular DNA triple helix and regions of base unpairing that confer sensitivity to chemical and enzymatic probes of single-stranded character in DNA. Stable H-DNA structures require some degree of mirror symmetry and are typically stabilized by supercoiling. Different H-DNA isomers arise from the formation of two mutually exclusive families of triple helices (Y-R-Y, pyrimidine motif; R-R-Y, purine motif). Perfect PMRs formally have the potential to adopt four different isomers (H-y3, H-y5, H-r3, and H-r5), depending on which half-element strand is donated to the intramolecular triplex (reviewed in Ref. 10). These isomers typically have very different energies, with the equilibrium strongly favoring one isomer, often H-y3. NMR spectroscopy has provided key information concerning the structures of triple-helical domains of H-DNA (11).

Although H-DNA can be detected under a variety of conditions using isolated plasmid DNA, it has only been detected under extreme conditions in bacteria (12), and it has never been demonstrated in living eukaryotes. On the other hand, perfect PMR sequences predicted to form very stable H-DNA structures (under appropriate conditions) are statistically overrepresented in the human genome (relative to yeast and Escherichia coli), with a frequency of once every 49 kilobase pairs (13). An antibody preparation thought to be specific for triplex DNA binds to metaphase chromosomes in fixed mammalian cells (14). In addition, PMRs have frequently been identified upstream of genes, a provocative location that suggests a possible role in transcriptional regulation.

The identification of imperfect PMR sequences upstream of
the human CFTR and MUC1 genes suggested that these sequences could form H-DNA. In previous experiments we demonstrated that these sites become sensitive to S1 nuclease at low pH (9). Subsequent high resolution studies of the PMR element upstream of the CFTR gene detected the presence of a non-B DNA structure related to the H-3 y3 isomer of H-DNA at low pH (13). In related work, single-strand-specific nuclear proteins were detected in nuclear extracts from cultured human cell lines (9). One of these factors, a 27-kDa protein, bound preferentially to purine-rich single strands including those within the PMRs from the CFTR and MUC1 genes, suggesting the possibility that this factor could stabilize certain unpaired DNA structures.

In the present study we more fully characterize the non-B DNA structure responsible for S1 nuclease sensitivity of M-PMR3.

EXPERIMENTAL PROCEDURES

Recombinant Plasmids—Subcloning of a 90-bp DNA fragment containing the 32-bp M-PMR3 element into plasmid pGEM-4Z (Promega) was performed after polymerase chain reaction using plasmid pMAH5 (9) as the template. Primers S-GCAAGTTGGCGCCGTTCACTTACGAC and S-GCAAGTTGGCGCCGTTCACTTACGAC and 5-bp M-PMR3 element modified to create perfect homopurine/homopyrimidine mirror symmetry. The resulting DNA fragment was cleaved with HindII and EcoRI and ligated into pGEM-3Z. Plasmid DNA at native superhelical density was extracted from DH5α cells and purified by CsCl equilibrium density gradient centrifugation in the presence of ethidium bromide (16).

Nuclease Sensitivity—S1 and P1 nuclease were obtained from Life Technologies, Inc. To map plasmid sites sensitive to double-strand cleavage by S1 nuclease, 2 μg of supercoiled DNA was treated in 50-μl reactions containing 30 mM sodium acetate buffer (pH 4.5), 50 mM NaCl, 1 mM ZnCl₂, 4 mM MgCl₂, 5% (v/v) glycerol, and 75 units of S1 nuclease. Reactions were incubated on ice for 30 min and then stopped by the addition of 10 μl of a solution containing 0.4 M Tris base and 0.25 M EDTA. Reactions were then diluted to 100 μl with H₂O, extracted with phenol, and precipitated using ethanol. For some experiments, the resulting DNA was resuspended and treated with ScaI endonuclease. For high resolution experiments, 5-μg samples of supercoiled plasmid DNA-P-PMR3 were sparingly nicked by nuclease treatment at 37°C for 30 min in 50-μl reactions at either pH 4.5 (0.3 units of S1 nuclease or 0.004 units of P1 nuclease) or pH 7.1 (0.4 units of P1 nuclease). Reactions at pH 4.5 contained 20 mM sodium acetate, 4 mM MgCl₂, and 100 mM NaCl. Reactions at pH 7.1 contained 25 mM MOPS, 4 mM MgCl₂, and 100 mM NaCl. Reactions were terminated by the addition of 50 μl of chilled H₂O, extracted with phenol, and precipitated with ethanol.

Chemical Sensitivity—Anhydrous hydrazine, CAA (50 wt % solution in H₂O), DEP, DMS, 2,2′-dipyridyl, formic acid, OT (4 wt % solution in H₂O), and piperidine were obtained from Aldrich and were used without further purification. Chemical probing of supercoiled pM-PMR3 was performed using modifications of published procedures (17). Supercoiled DNA samples (10 μg) were dissolved in 100 μl of pH 4.5 or pH 7.1 buffers (described above). For CAA reactivity, duplicate plasmid samples were treated with 2 μl of CAA solution (0.12 μ g final concentration) at 37°C for 1 h. For DEP reactivity, plasmid samples were treated with 3 μl of DEP (0.2 μl final concentration) at 24°C for 30 min with agitation. For OT reactivity, plasmid samples were treated for 15 min at 37°C with 7.6 μl of a solution made by mixing 14.2 μl of 0.5% 2,2′-dipyridyl and 3 μl of OT solution (final concentrations of both OT and 2,2′-dipyridyl were 1 μl final concentration). Reactions were terminated by the addition of 25 μl of DMS stop solution (16), followed by two ethanol precipitations.

High Resolution Mapping—Enzyme-treated or chemically modified pM-PMR3 or pM-PMR3* DNA was treated with HindII and EcoRI to release the corresponding 90-bp fragment containing the M-PMR3 or M-PMR3* elements. These fragments were purified by electrophoresis through a native 5% polyacrylamide gel, followed by elution and sequential precipitations using ethanol and spermine (18). The restriction fragments were selectively radiolabeled at one terminus of either the purine-rich or pyrimidine-rich strand using the Klenow fragment of E. coli DNA polymerase I and appropriate [32P]dideoxynucleoside triphosphates. After precipitation with spermine, untreated and CAA-treated samples were then treated with either piperidine formate (pH 2) or hydrazine in high salt followed by sequential ethanol precipitations to provide purine or C > T sequence ladders (16). Chemically modified DNA samples were then treated with 100 μl of 10% (v/v) piperidine at 90°C for 30 min, frozen, and lyophilized overnight. Samples containing equivalent amounts of radioactivity were analyzed by electrophoresis through denaturing 8% polyacrylamide sequencing gels. Radioactive signals were imaged and quantified using a Molecular Dynamics PhosphorImager. Single-hit kinetics were verified by detection of greater than 70% of total radioactivity in the full-length DNA fragment.

RESULTS

S1 Nuclease Hypersensitivity in the MUC1 Promoter Maps to the M-PMR3 Element—In previous experiments, we studied the S1 nuclease sensitivity of DNA fragments derived from the CFTR and MUC1 gene promoters (9). In these studies supercoiled (but not linearized) plasmids containing a MUC1 promoter fragment that included both the M-PMR2 and M-PMR3 elements displayed S1 nuclease hypersensitivity mapping in this region, whereas a plasmid containing M-PMR1 did not. The sequence of M-PMR3 is similar to the S1 nuclease sensitive C-PMR1 element from the human CFTR promoter. We therefore sought to determine if M-PMR3 displayed S1 nuclease sensitivity when isolated from the MUC1 promoter. A 90-bp DNA fragment containing M-PMR3 was removed from the MUC1 promoter and placed into a plasmid vector (Fig. 1) to create pM-PMR3.

To detect non-B DNA in pM-PMR3, supercoiled pM-PMR3 plasmid DNA or the supercoiled vector (lacking the M-PMR3 insert) were treated with S1 nuclease as shown in Fig. 2A. Supercoiled vector and pM-PMR3 plasmids comigrate (Fig. 2A, compare lanes 1 and 5). Linearization with Scal provides the reference band shown in lanes 2 and 6 of Fig. 2A. Treatment of supercoiled DNA with S1 nuclease at pH 4.5 caused substantial nicking and linearization of both vector and pM-PMR3 plasmids, indicating that both contained one or more non-B DNA structures (Fig. 2A, lanes 3 and 7). To precisely map sites of S1 nuclease sensitivity, S1-treated plasmids were digested with Scal to deplete the plasmid DNA at a unique site. These results are shown in Fig. 2A (lanes 4 and 8). Analysis of the estimated lengths of released DNA fragments indicates that the S1 cleavage site in the vector DNA lies near the origin of replication (Fig. 2A, lane 4), as indicated by (I) in Fig. 2B. A cruciform structure can be extrapolated at this site during the alkaline lysis plasmid preparation procedure (19, 20). In contrast, S1 sensitivity maps within the MUC1 promoter insert in pM-PMR3 (compare Fig. 2A, lanes 4 and 8). Fragment length analysis maps the S1 cleavage site to site (II) as shown in Fig. 2B. Thus, as previously observed for the C-PMR1 element from the human CFTR gene (21), insertion of the PMR into a plasmid vector alters the major site of S1 nuclease attack. S1 sensitivity initially mapping to the plasmid origin is completely replaced by S1 sensitivity mapping to the PMR element.

High Resolution Mapping of Nuclease-Sensitive Sites—Samples of plasmid pM-PMR3 were treated with limiting amounts of S1 nuclease at pH 4.5 or P1 nuclease at pH 4.5 or 7.1 to allow high resolution mapping of the nuclease-sensitive sites. Nicked DNA from the region of the M-PMR3 insert was isolated, labeled uniquely on one strand, denatured, and analyzed on sequencing gels as described under “Experimental Procedures.” The results are shown in Fig. 3. The data in Fig. 3 reveal several patterns of nuclease sensi-
tivity. First, a non-B DNA conformation occurs in the M-PMR3 element at pH 4.5 as detected by the S1 nuclease sensitivity of both DNA strands (Fig. 3, lanes 5 and 14). Cleavage by S1 nuclease occurs only within boundaries of the M-PMR3 element. On the purine (R) strand, cleavage is centered on the center of mirror symmetry and is particularly notable at the sites of imperfection in this symmetry (Fig. 3, lane 5). Cleavage on the pyrimidine (Y) strand is more subtle, limited to the 5'-half of the element and occurs most clearly at sites of mirror imperfection (Fig. 3, lane 14). Second, at pH 4.5, P1 nuclease cleaves weakly to produce a somewhat different pattern of sites than S1 (Fig. 3, lanes 7 and 16). To address non-B DNA structures at neutral pH, levels of P1 nuclease were increased to produce the same overall plasmid nicking at pH 7.1 as had occurred at pH 4.5 (data not shown). In no case were nicks observed in or near the M-PMR3 element, demonstrating that the non-B DNA structure in this region was absolutely dependent on low pH (Fig. 3, lanes 8, 9, 17, and 18). S1 and P1 nucleases therefore identify a non-B DNA conformation within M-PMR3 that is stabilized by protons.

High Resolution Mapping of Nucleotides Modified by Chemical Probes—We further characterized the non-B conformation of the native M-PMR3 element using four chemical probes of unpaired bases previously shown to be of value in detecting non-B DNA structures (17, 22–25, for a review see Ref. 26). DMS is sensitive to Hoogsteen hydrogen bonding because the N7 position of guanine is protected from methylation by DMS if guanine N7 is involved in a stable hydrogen bond (e.g. C-G-C...G-C).

**Fig. 1.** Sequence and subcloning of M-PMR3 element from the human MUC1 mucin promoter region. The 32-bp M-PMR3 element (upper boxed sequence) is an imperfect homopurine/homopyrimidine mirror repeat centered at position -118 relative to the major startpoint of MUC1 transcription. In the depicted DNA sequence, the center of the element is indicated by +, and imperfections in mirror symmetry are indicated by x. The 32-bp M-PMR3* element (lower boxed sequence) is a perfect homopurine/homopyrimidine mirror repeat centered between positions -123 and -124 relative to the major startpoint of MUC1 transcription. The center of the element is indicated by +. Recombinant plasmid pM-PMR3 (below) contains the M-PMR3 element subcloned as a ~90-bp DNA fragment between HindIII and EcoRI sites of pGEM-4Z. Plasmid pM-PMR3* is similar but contains the M-PMR3* element.

**Fig. 2.** Low resolution mapping of S1 nuclease sensitive sites in pM-PMR3. A, samples of supercoiled plasmid DNA lacking (lanes 1–4) or containing (lanes 5–8) the M-PMR3 element were treated under various conditions, separated by agarose gel electrophoresis, and detected by staining with ethidium bromide. Samples were linearized with Scal (lanes 2 and 6), treated with S1 nuclease at low pH (lanes 3 and 7), or treated with S1 nuclease at low pH followed by Scal (lanes 4 and 8). Sizes of phage λ HindIII reference markers are indicated at left. Estimated plasmid fragment sizes are shown at right. Linearized plasmid DNA is indicated (L). B, map of pM-PMR3 indicating S1 nuclease cleavage sites I and II, and estimated fragment sizes.

**Fig. 3.** High resolution mapping of S1 and P1 nuclease-sensitive sites in pM-PMR3. Lanes 1-9 depict reactivities of the purine-rich (R) strand of the M-PMR3 element (refer to lane 2 for A-G reference ladder). Lanes 10-18 depict reactivities of the pyrimidine-rich (Y) strand (refer to lane 12 for C-T reference ladder). Element boundaries are labeled at right and left of the figure. The center of mirror symmetry is indicated by +. Imperfections in mirror symmetry are indicated by gaps. S1 nuclease treatment was performed at pH 4.5. P1 nuclease treatment was performed at pH 7.1. In both cases, enzyme amounts were chosen so as to nick less than one-third of the supercoiled plasmids in each sample.
triplet). CAA preferentially reacts with adenine and cytosine (and to a lesser extent, guanine) to form their etheno derivatives when these bases occur in an unpaired or strained context, whereas DEP reacts with unpaired or strained purine residues (including, but not limited to, Z-DNA). OT reacts primarily with unpaired thymine bases.

DMS reactivity of guanines in the purine strand is not greatly different between pH 4.5 and pH 7.1 (Fig. 4, compare lanes 4 and 5). However, methylation appears slightly suppressed over the 5' -half of the element relative to the 3' -half (Fig. 4, lane 4). Complete protection of guanine N7 atoms within a putative intramolecular triplex may require that the structure be extremely stable.

CAA reactivities on the purine strand are best seen when superimposed with the C.A. chemistry (Fig. 4, compare lanes 8 and 9). At low pH, bases attacked by CAA clearly map only to purines on the 3' -half of the M-PMR3 element (Fig. 4, lane 8). No reactivity is observed at neutral pH. On the pyrimidine strand, CAA reactivity is best visualized when superimposed with A.G chemistry (Fig. 4, compare lanes 22 and 23). In this case, CAA reaction marks the two A residues that create imperfections in mirror symmetry on the 5' -half of the pyrimidine strand. No CAA hyperreactivity is observed at pH 7.1.

DEP treatment at pH 4.5 modified bases in a pattern very similar to that observed for CAA (Fig. 4, compare lanes 8 and 10, and lanes 22 and 26). These results are consistent with hyperreactivity of the 3' -half of the purine strand relative to the 5' -half, and emphasize the unique reactivity of A residues creating imperfections in mirror symmetry on the 5' -half of the pyrimidine strand (Fig. 4, lane 26). No DEP hyperreactivity is observed at pH 7.1 (Fig. 4, lane 27).

OT provides the most distinct signature within M-PMR3 at pH 4.5. This reagent strongly reacts with two T residues that create imperfections in the 3' -half of the purine strand of M-PMR3 (Fig. 4, lane 12) but does not identify these T residues at pH 7.1. T residues near the center of mirror symmetry on the pyrimidine strand are highly reactive to OT, as is a single T residue at position –101 just beyond the 5'-border of M-PMR3 on the pyrimidine strand (Fig. 1; Fig. 4, lane 28).

Effect of Mg\textsuperscript{2+} on Non-B DNA Structures—The stability of non-B DNA structures can depend on the nature and concentration of divalent cations (27, 28). Our initial experiments included 4 mM Mg\textsuperscript{2+}, previously shown to enhance S1 nuclease reactivity of the C-PMR1 element (21). To explore the importance of Mg\textsuperscript{2+} in stabilizing the non-B DNA structure within M-PMR3, patterns of CAA and OT reactivity were compared at pH 4.5 in the presence and absence of 4 mM MgCl\textsubscript{2} (Fig. 5). Hyperreactivity to CAA is observed for certain bases within M-PMR3 independent of the presence of Mg\textsuperscript{2+} (Fig. 5, compare lanes 3, 4, and 5). In contrast, the extent of OT modification of bases on the pyrimidine strand of M-PMR3 was significantly enhanced in the presence of Mg\textsuperscript{2+} (Fig. 5, compare lanes 9 and 10). These data suggest that the non-B structure in M-PMR3 can be stabilized by 4 mM Mg\textsuperscript{2+} but does not require its presence.

H-DNA Structural Model for Native M-PMR3 Element at Low pH—To create a reasonable model for the non-B DNA structure formed by M-PMR3 under conditions of supercoiling and low pH, nuclease and chemical sensitivity data are summarized in Fig. 6. This analysis shows that many probes of non-B structure recognize sequences near the center of M-PMR3. Moreover, most of the reactive sites in the element map above the center of mirror symmetry in Fig. 6. This corresponds to the 5' -half of the pyrimidine strand and the 3' -half of the purine strand. Bases in the opposite half of the element (3' -half of the pyrimidine strand and 5' -half of the purine strand)
H-DNA in the Human MUC1 Gene Promoter

Altered H-DNA Conformation in M-PMR3* with Perfect Ho-9arity. These data were used to discriminate between H-DNA models representing the conventional H-93 and H-95 isomers. The pH dependence and divalent cation independence of the M-PMR3 structure suggest that H-r isomers are not formed. Fig. 8, A and B, demonstrates that major sites of reactivity probes specific for non-B DNA (boxed residues) are consistent with an H-95 structure (Fig. 8B) but not with an H-93 structure (Fig. 8A).

FIG. 5. Effects of Mg2+ on chemical reactivity of M-PMR3. Lanes 1-5 depict reactivities of the purine-rich (R) strand of the PMR3 element (refer to lane 2 for A + G reference ladder). Lanes 6-10 depict reactivities of the pyrimidine-rich (Y) strand (refer to lane 8 for C > T reference ladder). Elements are labeled as in Fig. 3.

To test the hypothesis that M-PMR3* would adopt the more conventional H-93 isomer at low pH, high resolution assays of reactivity to OT, DEP, and CAA were performed with this element. These data are shown in Fig. 7 and are summarized in Fig. 8, C and D. OT reactivity with the purine strand of M-PMR3* was limited to two T residues just outside the boundaries of the element (Fig. 7, lane 3). DEP reactivity with the purine strand mapped to the center and 5'-border of the element (Fig. 7, lane 5). Interestingly, traces of DEP reactivity were also observed in samples treated at pH 7.1, suggesting that the stability of the non-B structure of M-PMR3* was enhanced at neutral pH relative to that of M-PMR3. CAA reactivity with the purine strand of M-PMR3* was limited to the 5'-half of the element (Fig. 7, lane 7). As was observed for DEP, some residual CAA reactivity was detected at pH 7.1 (Fig. 7, compare lanes 7 and 8). It is noteworthy that DEP and CAA reactivity was not evenly distributed within the purine region predicted to be unstructured in the H-93 isomer (Fig. 8C). This behavior is similar to the uneven reactivity of the purine strand of M-PMR3 (Fig. 8B) and that previously observed within a PMR element in the CFTR gene promoter (21). These observations suggest that nucleotides presumed to be "unstructured" are actually constrained by folding or packing against the intramolecular triplex such that only a subset of the sequence is exposed. The pyrimidine strand of M-PMR3* was uniformly unreactive to DEP and CAA (Fig. 7, lanes 14-19). OT strongly reacted with the four T residues proximal to the center of mirror symmetry but with no other T residues of the pyrimidine strand (Fig. 7, lane 12).

Chemical probing data for M-PMR3* were recorded and mapped onto conventional H-93 and H-95 models for H-DNA. The result is shown in Fig. 8, C and D. In contrast to the reactivity of the native M-PMR3 sequence, the pattern of M-PMR3* reactivity suggests an H-93 structure. This result is consistent with other studies in which PMR elements with perfect mirror symmetry have been shown to favor the H-93 isomer. This result confirms the hypothesis that that the unusual folding of M-PMR3 is related to its imperfect mirror symmetry.

DISCUSSION

PMRs Upstream of MUC1 and CFTR Display Mirror Symmetry in a Nonrepetitive Pattern—The best-studied examples of PMR sequences capable of adopting H-DNA structures display perfect mirror symmetry and are repetitive in nature, such as d(GA)16–18 (22, 23, 29, 30), d(G)30 (27), or other di-, tri-, or tetra-nucleotide repeats (25, 31). The short PMR elements in
the human CFTR and MUC1 gene promoters are not simple sequence repeats and contain imperfections in mirror symmetry. Both C-PMR1 and M-PMR3 are hypersensitive to S1 nuclease and other probes of unpaired bases at low pH. A previous study involving high resolution chemical and enzymatic probing of C-PMR1 indicated that the reactivity of this element resembled H-DNA in some ways (e.g. unpaired nucleotides at the center of mirror symmetry) but did not clearly conform with either H-y3 or H-y5 isomer. In contrast, the present chemical and enzymatic reactivity data for the native M-PMR3 element are quite clearly diagnostic of the H-y5 isomer.

**Hy-y3 Versus H-y5 Conformations in H-DNA**—The original studies of PMR sequences at low pH and native levels of supercoiling clearly demonstrated that the H-y3 isomer is favored over the H-y5 isomer (22, 23, 25, 27, 32–34). A plausible explanation for this preference involves details of strand rearrangements in each isomer. Topological modeling suggests that the H-y3 isomer releases one extra supercoil relative to the H-y5 isomer (29). However, details of divalent ion conditions and the exact nucleotide sequence of the loop region can influence the equilibrium between isomers (35–37).

The data obtained for the native M-PMR3 element conform to the more rare H-y5 isomer. Although initially unexpected, this preference may be understood by referring to the putative structures shown in Fig. 8 (compare A and B). Because of the presence of two T residues in the 3'-half of the purine-rich strand of the PMR, folding into the H-y3 isomer creates a situation in which the Hoogsteen strand of the putative triplex is too short to form continuous triplets (Fig. 8 A). The resulting geometry is likely to be very unstable. In contrast, the H-y5 isomer accommodates the extra T residues in the more reactive 3'-half of the purine strand (Fig. 8 B). A single A residue in the 5'-half of the pyrimidine strand is predicted to be extruded from the putative triplex, and this residue is indeed hyperreactive to DEP (Fig. 4, lane 26; Fig. 8 B).

Evidence for the role of sequence asymmetry in the H-y3: H-y5 equilibrium in M-PMR3 was obtained by studying H-DNA formation in the fully symmetric sequence, M-PMR3*. When sequence asymmetry in M-PMR3 was corrected in M-PMR3*, the element adopted a canonical H-y3 conformation. These results suggest that, together with ionic conditions and loop sequences, details of sequence asymmetry can determine the relative stabilities of H-y5 and H-y3 isomers.

**Conditions Required for H-DNA in the MUC1 Promoter**—As is typical for PMR elements containing mixtures of G and A residues, cytosine protonation was important in stabilizing the non-B structure of M-PMR3. H-DNA was detected at pH 4.5 but not at pH 7.1. It is noteworthy that there was no evidence for H-r5 or H-r3 isomers (involving R\_zR\_zY triplets) at neutral pH even in the presence of Mg\_2\^1. That supercoiling was required for the S1 sensitivity of M-PMR3 was previously established (9).

**Possible Roles of Non-B DNA in MUC1 Promoter Function**—Possible functions for H-DNA structures have been proposed in the regulation of DNA replication (38–41), recombination (42–44), and transcription (reviewed in Ref. 10). Unpaired nucleotides in H-DNA might be recognized by transcription factors,

---

Fig. 7. Chemical reactivity of M-PMR3* element with perfect homopurine/homopyrimidine mirror symmetry. A, lanes 1-8 depict reactivities of the purine-rich (R) strand of the PMR3* element (refer to lane 2 for A + G reference ladder). Lanes 9-19 depict reactivities of the pyrimidine-rich (Y) strand (refer to lane 11 for C + T reference ladder). Element boundaries are labeled at left of each panel. Centers of mirror symmetry are indicated by +.

Fig. 8. Summary of reactivity data and putative H-DNA conformations. Data for the native M-PMR3 element are mapped onto H-y3 (A) and H-y5 (B) H-DNA structural models. Boxes indicate nucleotide hyperreactive to enzymatic and chemical probes. The data are most consistent with the H-y5 conformation for the native M-PMR3 sequence at low pH. Data for the M-PMR3* element with perfect homopurine/homopyrimidine mirror repeat symmetry are mapped onto the H-y3 (C) and H-y5 (D) H-DNA structural models. The data are most consistent with the H-y3 conformation for the symmetric M-PMR3* sequence at low pH.
H-DNA formation might create a flexible hinge for chromatin folding, or the structure might act as a buffer for superhelical strain in transcription. Single-strand-specific DNA binding proteins that recognize homopurine or homopyrimidine sequences have been identified (9, 45–48). Such factors might stabilize certain H-DNA isomers.

Despite these intriguing possibilities, specific H-DNA structures have not been detected in eukaryotic cells, and the putative functions of H-DNA remain unknown. Indeed, when examined by genetic and biochemical experiments, PMRs have been shown to act as conventional B-form duplexes that may be recognized by DNA binding proteins (10, 15, 49, 50).

In the case of the short PMR elements in the promoters of the human CFTR and MUC1 mucin genes, little evidence is available from which to conclude whether H-DNA might be involved in the transcriptional regulation of these genes. Data from transient transfection experiments using promoter-reporter constructs suggest that M-PMR3 contains sequences that have a weakly negative effect on transcription initiation, since deletion of this element increased transcription in some cell lines. However, mirror symmetry may not be required for this activity. Combinations of genetic and biochemical studies will be needed to better address whether PMRs adopt H-DNA structures in living cells and to explore the function(s) of such structures.

Acknowledgments—We acknowledge expert suggestions from B. Johnston, H. Htun, and S. Mirkin.

REFERENCES
1. Zotter, S., Hageman, P. C., Lossnitzer, A., Modl, W. J., and Hilgers, J. (1988) Cancer Res. 12, 55–101
2. Chambers, J. A., Hollingsworth, M. A., Trezise, A. E. O., and Harris, A. (1994) J. Cell Sci. 107, 415–424
3. Hollingsworth, M. A., Strawhecker, J. M., Caffrey, T. C., and Mack, D. R. (1994) Int. J. Cancer 57, 198–203
4. Engelhardt, J. F., Yankaskas, J. R., Ernst, S. A., Marino, C. R., Boucher, R. C., Cohn, J. A., and Wilson, J. M. (1993) Nat. Genet. 2, 240–247
5. Trezise, A. E. O., Chambers, J. A., Wardle, C. J., Gould, S., and Harris, A. (1993) Hum. Mol. Genet. 2, 213–218
6. Abe, M., and Kufe, D. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 282–286
7. Kozarik, A., Peat, N., Wilson, D., Gendler, S. J., and Taylor-Papadimitriou, J. (1993) J. Biol. Chem. 268, 9917–9926
8. Shirioti, K., Taylor-Papadimitriou, J., Gendler, S. J., and Irimura, T. (1994) J. Biol. Chem. 269, 10303–10306
9. Hollingsworth, M., Closken, C., Harris, A., McDonald, C., Pahwa, G., and Maher, L. J. (1994) Nucleic Acids Res. 22, 1138–1146
10. Frank-Kamenetski, M. D., and Mirkin, S. M. (1995) Annu. Rev. Biochem. 64, 85–95
11. Radhakrishnan, L., and Patel, D. (1994) Biochemistry 33, 11405–11416
12. Kohwi, Y., Makhosyan, S., and Kohwi-Shigematsu, T. (1992) J. Mol. Biol. 223, 817–822
13. Schroth, G. P., and Ho, P. S. (1995) Nucleic Acids Res. 23, 1977–1983
14. Lee, S., Burkholder, G. D., Latimer, J. L. P., Haug, B. L., and Braun, R. P. (1987) Nucleic Acids Res. 1047–1061
15. Raghu, G., Tevosian, S. A., Anant, S., Subramanian, K. N., George, D. L., and Mirkin, S. M. (1994) Nucleic Acids Res. 22, 3271–3279
16. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
17. Pestov, D., Dayn, A., Siyanova, E., George, D., and Mirkin, S. (1991) Nucleic Acids Res. 19, 6527–6532
18. Hopps, B., and McClure, W. (1981) Nucleic Acids Res. 9, 5493–5504
19. Lilley, D. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 6468–6472
20. Tomizawa, J. (1984) Cell 38, 861–870
21. McDonald, C. D., Hollingsworth, M. A., and Maher, L. J. (1994) Gene (Amst.) 150, 267–274
22. Htun, H., and Dahlberg, J. (1988) Science 241, 1791–1796
23. Johnston, B. H. (1988) Science 241, 1800–1804
24. Kohwi-Shigematsu, T., and Kohwi, Y. (1985) Cell 43, 199–205
25. Hanvey, J., Klysik, J., and Wells, R. (1988) J. Biol. Chem. 263, 7386–7396
26. Wells, R. D., Coller, D. A., Hanvey, J. C., Shimizu, M., and Wohlrab, F. (1988) FASEB J. 2, 2939–2949
27. Kohwi, Y., and Kohwi-Shigematsu, T. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 3781–3785
28. Bernard, J., Betran, R., Casasnovas, J., and Azorin, F. (1990) Nucleic Acids Res. 18, 4067–4073
29. Htun, H., and Dahlberg, J. E. (1989) Science 243, 1571–1576
30. Lyamichev, V., Mirkin, S., and Frank-Kamenetski, M. (1986) J. Biomol. Struct. & Dyn. 3, 667–669
31. Coller, D. A., and Wells, R. D. (1990) J. Biol. Chem. 265, 10652–10658
32. Voloshin, O. N., Mirkin, S. M., Lyamichev, V. I., Befeleroskovski, B. P., and Frank-Kamenetski, M. D. (1988) Nature 333, 475–476
33. Hanvey, J., Shimizu, M., and Wells, R. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 6292–6296
34. Votikova, M., Mirkin, S., Lyamichev, V., Voloshin, O., Frank-Kamenetski, M., and Palacek, E. (1988) FEBS Lett. 234, 295–299
35. Kang, S., Wohlrab, F., and Wells, R. (1992) J. Biol. Chem. 267, 1259–1264
36. Kang, S., and Wells, R. D. (1992) J. Biol. Chem. 267, 20889–20891
37. Shimizu, M., Kubo, K., Matsumoto, U., and Shindo, H. (1994) J. Mol. Biol. 235, 185–197
38. Dayn, A., Samadashwily, G. M., and Mirkin, S. M. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 11104–11110
39. Baran, N., Lapidot, A., and Manor, H. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 507–511
40. Lapidot, A., Baran, N., and Manor, H. (1989) Nucleic Acids Res. 17, 883–900
41. Peleg, M., Kopel, V., Borowicz, J. A., and Manor, H. (1995) Nucleic Acids Res. 23, 1292–1299
42. Kohwi, Y., and Panchenko, Y. (1993) Genes Dev. 7, 1766–1778
43. Caddle, M., Lussier, R., and Heintz, N. (1990) J. Mol. Biol. 211, 19–33
44. Rooney, S. M., and Moore, P. D. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 2141–2144
45. Kolliari, K., Torrey, T., and Kinniburgh, A. (1992) Nucleic Acids Res. 20, 111–116
46. Brunel, F., Alzari, P. M., Ferrara, P., and Zakim, M. (1991) Nucleic Acids Res. 19, 5237–5245
47. Yee, H. A., Wong, A. K. C., van de Sande, J. H., and Rattner, J. B. (1991) Nucleic Acids Res. 19, 949–953
48. Murasito, T., Nomoto, S., Yamazaki, H., Mishima, Y., and Komariki, R. (1992) Nucleic Acids Res. 20, 6631–6635
49. Gilmour, D., Thomas, G., and Elgin, S. (1989) Science 245, 1487–1490
50. Glaser, R., Thomas, G., Siegfried, E., Elgin, S., and Lis, J. (1990) J. Mol. Biol. 211, 751–761

H-DNA in the Human MUC1 Gene Promoter