Regulation in Vitro of Metallothionein Gene Binding Factors

CARL SEGUIN AND DEAN H. HAMER*

Mouse nuclear factors that bind to an upstream metal regulatory element of the mouse metallothionein-I gene have been identified by DNA footprinting and oligonucleotide band shift assays. The formation of complexes at this site can be activated 20- to 40-fold by the in vitro addition of ionic cadmium. The activation reaction is rapid, reversible by a metal chelator, and may involve multiple proteins. These results suggest that the initial step in cadmium detoxification is an interaction between the metal and nuclear DNA-binding factors leading to an increase in metallothionein gene transcription. The ability to observe metal activation in vitro makes this a powerful system to study the biochemistry of eukaryotic gene regulation.

THE TRANSCRIPTION OF EUKARYOTIC genes can be activated or repressed by various intra- and extracellular signals including nutrients, ions, toxins, and hormones. Cells can recognize the presence or absence of such signal molecules and transmit this information to the transcriptional machinery by, for example, (i) increasing the synthesis and hence concentration of a transcriptional regulatory protein (*1*); (ii) increasing the affinity of the regulatory molecule for its target DNA sequence, either directly through a ligand-protein interaction or indirectly through some metabolic route (*2*); or (iii) altering the ability of the regulatory molecule to interact with other components of the transcriptional apparatus (*3*). In distinguishing among such pathways it is critical to develop in vitro systems that are responsive to the signal molecules. Although many such systems have been established for prokaryotes, few if any are available in eukaryotic organisms.

The metallothioneins (MT's) provide an especially useful system for studying eukaryotic gene regulation because the synthesis of these small metal-binding proteins is inducible, at the transcriptional level, by many agents, such as heavy metal ions including cadmium, zinc, and copper, glucocorticoid hormones, α-interferon, interleukin-1, and still unidentified products of activated *Ha-ras* oncogenes and of inflammatory stress. Metal ions are the most general and potent of these various inducers and are the only agent for which regulation plays a known physiological role, namely protection of cells against metal toxicity (*4*).

The ability of vertebrate MT genes to be induced by heavy metals is controlled in the cis figuration by a short DNA sequence present in multiple imperfect copies in the 5′ flanking region (*5, 6*). In vivo competition experiments suggested that a positively acting transcription factor interacts with these metal regulatory DNA sequences (*7*), but the biochemical nature of this trans-acting factor has not previously been addressed. We now report the detection of a mouse nuclear factor (or factors) that binds to an MT gene metal control sequence and demonstrate that this interaction can be regulated in vitro by cadmium.

The arrangement and sequences of the mouse MT-I gene metal regulatory elements (MRE's) are shown in Fig. 1. Experiments in which synthetic MRE sequences were placed upstream of the TATA box of the mouse MT-I gene or the heterologous HSV-TK gene have revealed that different MRE's have different transcriptional efficiencies. MREd is the strongest, MREa and MREc are 50 to 80 percent weaker, MREb is very weak, and MREe is apparently nonfunctional. Presumably these differences reflect variations in the sequences of the MRE's outside of the conserved core element TGCPuC (Pu, purine). Studies with MREa have also shown that multiple copies of the regulatory sequence give a better response than a single copy (*6*).

To detect interactions between mouse nuclear factors and MRE sequences, we initially performed DNA protection (footprinting) experiments (*8*) with exonuclease III and deoxyribonuclease I (DNase) (*8*) using nuclear extracts from mouse L cells grown in the absence or presence of an inducing dose of cadmium for 10 hours. In these experiments, the noninduced cell extract (extract −/E) was prepared with a final dialysis buffer containing 0.1 m*M* EDTA, a metal chelator, to minimize problems due to heavy metal contamination or exchange reactions. In contrast, the induced cell extract (extract +/Cd) was prepared with all buffers containing 5 μ*M* $CdCl_2$ to stabilize potential cadmium-protein interactions. These extracts were briefly incubated with a trace amount of a ^{32}P-labeled MT-I gene fragment in the presence of excess Hae III–digested φX174 RF DNA to compete for nonspecific binding, then digested with exonuclease III or DNase I, deproteinized, and subjected to electrophoresis on a sequencing gel to reveal protected regions of the DNA.

In exonuclease III footprinting of the noninduced (−/E) and induced (+/Cd) cell extracts with a fragment ^{32}P-labeled on the bottom strand of the MT-I gene at position +64 and extending to position −200 (Fig. 2A), the only obvious difference between the −/E and +/Cd extracts was the presence of a new band observed only with the cadmium-induced extract (arrow) although both extracts protected a number of bands not found in the naked DNA control. This DNA fragment had a length of 217 bases, thus mapping the 5′ border of the binding site to position −153. The specificity of this band was demonstrated by its efficient com-

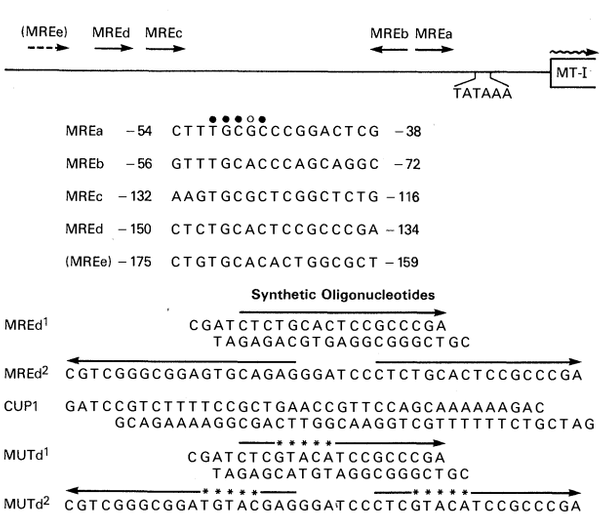

Fig. 1. Sequences of metal regulatory elements and synthetic oligonucleotides. The top line shows the arrangement of the four functional metal regulatory elements (MREa to d) and the apparently nonfunctional element (MREe) of the mouse MT-I gene (*5, 6*). The sequences of the MRE's are shown with dots indicating the conserved core sequence TGCPuC (Pu, purine). Oligonucleotides MREd[1], CUP1, and MUTd[1] were synthesized as two complementary strands, whereas MREd[2] and MUTd[2] were synthesized as self-complementary single strands (*17*). Arrows indicate MREd sequences, and asterisks indicate the positions of base substitutions in the mutant oligonucleotides.

C. Seguin, Molecular Endocrinology Laboratory, Laval University Medical Center, Ste. Foy, Quebec, Canada G1V4GZ.
D. H. Hamer, Laboratory of Biochemistry, National Cancer Institute, Bethesda, MD 20892.

*To whom correspondence should be addressed.

petition by plasmids containing the MT-I promoter but not control plasmids. Mapping of the 3′ border of the site with a probe labeled on the top strand at position −360 and extending to position +64 revealed a band occurring only in the induced cell extract mapped at position −127 (Fig. 2B, bottom of doublet marked by arrow). To determine whether in vivo exposure of cells to cadmium is sufficient to give protection of this binding site, we prepared an extract from cadmium-induced cells with buffers containing EDTA. This +/E extract gave much weaker protection of the −153 band than did the +/Cd extract prepared with $CdCl_2$ in the buffer (Fig. 2C). This suggests that a reversible metal-protein interaction is required for the observed binding activity in the MREd region.

The DNase I footprints (Fig. 2D) revealed two differences between the noninduced −/E and induced +/Cd extracts. (i) The induced cell extract gave decreased cleavage of three bands mapping between −141 and −146. (ii) The induced cell extract also gave increased cleavage of a band at position −153 (Fig. 2E). With exonuclease III, the protected region covers all of MREd together with the distal portion of MREc. With DNase I, the observed differences lie in the middle and just upstream of MREd with the distal boundary corresponding to that obtained with exonuclease III. Thus both enzymes map a metal regulated binding site that overlaps the metal-responsive transcriptional element MREd.

The interaction between nuclear factors and MRE sequences was further analyzed by electrophoretic band shift assays with oligonucleotide probes (*9*). We chose to synthesize oligonucleotides corresponding to MREd because of the footprinting results described above and because it is the strongest element for metal-induced transcription in vivo (*6*). Oligonucleotide $MREd^1$ contains a single copy of the metal regulatory DNA sequence in a double-stranded DNA molecule with short single-stranded termini (Fig. 1). Oligonucleotide $MREd^2$ contains two tail-to-tail copies of the regulatory sequence separated by a Bam HI restriction site. This rotationally symmetric oligonucleotide, which has the capability to exist in solution as either a double-stranded or hairpin structure, was synthesized because precisely the same duplicated sequence has been shown to be functional in vivo (*6*). Oligonucleotides $MUTd^1$ and $MUTd^2$ are mutated versions of $MREd^1$ and $MREd^2$ in which all five of the most strongly conserved nucleotides of the MRE consensus sequence have been altered. Oligonucleotide CUP1, used as a heterologous control, contains an upstream control sequence of the yeast copper metallothionein gene (*10*). These ^{32}P-labeled oligonucleotides were briefly incubated with nuclear extracts, in the absence of any competitor DNA to increase sensitivity, then subjected to electrophoresis through a nondenaturing polyacrylamide gel to separate DNA-protein complexes from free oligonucleotide.

Assays of the −/E noninduced cell extract with EDTA and the +/Cd induced cell extract with $CdCl_2$ after incubation with oligonucleotides $MREd^1$, $MREd^2$, and CUP1 are shown in Fig. 3A. A strong signal (band O) was observed at the origin of the gel. This presumably represents nonspecific DNA binding because it was not dependent on the sequence of the oligonucleotide and was also observed with yeast cell extracts. Oligonucleotide $MREd^2$ gave, in addition to the nonspecific band, two bands (A and B) that entered into the gel. Densitometry of the autoradiogram showed that band A was six times more intense in the induced (+/Cd) than in the uninduced (−/E) extract, whereas band B was the same. $MREd^1$ gave a single band (C) that entered the gel

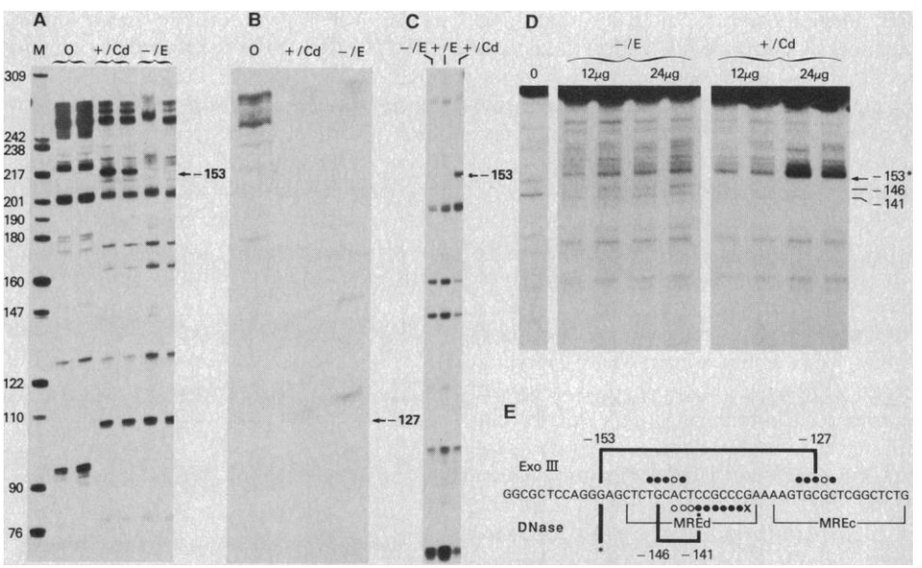

Fig. 2. Footprinting experiments. Binding reactions (50 μl) contained 5 fmol (1000 cpm) of ^{32}P-phosphorylated probe, 250 ng of Hae III–digested ϕX174 DNA, 10 μg of yeast transfer RNA, 1 μg of mixed deoxynucleotides, 8m*M* Hepes (*p*H 7.9), 9 m*M* tris (*p*H 7.9), 65 m*M* NaCl, 3 m*M* $MgCl_2$, 0.5 m*M* DTT, 1.2 m*M* sodium phosphate, *p*H 7.0, 0.2 m*M* PMSF, 11 percent glycerol, the indicated amount of nuclear extract (*18*), and either 0.1 m*M* EDTA or 5 μ*M* $CdCl_2$. After 10 minutes at 24°C, exonuclease III or DNase I was added, and incubation was continued at 30°C for 10 minutes. The DNA was extracted with phenol, precipitated with ethanol, and analyzed on a 5 percent polyacrylamide 8*M* urea sequencing gel. (**A**) Exonuclease III mapping the 5′ border with a probe 5′ end-labeled at +64 and extending to −200. The binding reactions contained buffer alone (lane 0), 54 μg of extract from cadmium-induced cells with $CdCl_2$ in the buffer (lane +/Cd) or 56 μg of extract from noninduced cells with EDTA in the buffer (lane −/E) and were digested with 200 units (left lane in each pair) or 300 units (right lane in each pair) of exonuclease III. Lane M contains pBR322–MspI marker fragments. The arrow indicates the cadmium-induced band mapping at position −153. (**B**) Exonuclease III mapping the 3′ border with a probe 5′ end-labeled at position −360 and extending to position +64. The binding reactions contained the same extracts as in (A) and were digested with 300 units of exonuclease III. The arrow indicates the cadmium-induced band mapping at position −127 as determined by comparison to marker fragments and DNA sequence ladders. (**C**) The effect of preparing an extract from cadmium-induced cells in the presence of EDTA. The binding reactions contained the same probe used in (A) together with 53 μg of extract from noninduced cells with EDTA in the buffer (lane −/E) or from cadmium-induced cells with EDTA in the buffer (lane +/E) or from cadmium-induced cells with $CdCl_2$ in the buffer (lane +/Cd). The mixtures were digested with 400 units of exonuclease III. The arrow to the right indicates the protected band at position −153. (**D**) DNase footprinting. The binding reactions contained 0, 12, or 24 μg of the indicated extract and the same probe used in (A). Digestion was with 25 ng (lane 0 and left lane in each pair) of 50 ng (right lane in each pair) of DNase I. The arrows indicate the enhanced cleavage at −153 and the protected region between −141 and −146. (**E**) Sequence of the protected region. Thin brackets indicate the positions of MREc and MREd. Dots above the sequence indicate the conserved MRE core sequence TGCPuC. Marks below the sequence indicate homology to the consensus Sp1 binding site (*16*); solid dots indicate agreements with unique nucleotides, hollow dots indicate agreements with ambigous nucleotides, and the cross (x) indicates disagreement at an ambiguous site. The heavy brackets indicate the areas protected from exonuclease III and DNase I digestion. The asterisk indicates the enhanced DNase I cleavage site.

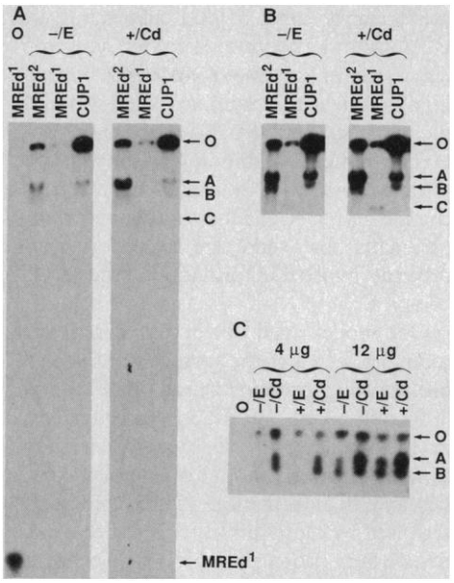

Fig. 3. The oligonucleotide band shift assay. (**A**) Assays of nuclear extracts from noninduced cells prepared with EDTA (−/E) and from cadmium-induced cells prepared with $CdCl_2$ (+/Cd) with various oligonucleotides. The binding reactions (22 μl) contained 0.1 fmol of ^{32}P-labeled oligonucleotide MREd¹ (300 cpm) or MREd² (700 cpm) or CUP1 (700 cpm), 16.7 m*M* Hepes, *p*H 7.9, 50 m*M* NaCl, 0.4 m*M* DTT, 0.4 m*M* PMSF, 16 percent glycerol, 10 μg of nuclear extract, and either 0.1 m*M* EDTA or 5 μ*M* $CdCl_2$. The mixtures were incubated at room temperature for 15 minutes then subjected to nondenaturing polyacrylamide gel electrophoresis (*9*). The lane at the far left contained no extract. The arrows point to the positions of nonspecifically bound material at the origin of the gel (O), specific complexes (A, B, C), and free oligonucleotide (MREd¹). (**B**) A longer exposure of (A) to show band C more clearly. Only the top portion of the gel is shown in this and all subsequent figures. (**C**) Assays of nuclear extracts from noninduced cells prepared with EDTA (−/E) or $CdCl_2$ (−/Cd) and from cadmium-induced cells prepared with EDTA (+/E) or cadmium (+/Cd). The 16-μl reaction mixtures contained 4 or 12 μg of nuclear extract and 0.2 fmol (1400 cpm) of ^{32}P-labeled MREd². The lane at the far left contained no extract.

and this was three times more intense in the +/Cd than in the −/E extract (this is best seen in the longer exposure shown in Fig. 3B). The CUP1 control oligonucleotide yielded two bands entering the gel and these had equal intensities in the two extracts. The formation of all of the bands was abolished by prior treatment of the extract with proteinase K.

To determine whether the difference between the −/E and +/Cd extracts was due to growth in the absence or presence of cadmium as compared to the different preparation buffers that were used, we prepared in parallel extracts from noninduced cells with EDTA (−/E) or cadmium (−/Cd) and extracts from induced cells with EDTA (+/E) or cadmium (+/Cd). These were assayed with oligonucleotide MREd² because it gave the greatest difference between the −/E and +/Cd extracts. Extracts prepared with cadmium (−/Cd and +/Cd) gave a stronger signal than extracts prepared with EDTA (−/E and +/E) regardless of whether or not the cells had been grown in cadmium (Fig. 2C). Hence, in agreement with the footprint results (Fig. 2C), it appeared that cadmium was not inducing the synthesis of complex-forming proteins during cell growth but rather was interacting with a preformed nuclear component.

This result encouraged us to test the effect of adding cadmium in vitro to reaction mixtures containing a noninduced cell extract prepared in buffer with no additions (extract −/0). The presence of $CdCl_2$ in the 15-minute binding reaction strongly stimulated the formation of MREd² band A while having no effect on bands B or O (Fig. 4). Quantitation of the autoradiogram by laser densitometry showed that the intensity of band A was increased 17-fold by 0.8 μ*M* $CdCl_2$ and 23-fold by 2.5 to 80 μ*M* $CdCl_2$. Concentrations of $CdCl_2$ higher than 100 μ*M* inhibited both specific and nonspecific binding, whereas zinc and copper had no effect between 1 and 500 μ*M*. The ability of cadmium to activate band A formation 20- to 40-fold was reproducibly observed in all of the five different −/0 extracts that we used. However, the ratio of the specific complex to nonspecific complexes (bands O and B) varied in different extracts, presumably due to differences in extraction efficiency and the lability of the specific binding proteins.

The sequence specificity of cadmium-activated complex formation was studied with mutated analogs of MREd¹ and MREd² (Fig. 5). The clustered point mutant MUTd¹ gave at least 20-fold less band C than the wild-type MREd¹ and the mutant MUTd² gave at least 20-fold less band A but the same amount of band B as MREd². The activation ratios were 20-fold for MREd¹ band C and 38-fold for MREd² band A. A long exposure of the autoradiogram revealed that the MUTd¹ band C and MUTd² band A were also inducible by cadmium, even though overall binding was very weak. The stronger binding to MREd² than to MUTd² was confirmed by competition experiments.

The effect of the metal chelator EDTA on complex formation was studied in an order of addition experiment (Fig. 6). EDTA at 1.6 to 16 m*M* completely inhibited the formation of MREd² band A in the presence of 83 μ*M* $CdCl_2$ and this effect was independent of the addition. Even when DNA, $CdCl_2$, and the −/0 nuclear extract were first incubated for 10 minutes, the addition of 4 m*M* EDTA for an additional 10 minutes abolished the formation of band A. Thus the cadmium-activated formation of complex A is fully reversible.

The in vitro activation of MREd binding parallels the in vivo induction of MT gene transcription in several respects. First, the concentration of cadmium required for maximal activation in vitro, approximately 1 μ*M*, is well within the range of intracellular cadmium concentrations found in induced cells in vivo (*11*). Second, the extent of activation in vitro is reproducibly 20- to 40-fold which compares favorably with the induction ratios reported for transcription of the endogenous MT-I gene in various mouse cell lines (*12*). Finally, the activation reaction, like transcriptional induction, is rapid. We have focused on cadmium because it is the most potent heavy metal inducer of MT gene transcription both in terms of the required dose and the extent of the response (*4*). It is not yet clear whether other metals, such as zinc and copper, recognize different factors or activate the formation of complexes not detected by our assay systems. Also unclear is whether the same or different factors recognize the metal regulatory sites other than MREd. Previous experiments have shown that MREd is the

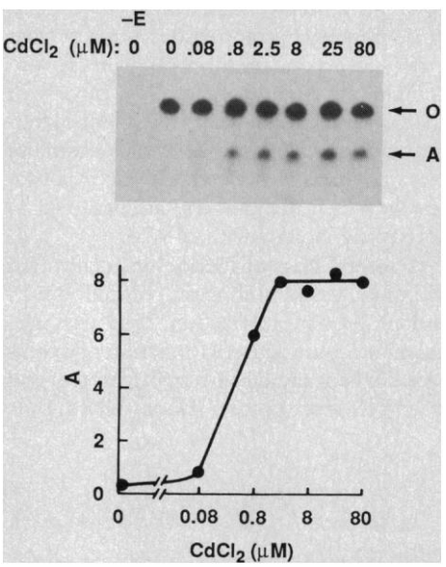

Fig. 4. Cadmium activates the formation of complex A in vitro. The 12-μl reaction mixtures contained 20 μg of nuclear extract from noninduced cells (extract −/0 in buffer III with neither EDTA nor $CdCl_2$), 0.35 fmol (2500 cpm) of ^{32}P-labeled MREd², and the indicated concentrations of $CdCl_2$. The oligonucleotide and $CdCl_2$ were mixed before the nuclear extract was added. (**Top**) The autoradiogram of the oligonucleotide band shift assay. The lane to the far left contained no extract. (**Bottom**) Laser densitometry quantitation of band A (in arbitrary absorbance units) as a function of added $CdCl_2$.

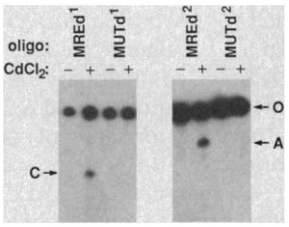

Fig. 5. Sequence specificity of complex formation. The 12-μl reaction mixtures contained 12 μg of nuclear extract from noninduced cells (extract −/0 in buffer III), 1.5 fmol of ^{32}P-labeled MREd[1] (10,000 cpm) or MUTd[1] (10,000 cpm) or MREd[2] (4000 cpm) or MUTd[2] (4000 cpm), and 0 or 83 μ*M* $CdCl_2$ as indicated.

strongest element for in vivo transcription (*6*) and that a DNA fragment containing MREd and MREc is an effective competitor for MT gene transcription factors whereas a fragment containing MREa and MREb is not (*7*). Thus it is possible that the same factors interact with MREa-c but that binding is too weak to be detected by our assays.

We can consider two models for the mechanism by which cadmium activates formation of the protein complex at MREd. One possibility is that cadmium directly interacts with a DNA-binding factor and thereby increases its affinity for the regulatory site. It has been shown that zinc ions are required for specific DNA binding by the *Xenopus* 5*S* transcriptional activator protein TFIIIA, possibly by acting as structural cores for repetitive DNA binding "fingers" (*13*). Sequence comparisons suggest that other eukaryotic regulatory proteins also contain metal-binding sites (*14*). In view of the similar coordination chemistry of cadmium and zinc, it is possible that cadmium plays an analogous structural role in a protein in which the "fingers" are designed to specifically interact with MREd.

A second class of models postulates that cadmium acts by stimulating or inhibiting a protein-protein interaction. For example, cadmium could activate formation of a complex between the DNA-binding protein and a "co-activator" protein thereby altering the conformation and MREd affinity of the DNA-binding protein. The "co-activator" either might or might not also bind to DNA. Conversely, cadmium could stimulate the disassociation of an "anti-activator" protein from the DNA-binding factor. We have some experimental indication from ion exchange and DNA affinity fractionation experiments that there are indeed multiple proteins involved in formation of the MREd complex.

One model ruled out by our data is that cadmium increases the synthesis of the complex-forming proteins during the 10-hour induction period. We found that cells grown in the presence or absence of cadmium contain the same amount of complex activities and that formation of the complex is activated by cadmium in vitro under conditions where protein synthesis would not be expected to occur. These data are consistent with the in vivo observations that induction of MT gene transcription is rapid and occurs in the presence of protein synthesis inhibitors (*15*). Thus it appears that cadmium interacts with a preformed nuclear component.

The mechanism by which cadmium-activated formation of a protein complex at MREd ultimately affects transcription is not known. One intriguing possibility is that a component of the complex contacts a transcription factor such as Sp1 (*16*). The 3′ portion of MREd is highly homologous to the Sp1 consensus binding sequence (Fig. 2D). Moreover, a G to A change at position −133, which generates complete homology to the consensus sequence, converts MREd to a strong constitutive transcriptional signal (*6*). The other MRE's show less homology to the Sp1 site, and it may be that this is the cause of their relative weakness as in vivo transcriptional elements when assayed in the absence of other upstream promoter sequences. While fractionation experiments suggest that the MREd complex does contain multiple components, identification of the proteins involved and characterization of their action requires further purification. The ability to activate formation of the MREd complex in vitro should facilitate detailed biochemical analysis of the interactions that ultimately lead to transcriptional activation.

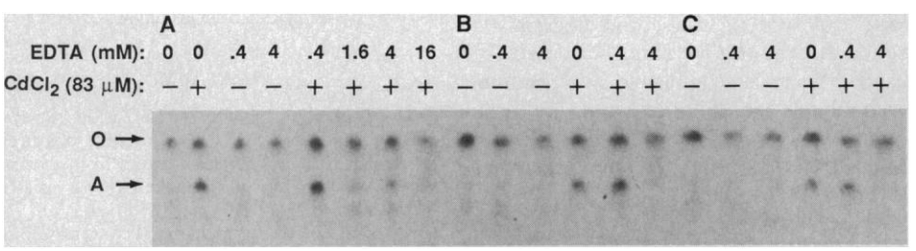

Fig. 6. Complex formation is inhibited by EDTA. The 12-μl reaction mixtures contained 11 μg of nuclear extract from noninduced cells (extract −/0 in buffer III), 0.5 fmol (450 cpm) of ^{32}P-labeled MREd[2], the indicated concentration of EDTA and either 0 or 83 μ*M* $CdCl_2$. The order in which the reagents were added to the DNA was varied as follows: (**A**) EDTA at 0 minutes, $CdCl_2$ at 10 minutes, extract at 11 minutes. (**B**) $CdCl_2$ at 0 minutes, EDTA at 10 minutes, extract at 11 minutes. (**C**) $CdCl_2$ at 0 minutes, extract at 10 minutes, EDTA at 21 minutes; electrophoresis in all cases was at 31 minutes.

REFERENCES AND NOTES

1. For example, general amino acid control in yeast [G. R. Fink, *Cell* **45**, 155 (1986)].
2. For example, steroid hormone receptors [K. Yamamoto, *Annu. Rev. Genet.* **19**, 209 (1985)].
3. A possible example is found in the yeast GAL genes [E. Giniger, S. Varnum, M. Ptashne, *Cell* **40**, 767 (1985)].
4. J. H. R. Kagi and M. Nordberg, *Metallothionein* (Birkhauser, Basel, 1979); D. H. Hamer, *Annu. Rev. Biochem.* **55**, 913 (1986).
5. A. D. Carter *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7392 (1984); M. Karin *et al.*, *Nature (London)* **308**, 513 (1984); G. W. Stuart, P. F. Searle, H. Y. Chen, R. I. Brinster, R. D. Palmiter, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7318 (1984).
6. P. F. Searle, G. Stuart, R. H. Palmiter, *Mol. Cell Biol.* **5**, 1480 (1985); G. W. Stuart, P. F. Searle, R. D. Palmiter, *Nature (London)* **317**, 828 (1985).
7. C. Seguin, B. K. Felber, A. D. Carter, D. H. Hamer, *Nature (London)* **312**, 781 (1984).
8. C. Wu, *ibid.* **317**, 84 (1985); D. J. Galas and A. Schmitz, *Nucleic Acids Res.* **5**, 3157 (1978).
9. M. M. Garner and A. Revzin, *Nucleic Acids. Res.* **9**, 3047 (1981); M. Fried and D. M. Crothers, *ibid.*, p. 6525; R. Schneider, I. Gonder, R. Muller, E. L. Winnacker, *ibid.* **14**, 1303 (1986).
10. D. J. Thiele and D. H. Hamer, *Mol. Cell. Biol.* **6**, 1158 (1986).
11. C. E. Hilderbrand, J. K. Griffith, R. A. Tobey, R. A. Walters, M. D. Enger, in *Biological Roles of Metallothionein*, E. Foulkes, Ed. (Elsevier, New York, 1982), pp. 279–303.
12. K. Mayo and R. D. Palmiter, *J. Biol. Chem.* **256**, 2621 (1981); D. H. Hamer and M. J. Walling, *J. Mol. Appl. Genet.* **1**, 273 (1982).
13. J. S. Hanas, D. J. Hazuda, D. F. Bogenhagen, F. Y. H. Wu, C. W. Wu, *J. Biol. Chem.* **258**, 14120 (1983); J. Miller, A. D. McLachlan, A. Klug, *EMBO J.* **6**, 1609 (1985).
14. J. M. Berg, *Science* **232**, 485 (1986).
15. D. M. Durnam and P. D. Palmiter, *J. Biol. Chem.* **256**, 5712 (1981); M. Karin, R. D. Anderson, E. Slater, K. Smith, H. R. Herschman, *Nature (London)* **286**, 295 (1980).
16. W. S. Dynan and R. Tjian, *Nature (London)* **316**, 774 (1985); K. A. Jones and R. Tjian, *ibid.* **317**, 179 (1985).
17. Oligonucleotides were synthesized by the phosphoramidite method, purified on 12 percent polyacrylamide 8*M* urea gels, then labeled with polynucleotide kinase and [γ-^{32}P]ATP to 700 cpm/fmol. Equimolar quantities of the two strands of MREd[1], CUP1 and MUTd[1], were mixed, boiled, and the double-stranded oligonucleotides were purified by electrophoresis through a nondenaturing 15 percent polyacrylamide gel. Oligonucleotides MREd[2] and MUTd[2] were synthesized as self-complementary single-stranded molecules which were boiled, annealed, and submitted to native gel electrophoresis as above. Typically 20 percent of the DNA migrated at approximately the same position as the double-stranded *CUP1* oligonucleotide and 80 percent migrated more rapidly at a rate slightly slower than the double-stranded MREd[1] oligonucleotide. We assume that the upper band represents double-stranded molecules and that the lower band represents hairpin structures. The upper band was eluted and used for binding studies.
18. An exponentially growing culture of suspension mouse L cells (5×10^5 to 6×10^5 cells per milliliter) was induced with 5 μ*M* $CdCl_2$ for 10 hours if indicated, collected by centrifugation, rinsed with phosphate-buffered saline (PBS), and frozen at −70°C. To prepare extract −/0, the frozen cell pellet was thawed in 5 volumes of buffer I [10 m*M* Hepes, *p*H 7.9, 10 m*M* KCl, 1.5 m*M* $MgCl_2$, 0.5 m*M* DTT (dithiothreitol), 0.5 m*M* PMSF (phenylmethylsulfonyl fluoride)] and disrupted in a Dounce A homogenizer [J. D. Dignam *et al.*, *Nucleic Acids Res.* **11**, 1475 (1983)]. Nuclei were collected by centrif-

ugation, rinsed in buffer I, resuspended in 1.7 volumes of buffer II (10 m*M* Hepes, *p*H 7.9, 0.4*M* NaCl, 1.5 m*M* $MgCl_2$, 0.5 m*M* DTT, 0.5 m*M* PMSF, 5 percent glycerol), agitated for 30 minutes, and then centrifuged 50 minutes at 34,000 rpm in a Beckman Ti50 rotor (*8*). The supernatant was dialyzed overnight against buffer III (20 m*M* Hepes, *p*H 7.9, 50 m*M* NaCl, 0.5 m*M* DTT, 0.5 m*M* PMSF and 20 percent glycerol), recentrifuged, and stored in equal portions at −70°C. Chelated extracts were prepared with 0.1 m*M* EGTA in buffer II and 0.1 m*M* EDTA in buffer III, whereas $CdCl_2$-treated extracts were prepared with 5 μ*M* $CdCl_2$ in all buffers.

19. We thank MaryJane Walling and Guy Rimnitz for technical support, Carl Wu for advice on exonuclease III footprinting, Michael Brownstein for oligonucleotide synthesis, Shankar Adhya and Dennis Thiele for comments on the manuscript, and Nadine Horne for secretarial assistance. Supported in part by a Centennial Fellowship from the Medical Research Council of Canada (C.S.).

2 September 1986; accepted 10 December 1986

Megabase-Scale Mapping of the HLA Gene Complex by Pulsed Field Gel Electrophoresis

SIMON K. LAWRANCE,* CASSANDRA L. SMITH, RAKESH SRIVASTAVA, CHARLES R. CANTOR, SHERMAN M. WEISSMAN

In the study of the genetic structure of mammalian chromosomes, there exists a "resolution gap" between molecular cloning experiments and meiotic linkage analyses. This gap has discouraged attempts to construct full-scale genetic maps of mammalian chromosomes. The organization of the human major histocompatibility complex was examined within this range by pulsed field gel electrophoresis. The data obtained indicate that the complex spans over 3000 kilobases and enable the construction of a megabase-scale molecular map. These results indicate that the techniques employed in DNA extraction, enzymatic digestion, electrophoresis, and hybridization are suitable for the efficient analysis of megabase regions of mammalian chromosomes and effectively bridge the resolution gap between molecular cloning and classical genetics.

THE MAJOR HISTOCOMPATIBILITY complex (MHC) is an extended multigene family composed of several classes of genes, many of which play central roles in the function of the immune system. The multiplicity of phenotypes controlled by the MHC has made the solution of its genetic structure a major goal of immunogeneticists. Considerable progress has been made in studies of both the human (HLA) and the murine (H-2) MHCs. Meiotic linkage analyses have defined three major regions and several subregions of the MHC (*1*). Molecular cloning and chromosome walking experiments have resolved the structure of individual genes and, in some cases, the organization of subregions (*2, 3*). Nevertheless, the limits of resolution of each technique have precluded the synthesis of a full-scale structural map of the HLA complex.

Recently, several methods for the molecular analysis of DNA sequences separated in the genome by distances of up to 1000 kb have been developed (*4, 5*). Pulsed field gel electrophoresis (PFG) has been used to separate and analyze DNA fragments as large as yeast chromosomes (*6*) and to generate a full-scale molecular map of the *Escherichia coli* genome (*7*). This report describes the extension of these methodologies to the human genome in an analysis of the organization of the human MHC.

The fragments of mammalian DNA generated by the majority of restriction enzymes are one to two orders of magnitude smaller than the size range necessary for effective mapping of large chromosomal regions such as HLA. The enzymes used in this study were initially selected as those which recognize eight–base pair cleavage sites and/or contain CpG in their recognition sequences. Because of the rarity of CpG in mammalian genomes, these enzymes were expected to generate fragments that are much larger than would be predicted statistically in random sequence DNA. Figure 1 shows PFG fractionation of human DNA digested with the restriction nucleases Not I, Sal I, and Nru I. Linear concatemers of λ *vir* DNA, coelectrophoresed with the digests of human DNA, enable precise determinations of the sizes of the human DNA fragments. In contrast to conventional gel electrophoresis, in which linear ladders of DNA become

S. K. Lawrance, R. Srivastava, S. M. Weissman, Department of Human Genetics, Yale University School of Medicine, New Haven, CT 06510.
C. L. Smith and C. R. Cantor, Department of Genetics and Development, Columbia University, New York, NY 10032.

*Present address: Department of Immunology, IMM-8, Scripps Clinic and Research Foundation, La Jolla, CA 92037.

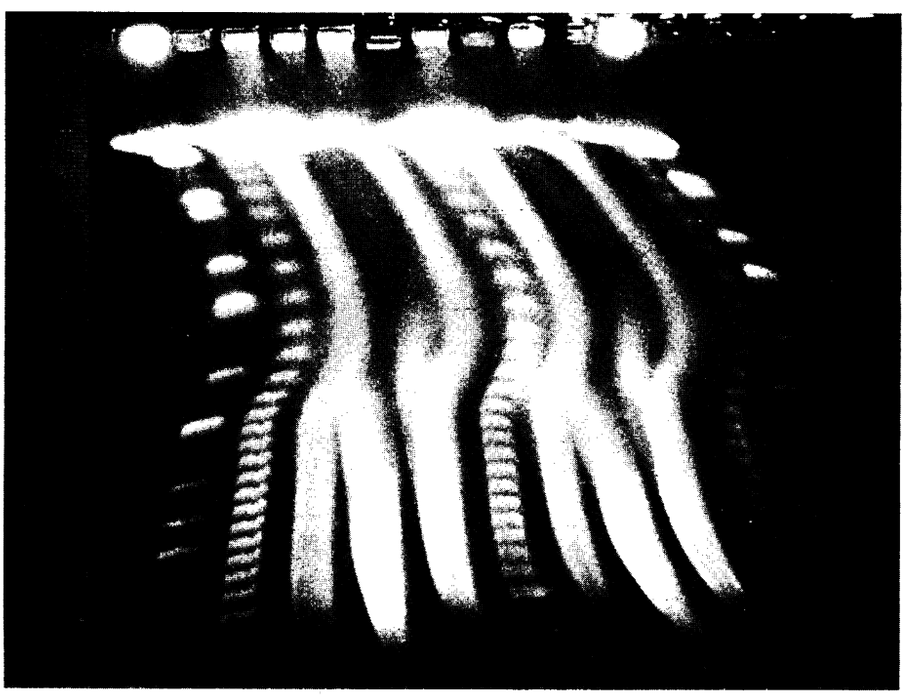

Fig. 1. Pulsed field gel electrophoresis of human DNA. The samples were separated in a 55-cm double, inhomogenous PFG apparatus for 72 hours at 500 volts with 2-minute pulses between 90° field reorientations. The lanes from left to right are: yeast chromosomal DNAs (*Saccharomyces cerevisiae*, strain DBY728); concatemers of λ *vir* (42.5-kb monomer); human (3.1.0) DNA digested with Not I, with Sal I, and with Nru I; λ *vir* concatemers; human (GM3104A) DNA digested with Not I, with Sal I, and with Nru I; λ *vir* concatemers; and yeast chromosomal DNAs. Further details of the DNA preparations, digestion, and electrophoresis are as described (*4*).