Characterization of a polypurine/polypyrimidine sequence upstream of the mouse metallothionein-I gene

Nicole A. Becker and L. James Maher, III*

Department of Biochemistry and Molecular Biology, Mayo Foundation, 200 First Street S.W., Rochester, MN 55905, USA

Received December 23, 1997; Revised and Accepted March 4, 1998

ABSTRACT

A 128 base pair long homopurine/homopyrimidine (R/Y) element is located −1.2 kb upstream of the transcription start point of the mouse metallothionein-I (MT-I) gene. We present a detailed in vitro structural characterization of the MT-I R/Y sequence as determined by enzymatic and chemical probes. An −190 bp fragment containing the MT-I R/Y sequence was subcloned into a recombinant vector. Low resolution analysis with S1 nuclease indicates that DNA in this region was unpaired in supercoiled plasmids treated at low pH. High resolution mapping with chemical probes selective for non-B DNA structures provides evidence that the MT-I R/Y sequence adopts one or more H-DNA structures. We also investigated this sequence to determine if it can influence transcriptional regulation. Promoter/reporter constructs were prepared in which the MT-I R/Y sequence was positioned in either orientation upstream of either the MT-I or HSV-TK promoters. Promoter/reporter activities were evaluated by transient transfection assays using mouse NIH3T3 cells. The MT-I R/Y sequence displayed no detectable activity as a cis-acting transcriptional regulatory element. These results demonstrate that although the MT-I R/Y sequence is able to adopt a non-B DNA structure under certain in vitro conditions, there is no evidence that this sequence plays a significant role in transcriptional regulation.

INTRODUCTION

Metallothioneins (MTs) are small cysteine-rich proteins found in all higher eukaryotes. MTs strongly bind heavy metals and play an important role in metal homeostasis (1,2). Heavy metal regulation of genes encoding MTs therefore serves as an important mechanism in cadmium and mercury detoxification (1–3). This regulation is also thought to play a significant role in maintaining physiological levels of essential heavy metals, such as copper and zinc (1.2). MT gene expression is induced by both heavy metals (e.g. cadmium) and by glucocorticoids (1.4,5). Metal induction is mediated by several copies of a 15 base pair (bp) consensus sequence (metal-responsive element, MRE) present in the promoter regions of MT genes (5,6). Exposure to heavy metals results in a rapid 5–20-fold increase in MT mRNA and protein levels in cultured cells and in tissues (5,6).

The mouse MT-I gene is a prototypical member of the MT gene family and has been well characterized. The MT-I promoter region contains six MREs in various orientations localized within the first 200 bp upstream of the transcription start point (tsp) (4,7). The mouse MT-I gene and its promoter have been studied extensively to determine the effect of heavy metals (typically cadmium) on the rates of MT-I transcription and translation (8,9).

We became interested in the mouse MT-I promoter after the identification of a peculiar homopurine/homopyrimidine (R/Y) sequence in its vicinity (10). We refer to this element as the MT-I R/Y sequence. This sequence is 128 bp in length and is centered 1184 bp upstream of the major tsp. The MT-I R/Y sequence is of particular interest because of its unusual sequence composition, its length and its provocative location upstream of the MT-I gene. As shown for other sequences (including our previous studies of shorter R/Y sequences from the CFTR and MUC1 promoters), such elements may adopt non-B DNA structures of unknown significance under certain in vitro conditions (11–13). The R/Y sequences near CFTR and MUC1 are similar to that near MT-I in that they display imperfect mirror symmetry consistent with the potential for isomerization to imperfect triplex (H-DNA) structures (12,13). The MT-I R/Y sequence contains no single plane of extensive mirror symmetry, but contains several imperfect purine/pyrimidine mirror repeats (PMRs) within it. It is of interest to determine what, if any, unusual structure(s) occur within this sequence near physiological conditions, and whether the sequence can influence transcription initiation from an adjacent promoter.

Certain PMRs have been shown to isomerize to H-DNA in vitro under conditions of negative supercoiling and acidic pH (11,14–17). H-DNA refers to a family of DNA structures characterized by an intramolecular DNA triplex and accompanying regions of base unpairing. These regions of base unpairing are sensitive to chemical and enzymatic probes that react differentially with single-stranded DNA. Chemical and enzymatic probing allows structural information to be deduced for such sequence elements (18–23). Stable H-DNA structures require some degree of mirror symmetry and are typically stabilized by negative supercoiling. Distinct H-DNA isomers arise through the formation of two mutually exclusive families of intramolecular triple
helices due to base triplets. Triplexes with Y•R•Y triplets where the two Y strands are antiparallel are termed the ‘pyrimidine motif’. Triplexes with R•R•Y triplets where the two R strands are antiparallel are termed the ‘purine motif’. The pyrimidine motif is stabilized under acidic conditions due to the fact that protonation of the cytosine N3 position produces a particularly stable C–G–C triple. This protonation allows for a second Hoogsteen hydrogen bond between cytosine and guanine residues and presumably reduces unfavorable electrostatics within the triplex. Perfect PMRs have the potential to adopt four different isomers (H-3y, H-5y, H-3r and H-5r), where ‘r’ or ‘y’ and ‘3’ or ‘5’ describe which half-element strand is donated to the intramolecular triplex (11). These isomers are not isoenergetic.

H-DNA structures have been detected under a variety of conditions in purified supercoiled plasmid DNA, but such structures have only been detected under extreme conditions in bacteria (24), and have not been conclusively demonstrated in living eukaryotic cells. Previous work has demonstrated that an antibody preparation thought to be specific for triplex DNA binds to metaphase chromosomes in fixed mammalian cells (25). It is also provocative that PMR sequences predicted to form stable H-DNA structures (under appropriate conditions) are statistically over-represented in eukaryotic genomes, with an estimated frequency of one occurrence per 49 kb (26). Such sequence elements have been found near functionally interesting locations, such as sites of frequent recombination and upstream of transcription start sites (26–28). These intriguing locations have contributed to the suggestion that H-DNA might have a functional role, but this claim has not been clearly substantiated.

In the present study we used enzymatic and chemical probes specific for single-stranded DNA to characterize the non-B DNA structure(s) present within the MT-I R/Y sequence. Data from the in vitro structural characterization confirm the formation of one or more pH-dependent, non-B DNA structures and allowed us to formulate one plausible model. We further report that the MT-I R/Y sequence displayed no discernible activity as a cis-acting regulator of transcription in transient transfection experiments.

MATERIALS AND METHODS

Recombinant plasmid constructs

pJ034. An ∼190 bp fragment of the mouse MT-I promoter containing the 128 bp R/Y sequence was generated by polymerase chain reaction (PCR) using plasmid pMTCAT (29) and primers 5′-GCA2GCT2CT2GTATCTA2C2AC2ATGA and 5′-GCGA2T2CTGCTGC2TA2C2AC2ATGA and 5′-GC-GCA2T2CTGCTG2TA2C2AC2ATGA and 5′-GC2TCAAGGCTG2CTATGACACTC to install EcoRI and HindIII restriction sites at the molecular termini. Insertion of the indicated fragment between the HindIII and EcoRI restriction sites of pGL3-Basic resulted in pJ164. Subcloning of an ∼190 bp DNA fragment containing the 128 bp MT-I R/Y sequence into plasmid pJ164 was performed following PCR using plasmid pMTCAT and primers 5′-GC2TCAAGGCTG2CTATGACACTC to install XhoI restriction sites at the molecular termini. Insertion of the indicated fragment in either orientation between the XhoI restriction site of pJ164 produced plasmids pJ165 and pJ166. Plasmid pJ166 contains the MT-I R/Y sequence in its native orientation relative to the transcription start point, whereas pJ165 contains the MT-I R/Y sequence in reverse orientation. Plasmid DNA for transient transfections was extracted from DH5α cells and purified by Qiagen resin after alkaline lysis preparation.

Enzymatic probing of non-B DNA structures

To map plasmid sites sensitive to cleavage by S1 nuclease, aliquots containing 2 μg supercoiled plasmid DNA were treated in 50 μl reactions containing 30 mM sodium acetate buffer (pH 4.5), 150 mM NaCl, 1 mM ZnCl2, 4 mM MgCl2, 5% (v/v) glycerol and 75 U S1 nuclease (Life Technologies, Inc.). Reactions were incubated on ice for 30 min and then terminated 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 Scal endonuclease.

Probing of non-B DNA structures

Anhydrous hydrazine, chloroacetaldehyde (CAA), dimethylsulfate (DMS), diethylpyrocarbonate (DEPC), 2,2′-bipyridyl, formic acid, osmium tetroxide (OT; 4 wt % solution in H2O) and piperidine were obtained form Aldrich and were used without further purification. The chemical reactivity of supercoiled plasmid pJ034 was performed using modifications of published procedures (18). Briefly, supercoiled plasmid DNA samples (15 μg) were dissolved in 100 μl of buffer. For experiments at pH 4.5, the buffer contained 20 mM sodium acetate, 4 mM MgCl2 and 100 mM NaCl. For the experiments at pH 7.1, the buffer contained 25 mM MOPS, 4 mM MgCl2 and 100 mM NaCl. To monitor CAA reactivity, duplicate plasmid samples were treated with diethylpyrocarbonate (DEPC), dimethylsulfate (DMS) and anhydrous hydrazine.
Mouse NIH3T3 fibroblasts were maintained at 37°C for 1 h. To analyze DEPC reactivity, plasmid samples were treated by adding 3 µl DEPC (~0.2 M DEPC final concentration) at 24°C for 30 min with periodic agitation. To monitor OT reactivity, plasmid samples were treated for 15 min at 37°C by adding 7.6 µl of a solution made by mixing 14.2 µl of 0.5% 2,2′-bipyridyl and 3 µl OT solution (final concentrations of both OT and 2,2′-bipyridyl were ~2 mM). CAA, DEPC and OT reactions were terminated by two ethanol precipitations. For DMS reactivity, plasmid samples were treated for 90 s at 24°C with 10 µl of a fresh 5% aqueous solution of DMS (48 mM final concentration). DMS reactions were terminated by the addition of 25 µl DMS stop solution [1.5 M sodium acetate (pH 5.2), 1 M β-mercaptoethanol, 2.5 µg tRNA] followed by two ethanol precipitations.

Mapping chemically-modified DNA at high resolution

Chemically modified pJ034 DNA was treated with HindIII and EcoRI endonucleases to release the ~190 bp fragment containing the MT-I/R/Y sequence. Restriction fragments were purified by electrophoresis through a native 8% polyacrylamide gel, eluted from the gel, and precipitated sequentially using ethanol and spermine (31). Purified restriction fragments were selectively radiolabeled at the 3′ terminus of either the purine-rich or pyrimidine-rich strand using the Klenow fragment of Escherichia coli DNA polymerase I and appropriate [32P]deoxyribonucleoside triphosphates. After precipitation with spermine, untreated and CAA-treated samples were then incubated with either formic acid or hydrazine in high salt followed by two ethanol precipitations to superimpose the CAA data on Maxam–Gilbert reference bands shown in Figure 2A, lanes 2 and 6. Treatment of plasmid DNA (32,33). In contrast, S1 sensitivity in pJ034 maps to a new sequence (compare Fig. 2A, lanes 4 and 8). Analysis of the estimated lengths of the 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 Figure 2B. Previous studies have shown that a stable cruciform structure can be formed at this site during the alkaline lysis preparation of plasmid DNA (32,33). In contrast, S1 sensitivity in pJ034 maps to a new sequence (compare Fig. 2A, lanes 4 and 8). Analysis of DNA fragment lengths suggests that the site of S1 nuclease sensitivity occurs at site (II) within the MT-I/R/Y sequence (Fig. 2B).

RESULTS

Low resolution S1 nuclease sensitivity identifies a non-B DNA structure in the MT-I/R/Y sequence under conditions of negative supercoiling and low pH

In previous experiments we have employed low resolution S1 nuclease sensitivity to demonstrate unusual structures in DNA sequences cloned from the CFTR and MUC1 gene promoters (12,13). These studies demonstrated that supercoiled (but not linearized) plasmids containing the PMR from these gene promoters displayed S1 nuclease sensitivity mapping in the region of the PMR. In the present study, we sought to determine if the 128 bp MT-I/R/Y sequence (not a perfect PMR) displayed S1 nuclease sensitivity when subcloned into a supercoiled plasmid. For this work, an ~190 bp DNA fragment containing the MT-I/R/Y sequence was amplified by PCR from the upstream region of the MT-I gene and subcloned to create pJ034 (Fig. 1).

To determine if one or more non-B DNA structures exist within pJ034, samples of supercoiled pJ034 plasmid DNA, or the supercoiled parent vector (lacking the MT-I/R/Y sequence insert) were treated with S1 nuclease as shown in Figure 2A. Results for untreated supercoiled vector and plasmid pJ034 are shown in Figure 2A, lanes 1 and 5. Linearization with ScaI provided the reference bands shown in Figure 2A, lanes 2 and 6. Treatment of supercoiled DNA with S1 nuclease at pH 4.5 resulted in almost complete linearization of both vector and pJ034 plasmids, indicating that both contained at least one non-B DNA structure with unpaired bases (Fig. 2A, lanes 3 and 7). In order to map site(s) of S1 nuclease sensitivity, S1-treated plasmid DNA was digested with ScaI to cleave the plasmid once. The resulting fragments are shown in lanes 4 and 8 of Figure 2A. Analysis of the estimated lengths of the 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 Figure 2B.

Cell culture and transient transfection

Mouse NIH3T3 fibroblasts were maintained at 37°C and 5% CO2 in Dulbecco’s minimal essential medium (DMEM) supplemented with 10% calf serum (Gibco). For transient transfections, mouse NIH3T3 cells were seeded into six-well tissue culture plates (~2 × 105 cells per well) and grown in 2 ml DMEM with 10% calf serum at 37°C and 5% CO2 to 50–60% confluency. Transfections were performed using Lipofectamine (Gibco) as directed by the manufacturer. At the time of transfection, cells were washed with 2 ml of Opti-MEM (Gibco) followed by the overlay of the lipid–DNA mixture and incubation for 5 h at 37°C in 5% CO2. After addition of calf serum to 10%, cells were further incubated at 37°C in 5% CO2 for a total transfection time of 48 h. In some experiments, MT-I gene expression was induced at 40 h with 2 µM CdCl2 for 8 h prior to assay of reporter gene expression.

Dual luciferase assay

Each culture well containing transfected NIH3T3 cells was washed once with 2 ml of phosphate buffered saline and lysed in 500 µl passive lysis buffer (Promega). Lysates were incubated at 24°C for 15 min with agitation. Assays for firefly and Renilla luciferase activities were performed sequentially from single extract preparations (Promega) and enzymatic activities were measured using a TD-20/20 luminometer (Turner Designs). For each experiment the luciferase values for the experimental construct were normalized to luciferase values for the control construct.

Transient transfection data calculations

Each transient transfection experiment measured reporter expression for four different test plasmid DNA concentrations prepared in duplicate. In addition, each test construct was assayed in four separate transient transfection experiments, using two different alkaline lysis plasmid preparations. Data from each transient transfection experiment were plotted (e.g. Fig. 5A). Slopes (normalized test reporter activity per microgram test construct) calculated by least-squares fitting for each construct were averaged, normalized to the first construct in each set, and the standard error of the mean was determined.
Figure 1. Subcloning of the MT-I R/Y sequence. The 128 bp MT-I R/Y sequence (underlined) is centered at position –1184 relative to the major MT-I tsp. The sequence contains no single plane of extensive mirror symmetry. A single pyrimidine interruption occurs within the purine strand (position –1176; vertical arrow). Recombinant plasmid pJ034 (shown below) contains the MT-I R/Y sequence cloned as an ∼ 190 bp DNA fragment between the HindIII and EcoRI sites of pGEM-4Z. Therefore, insertion of the MT-I R/Y sequence into a plasmid vector alters the major site of S1 nuclease sensitivity; S1 nuclease sensitivity initially mapping to the plasmid origin of replication is completely replaced by S1 nuclease sensitivity mapping to the inserted R/Y sequence.

High resolution mapping of base unpairing in the MT-I R/Y sequence using chemical probes

We used chemical probes that react preferentially with unusual DNA structures in order to further characterize the non-B DNA conformation of the MT-I R/Y sequence. Chemical probes that react with unpaired or unstacked bases have previously been shown to be of value in detecting and evaluating non-B DNA structures (18–23). CAA reacts with adenine and cytosine to form their respective etheno derivatives when these bases occur in unpaired or unstacked conformations. OT in the presence of 2,2′-bipyridyl reacts primarily with unpaired or unstacked thymine bases. DMS modifies the N7 position of guanine within a DNA duplex. In contrast, guanine is protected from DMS modification when involved in Hoogsteen hydrogen bonding (e.g. within a C+G:C triplet) because the N7 position of guanine is hindered. DEPC reacts with unpaired or unstacked purine residues.

Differences in CAA reactivities were detected when comparing plasmid samples treated at low pH relative to samples treated at neutral pH. CAA reactivities on the pyrimidine strand are best seen when superimposed on a G+A ladder (Fig. 3A, compare lanes 6 and 7). At pH 4.5, bases attacked by CAA are localized to regions –1175 to –1225 (Fig. 3A, lane 6). At pH 7.1, these reactivities were greatly reduced (Fig. 3A, lane 7). CAA reactivities on the purine strand are best seen when superimposed on a C>T ladder (Fig. 3B, compare lanes 5 and 6). At low pH, bases modified by CAA were localized to regions –1140 to –1200 (Fig. 3B, lane 5). At neutral pH, the reactivities of the bases within this region were again greatly reduced (Fig. 3B, lane 6). DEPC treatment at pH 4.5 modified bases in a pattern similar to that observed for CAA (Fig. 3B, compare lanes 5 and 11). OT modifications provided the most distinctive signature within the MT-I R/Y sequence. Treatment of pJ034 at pH 4.5 revealed strongly reactive T residues in the region from –1160 to –1221 on the pyrimidine strand (Fig. 3A, lane 6). At pH 7.1, this reactivity was greatly reduced (but not abolished) at pH 7.1 (Fig. 3A, compare lanes 11 and 12).

Figure 2. Low resolution mapping of S1 nuclease sensitive sites in pJ034. (A) Samples of supercoiled plasmid DNA lacking (lanes 1–4) or containing (lanes 5–8) the MT-I R/Y sequence were treated as indicated, separated by agarose gel electrophoresis, and detected by staining with ethidium bromide. Samples were linearized with ScaI (lanes 2 and 6), treated with S1 nuclease at pH 4.5 (lanes 3 and 7), or treated with S1 nuclease at pH 4.5 followed by ScaI digestion (lanes 4 and 8). Fragments released by S1 and ScaI digestion are indicated by white dots (left of lanes 4 and 8). Sizes of phage λ HindIII reference markers (lanes M) are indicated at left. Plasmid fragment sizes estimated by interpretation from standard curves are shown at right. Mobility of linearized plasmid DNA is indicated (L). (B) Map of pJ034 indicating S1 cleavage sites I and II, and estimated fragment sizes.
region is slight, its reproducibility suggests the location of a relatively stable intramolecular triplex. Complete protection of guanine N7 atoms throughout a putative intramolecular triplex region may require that the structure be particularly stable.

The data obtained from the chemical reactivity experiments was combined to predict one plausible folded structure for the \textit{MT-I} R/Y sequence (Fig. 4). The indicated hypothetical structure corresponds to an H-\gamma3 conformer of H-DNA. Despite its ability to accommodate much of our data, the proposed model is not meant to imply that a single non-B DNA structure predominates within this complex R/Y sequence at low pH. It appears likely that the model structure may co-exist or be in dynamic equilibrium with one or more alternative non-B DNA structures.

**Transient transfections to assay effects of the \textit{MT-I} R/Y sequence on transcription**

Previous data have shown R/Y and PMR sequences to adopt non-B DNA structures \textit{in vitro}, but the functional significance, if any, of such sequences \textit{in vivo} has not been clearly defined (11,34,35). Since such sequences are often found upstream of transcription start sites, it has been proposed that R/Y and PMR sequences might play a role in transcriptional regulation (11). Our \textit{in vitro} structural characterization indicates that a non-B DNA structure occurs within the \textit{MT-I} R/Y sequence, and is stabilized at low pH. Is this unusual structure an artifact of solution pH, or could it have relevance for transcriptional regulation \textit{in vivo}? To address this question we performed simple assays to determine if the \textit{MT-I} R/Y sequence can influence the transcriptional activity of the \textit{MT-I} promoter or heterologous promoters.

We created a series of promoter/reporter constructs to test if the \textit{MT-I} R/Y sequence modulates transcriptional activity in transient transfection assays (Fig. 5B–D, left). Constructs pJ105 and pJ106 were used to determine if deletion of a large DNA fragment including the \textit{MT-I} R/Y sequence from the native \textit{MT-I} 5′ flanking region would affect the transcriptional activity of the \textit{MT-I} promoter. Plasmid pJ105 includes \textasciitilde 1775 bp upstream of the \textit{MT-I} gene. This region contains the \textit{MT-I} R/Y sequence, MREs and the \textit{MT-I} basal promoter. Plasmid pJ106 contains an \textasciitilde 1100 bp deletion that encompasses the 128 bp \textit{MT-I} R/Y sequence, but leaves intact the MREs and basal \textit{MT-I} promoter (Fig. 5B, left). Transient transfection experiments indicated that deletion of the
The MT-I R/Y sequence (pJ106) did not significantly affect either MT-I basal transcription (Fig. 5B, compare shaded boxes in right panel) or transcription induced by exposure of the transfected cells to 2 μM CdCl₂ (Fig. 5B, compare open boxes in right panel). Thus, the MT-I R/Y sequence does not modulate activity in the context of the basal MT-I promoter in transient transfection experiments.

We next wished to determine the effect of the MT-I R/Y sequence in heterologous contexts, i.e. when positioned in either orientation upstream from the strong constitutive HSV-TK promoter (Fig. 5C, left). Plasmid pRL-TK contains the HSV-TK promoter and is used to normalize the data from the constructs containing the MT-I R/Y sequence. Plasmids pJ111 and pJ112 contain the MT-I R/Y sequence in either orientation immediately upstream of the HSV-TK promoter (Fig. 5C, left). As shown in Figure 5C (compare shaded boxes in right panel) plasmids pJ111 and pJ112 were not significantly different with respect to promoter activity when compared to pRL-TK (containing the isolated HSV-TK promoter). This analysis again indicates that the MT-I R/Y sequence is not an important regulator of the transcriptional activity of the HSV-TK promoter in transient transfection experiments.

In a final study, we determined if proximity of the MT-I R/Y sequence to the MT-I promoter might reveal regulatory effects. For these experiments we used promoter/reporter constructs that contained the MT-I basal promoter and MREs (based on pJ164; Fig. 5D). Plasmids pJ165 and pJ166 contained the MT-I R/Y sequence in either orientation immediately upstream of the MT-I basal promoter and MREs (Fig. 5D, left). In the absence of metal induction, all promoter constructs displayed similar activities (Fig. 5D, compare shaded boxes in right panel). When the transfected cells were induced with 2 μM CdCl₂, a slight increase in promoter activity was observed when the MT-I R/Y sequence was present (Fig. 5D, compare open boxes in right panel). This modest increase in promoter activity under induced conditions was observed for both orientations of the MT-I R/Y sequence.

Thus, although the MT-I R/Y sequence is capable of adopting a non-B DNA conformation in vitro, transient transfection

Figure 4. Summary of reactivities to chemical probes and a plausible H-DNA conformation for the MT-I R/Y sequence. Chemical reactivity data are mapped onto an H-3 H-DNA structural model. White letters within black boxes indicate nucleotides hyperreactive to one or more chemical probes. White ‘+’ within black boxes indicate slight protection of guanosines from DMS modification.

Figure 5. Transient transfection analysis to monitor the effect of the MT-I R/Y sequence on transcriptional activities from various promoter constructs. (A) Example of data from a single transfection experiment. Data points reflect the average firefly luciferase activity (light units normalized to light units encoded by a Renilla luciferase internal control) from duplicate wells. Standard deviations are indicated. Slopes of linear fits to the data (normalized test reporter activity per μg test construct) were averaged to provide the data in (B)–(D). (B)–(D) Plasmid constructs and data from transient transfection analyses. Vector DNA (thin line), MT-I upstream promoter sequence (white rectangle), herpes simplex virus thymidine kinase (HSV-TK) promoter (shaded rectangle), MT-I R/Y sequence (hatched arrow), and transcriptional start sites (black arrows) are shown. Promoter activity per μg template is shown at right (normalized to the first construct in each set). Filled rectangles indicate basal promoter activities. Open rectangles indicate promoter activities in the presence of 2 μM CdCl₂.
analysis indicates that this sequence does not play a detectable role in transcriptional regulation under the conditions studied.

DISCUSSION

Characteristics of the MT-I R/Y sequence

The best-studied examples of DNA sequences capable of adopting H-DNA structures possess perfect mirror symmetry and are often repetitive in nature [e.g. d(GA)16-18, (14,17,20,36); d(G)30, (37) and other di-, tri- or tetra-nucleotide repeats, (22,38)]. In this work we wished to characterize aspects of the structure and any detectable function of a remarkable homopurine/homopyrimidine sequence located upstream of the MT-I gene. The 128 bp MT-I R/Y sequence has multiple small mirror repeats, but does not contain any large regions of perfect mirror symmetry. This element was of interest because of its exceptional length and provocative position upstream of a well-studied gene. The extreme purine/pyrimidine strand bias within this long element suggests that it did not originate through a random process, and that it might have a biological function.

Comparison with previous studies

The MT-I R/Y sequence was originally identified by Bacolla and Wu (10) and probed at low pH using mung bean nuclease (10). By subcloning and sequencing DNA fragments generated by nuclease treatment, these authors concluded that non-B DNA structures could be detected under conditions of negative supercoiling and low pH, and further proposed that H-DNA structures were present in more than one configuration (10). We revisited these structural issues in the present study using a more comprehensive approach. Low resolution S1 nuclease analysis and high resolution chemical probing experiments confirm that the formation of one or more non-B DNA structures is stabilized by negative supercoiling and acidic pH. Traces of base unpairing in supercoiled plasmids were detected even at physiological pH.

Our chemical probing data are in partial agreement with the results previously reported by Bacolla and Wu (10). Differences might be explained in part by the different reagents used (18–23). Our analysis lead us to propose a single plausible folded H-DNA structure that is generally consistent with the chemical reactivity data (Fig. 4). Although we propose a single structure for this element, it is possible that multiple structures exist in equilibrium within an element of this size and sequence complexity. Our model corresponds to the common H-y3 isomer of H-DNA in which a partially-protonated pyrimidine strand is donated to a duplex, consistent with the obvious pH-dependence of the structure. Previous studies have shown the H-y3 isomer of H-DNA to be intrinsically more stable than the H-y5 isomer (36).

Much of our data support the proposed model, while certain hyperreactivities are more difficult to accommodate within this single folded structure. For example, hyperreactive bases between –1172 and –1140 of the pyrimidine strand are perhaps unexpected, but might be due to a slight unwinding of the DNA duplex to accommodate the third strand in the major groove. The CAA...