Metallothionein genes in *Drosophila melanogaster* constitute a dual system

(cDNA sequence/differential regulation)

RAJA MOKDAD*[†], ALAIN DEBEC[‡], AND MAURICE WEGNEZ*[†]

*Centre de Génétique Moléculaire, Centre National de la Recherche Scientifique, 91190 Gif-sur-Yvette; [†]Laboratoire d'Embryologie Moléculaire, Unité Associée 1134 du Centre National de la Recherche Scientifique, Université Paris 11, bâtiment 445, 91405 Orsay; and [‡]Laboratoire de Zoologie, Université Paris 6, 7 Quai Saint Bernard, 75005 Paris, France

Communicated by J.-M. Lehn, December 29, 1986 (received for review September 17, 1986)

ABSTRACT We have selected a metallothionein (MT) cDNA clone from a cadmium-resistant *Drosophila melanogaster* cell line. This clone includes an open reading frame coding for a 43-amino acid protein whose characteristics are a high cysteine content (12 cysteines, 28% of all residues) and a lack of aromatic amino acids. This protein differs markedly from the *Drosophila* MT (*Mtn* gene) previously reported [Lastowski-Perry, D., Otto, E. & Maroni, G. (1985) *J. Biol. Chem.* 260, 1527–1530]. The MT system of *Drosophila* thus consists of at least two distantly related genes, in sharp contrast with vertebrate MT systems, in which the different members of MT gene families display high similarity. The gene corresponding to our MT cDNA (*Mto*) is inducible in *Drosophila* cell lines and in both larval and adult flies.

Metallothioneins (MTs) are small metal-binding proteins found in eukaryotic cells and probably involved in zinc ion homeostasis. These proteins are mainly described as copper MTs in lower eukaryotes (*Neurospora*, *Saccharomyces*) and as zinc or cadmium MTs in higher eukaryotes (vertebrates), in reference to the predominant type of metallic ions that are bound to these molecules *in vivo* (1). The significance of this difference is not yet clear, since the exact role of MT is not known (2). Most of our knowledge about MT comes from work done on vertebrates. A better insight into the function of these ubiquitous proteins could be achieved by studying species such as *Drosophila melanogaster* that offer an easy genetical approach.

We have previously established a stable cadmium-resistant *Drosophila* cell line whose cells produce high levels of MT (3). Starting with mRNA from these cells, a cDNA clone has been obtained and sequenced; it displays the main characteristics of MT cDNA but differs markedly from another *Drosophila* MT cDNA (4). The two MT cDNAs code for proteins of different sizes (40 and 43 amino acids, respectively) that are very different in sequence. We compare in this paper the MT system of *Drosophila* with the systems of other eukaryotes.

The terminology used in sequence comparisons is very often ambiguous. In this paper, we use the term *homology* (*homologous*) to indicate that different sequences have evolved from a common ancestral sequence. When comparing and quantifying the degree of similarity between DNA or protein sequences, we use the term *isology* (*isologous*), which lacks implications about evolutionary relationships. This term was initially used for the comparison of chemical structures (5).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

***D. melanogaster* Cell Lines.** The two cell lines used in this study, the D line (cadmium sensitive) and the Cd R200 line (cadmium resistant), are characterized in the preceding paper (3).

Nucleic Acid Isolation. Total cellular RNA was prepared according to Auffray and Rougeon (6) by lysis of cells in 8 M urea/4 M LiCl/10% NaDodSO₄/10 mM sodium acetate, pH 5.0. RNA was precipitated from this solution at 4°C for 48 hr. RNA from larvae and adults was extracted from batches of 50 individuals disrupted at 0°C in 1 ml of 100 mM NaCl/10 mM EDTA/0.5% NaDodSO₄/10 mM Tris·HCl, pH 7.4, and then extracted with phenol/chloroform (7). RNA was precipitated at 4°C by addition of LiCl to 4 M. The method of Aviv and Leder (8) was used to purify poly(A)⁺ RNA. Genomic DNA was prepared according to the method of Junakovic (9).

***In Vitro* Translation.** Both total cellular RNA and poly(A)⁺ RNA were translated in a wheat germ cell-free system as described by the supplier (Amersham). [³⁵S]Cysteine (1000 Ci/mmol, New England Nuclear; 1 Ci = 37 GBq) was used as the radioactively labeled amino acid. After translation, the temperature was raised to 65°C for 10 min. Thermoresistant proteins were then carboxymethylated and electrophoresed on 8 M urea/0.1% NaDodSO₄/20% polyacrylamide gels (10).

Construction and Screening of a Cd R200 cDNA Library. A cDNA library was constructed starting from 5 µg of poly(A)⁺ RNA purified from Cd R200 cells (11). Insertion of the cDNA at the *Eco*RI site of vector plasmid pUC8 was performed after addition of *Eco*RI linkers (12). The protocol of Mandel and Higa (13) was used to transform *Escherichia coli* strain JA221 (14). The colony hybridization method of Grunstein and Hogness (15) was employed to identify cDNA plasmids exhibiting hybridization with a Cd R200 [³²P]cDNA probe but not with a D [³²P]cDNA probe.

RNA and DNA Gel Blotting and Hybridization. RNA was electrophoresed on 1.2% agarose/formaldehyde gels and blotted as described by Thomas (16) onto Schleicher & Schuell BA 85 nitrocellulose membranes. DNA was electrophoresed on 0.7% agarose gels and blotted as described by Smith and Summers (17). Hybridization conditions were those of Wahl *et al.* (18). Nick-translated probes (10⁸ cpm/µg) were used at 30 ng/ml.

DNA Sequencing. Plasmid DNA was purified by the method of Birnboim and Doly (19). Restriction fragments to be sequenced were dephosphorylated, fractionated on 5% polyacrylamide gels, and labeled at their 5' ends by using T4 polynucleotide kinase (20). Labeling at the 3' end was done with the Klenow fragment of DNA polymerase (11). Fragments with a unique ³²P-labeled end were obtained by either

Abbreviation: MT, metallothionein.

strand separation or restriction digestion. Sequencing reactions were performed according to the Maxam and Gilbert procedure (20).

RESULTS

Cloning and Sequencing of a *Drosophila* MT. The cadmium-resistant Cd R200 cell line derives from the cadmium-sensitive D line (3). Poly(A)⁺ RNA purified from both cell lines was assayed in a wheat germ translation system for the presence of MT mRNA. We used [³⁵S]cysteine, the predominant amino acid component of MT, as a radioactive label. A low molecular weight protein, approximately 5000, was shown to be the major translational product of the Cd R200 mRNA (Fig. 1, lane 2). This protein was not detected in the control experiment with mRNA from D cells (Fig. 1, lane 4). Addition of CdCl₂ (30 μM) to the incubation mixture increased the amount of [³⁵S]cysteine incorporated into MT by approximately 3-fold (Fig. 1, lane 3). An "induction" at the translational level has already been observed in the case of the rat MT mRNA (21). This may conceivably result from an increase in the rate of release of MT from ribosomes upon binding of cadmium ions to nascent polypeptide chains. Cadmium could also stabilize MT against oxidation and degradation.

Putative MT clones were isolated by differential screening of a Cd R200 cDNA library constructed in the pUC8 plasmid. Five out of 1000 clones gave a very intense signal when probed with Cd R200 [³²P]cDNA and no signal when probed with D [³²P]cDNA (data not shown). Two of the five clones, all of which cross-hybridized, were selected for sequencing. The strategy of sequencing and the derived sequence are shown in Fig. 2. This 301-base-pair (bp) sequence is the combination of pCd2 and pCd14 sequences. Both clones share an identical segment of 191 bp, thus allowing the derivation of a single sequence that contains an open reading

FIG. 1. *In vitro* translation of poly(A)⁺ RNA from D and Cd R200 cells. The poly(A)⁺ RNA was translated, using [³⁵S]cysteine, in a cell-free wheat germ system. The products were analyzed, after carboxymethylation, on a 20% polyacrylamide gel. Translation conditions: lane 1, control without added RNA; lane 2, 5 μg of Cd R200 poly(A)⁺ RNA; lane 3, 5 μg of Cd R200 poly(A)⁺ RNA + 30 μM CdCl₂; lane 4, 5 μg of D poly(A)⁺ RNA. The arrow indicates the main translational product of Cd R200 poly(A)⁺ RNA.

frame corresponding to a 43-amino acid protein (Fig. 2). This protein displays the main structural characteristics of MT—i.e., a high cysteine content (12 cysteines, 28% of all residues) and no leucine, isoleucine, histidine, or aromatic amino acids. This MT-encoding sequence will be referred to as *Mto* in the next part of the report.

Comparison of *Mto* with Other MT-Encoding Sequences. We have used the program EPURE (22) to compare the amino acid sequence of *Mto*-encoded MT with the other reported MT sequences (23). From all binary comparisons, we found only two alignments of sequences that appear as significant. The first was found for a segment spanning 24

FIG. 2. Primary structure of *Mto* cDNA. (A) Strategy used for determining *Mto* cDNA sequence. The upper line shows the combined restriction map of pCd2 (213 bp) and pCd14 (279 bp). The thicker line corresponds to the open reading frame. The arrows below both pCd14 and pCd2 indicate the direction and extent of sequenced fragments. (B) Combined nucleotide sequence (301 bp) of the antisense strand of pCd14 and pCd2. The deduced protein sequence is written below the open reading frame. The three restriction sites shown in A are underlined. The two AACAAG (see the text) and the polyadenylation site are overlined.

amino acids of the two *Drosophila* MTs (Cys⁶ to Cys²⁹ in *Mto* product, Cys⁵ to Cys²⁸ in *Mtn* product; Fig. 3A). The second alignment is between *Drosophila* and *Scylla* (crab) MTs (Fig. 3A).

The program EPURE has also been used to compare the sequences of *Mtn* and *Mto* cDNAs at the nucleotide level (Fig. 3B). In this case, alignment was found to start at the codon corresponding to the first amino acid of the segment defined from protein comparisons and to fall one codon short of the TGA stop codon of the *Mtn* sequence and three codons short of the TAA stop codon of *Mto* (Fig. 3). No alignment of sequences can be proposed in the 5' and 3' noncoding regions. However, a tandemly duplicated 6-nucleotide box (AACAAG) does occur in the 5' noncoding region of both sequences. Respective positions, relative to ATG initiation codons, are -39 in *Mtn* (4) and -67 in *Mto* (Fig. 2).

Expression of *Mto*. Inducibility of *Mto* in cell lines, larvae, and adults. Expression of *Mto* in Cd R200 cells was assayed by probing "Northern" blots (16, 18) with pCd14. An abundant RNA species, of approximately 400 nucleotides, was detected starting from either total or poly(A)⁺ cellular RNA (data not shown). This mRNA was not detectable in D cells grown under normal conditions but was clearly induced in these cells after addition of CdCl₂ to 10 μ M in the culture medium (Fig. 4). The *Mto* mRNA concentration, low 4 hr after cadmium treatment, increased sharply between 4 and 8 hr and then rose slowly up to 48 hr. *Mto* mRNA levels in 10 μ M CdCl₂-induced D cells (48 hr) and in Cd R200 cells grown in the presence of 200 μ M CdCl₂ were compared by dot-blot analysis (data not shown). In these conditions, Cd R200 cells were found to display 3 to 4 times more *Mto* mRNA than induced D cells. Wild-type *Drosophila* larvae and adults do not express *Mto* at a detectable level, but the gene was shown to be inducible in both stages after addition of cadmium to the culture medium (data not shown).

Differential response of *Mto* to cadmium and copper induction. The two distinct *Drosophila* MT cDNAs were cloned starting from very different systems, either copper-induced larvae (*Mtn*; ref. 4) or cadmium-resistant cells (*Mto*; this work). This suggested the possibility of specific induction for each of the two MT genes. To test this possibility directly, we looked for a differential induction of *Mto* in response to cadmium or copper ions. As shown in Fig. 5, *Mto* is poorly inducible by copper in D cells. The amount of *Mto* mRNA

FIG. 4. Time course of accumulation of *Mto* mRNA in cadmium-induced D cells. D cells were cultured in medium supplemented with 10 μ M CdCl₂ for various lengths of time. Total RNA (30 μ g per lane) was fractionated on 1.2% agarose gels and probed with pCd14 (*Mto*). RNA concentration in each lane was measured by probing with λ Dm 1632, which includes *Drosophila* actin II and actin III genes (26). Hybridization with actin mRNAs [1.7 and 1.9 kilobases (kb)] is shown in A (16-hr autoradiography) and hybridization with *Mto* mRNA (0.4 kb) in B (72-hr autoradiography). As shown in A, there was a slight excess of RNA in lane 3. Cadmium treatment was 2, 4, 8, 16, 24, 32, and 48 hr for lanes 1-7, respectively.

detectable in these cells grown in the presence of 1 mM CuSO₄ is negligible compared with the mRNA level in cells induced with 10 μ M cadmium. There is also a significant, but less dramatic, difference in inducibility of *Mto* by cadmium or copper in adult *Drosophila*. In this case, similar levels of *Mto* mRNA could be detected when copper concentration was 5 times that of cadmium (data not shown).

DISCUSSION

We report in this paper the sequence of a *Drosophila* cDNA clone (referred to as *Mto*) that has an open reading frame coding for a protein possessing the main structural characteristics of MTs. The structure of this clone differs significantly from that reported by Lastowski-Perry et al. (4) for another *Drosophila* MT (referred to as *Mtn*). The genome of *Drosophila* thus contains at least two very different MT

FIG. 3. Comparison of *Mto* and *Mtn* sequences and their deduced amino acid sequences. (A) Amino acid sequence comparison of *Mto* (this work), *Mtn* (4), and *Scylla* (24) MTs. Protein sequences are written with the single-letter code (25). The two upper arrows indicate the borders of the two *Drosophila* sequences, which were aligned according to EPURE (22). Asterisks between *Mto* and *Mtn* sequences mark shared residues. Triangles above *Mto* and *Mtn* indicate cysteine residues that are not shared by both sequences. Double asterisks between *Scylla* MT-I and *Drosophila* MT point to residues common to the three sequences. Single asterisks between *Scylla* and *Drosophila* MTs show the residues that are common to the crab and one of the two fly MTs. Triangles below the *Scylla* sequence indicate cysteine residues found only in that sequence. (B) cDNA sequence comparison of *Mto* and *Mtn*. Sequences are aligned according to the data yielded by EPURE. The two sequences correspond to the amino acid sequences underlined in A. Asterisks denote shared nucleotides.

FIG. 5. Differential induction of *Mto* by cadmium and copper ions. D cells were grown for 48 hr in the presence of the following salts: lane 1, 10 μ M CdCl₂; lane 2, 10 μ M CuSO₄; lane 3, 100 μ M CuSO₄; lane 4, 1000 μ M CuSO₄. The cells were viable after incubation in all ionic concentrations. Total RNA was extracted, fractionated on a 1.2% agarose gel (30 μ g per lane), and probed first with pCd14 (A) and then with λ Dm 1632 (B) to measure RNA concentrations. Autoradiography was 16 hr in A and 72 hr in B.

genes. This result is in sharp contrast with the other MT systems studied so far, where MT gene families consist of members displaying high levels of isology.

Evolution of MT Genes in *D. melanogaster*. The two *Drosophila* MTs, *Mtn* and *Mto*, are remarkably different in sequence (Fig. 3). They differ in size (40 and 43 amino acids, respectively), number of cysteine residues (10 and 12, respectively), and amino acid composition (no valine residue in *Mtn*, two valine residues in *Mto*). Comparison of *Mtn* and *Mto* nevertheless allows us to align both sequences without the need of introducing gaps (Fig. 3). It can be calculated, from this alignment, that *Mto* and *Mtn* proteins are 26% isologous. This low level of isology explains why Maroni *et al.* (27) did not detect *Mto* by using *Mtn* as a probe and why we did not detect *Mtn* by using the *Mto* probe. It is, however, likely that the two genes are homologous. If so, then the duplication of the ancestral MT gene that led to *Mtn* and *Mto* was a very ancient evolutionary event. It is interesting to note that the *Scylla* MT-I (24) appears as intermediate between the two *Drosophila* MTs. As shown in Fig. 3, seven amino acids, all of which are cysteine residues, are common to the three sequences. Five additional amino acids are found at similar positions in *Mtn* and *Scylla* MT-I and four other in *Mto* and *Scylla* MT-I (Fig. 3). A similar result is obtained when the *Scylla* MT-II (24) is compared with the two *Drosophila* sequences. This suggests that duplication of the ancestral *Drosophila* MT gene may have occurred after separation of the crustacean and insect evolutionary lines.

These results contrast with the data reported for vertebrates, in which the different MTs within any species are much alike, having the same number and the same positions of cysteine residues as well as exactly the same size (1, 23). Isology levels between the different MTs (61 amino acids) within any vertebrate species range from 77% to 92%. A similar situation is found for the two MT isoforms from the crustacean *Scylla* (24). In this case the two MTs slightly differ in size (57 and 58 amino acids, respectively) but are still 81% isologous.

Expression of *Mto*. We show in this paper that cadmium ions strongly induce *Mto* mRNA synthesis in the D cell line as well as in adults and larvae. As shown in Fig. 4, the time course of *Mto* mRNA accumulation in D cells, after cadmium induction, is similar to that observed in other MT systems (28). These results are in good agreement with the characterization in the same cells of a protein displaying all MT properties (3). The cadmium-resistant cell line (Cd R200) was

shown to display MT levels 22 times higher than the cadmium-induced D line (3). Here we show that *Mto* mRNA levels are only 3 to 4 times higher in Cd R200 than in induced D cells. A duplication of *Mto* in Cd R200 cells (preliminary data) most probably accounts for the increased level of *Mto* mRNA in these cells. The much higher induction at the protein level could be due to some translational regulation or to a higher stability of MT in the presence of 200 μ M CdCl₂. Interpretation of the Cd R200 cadmium resistance phenotype will evidently also require the study of *Mtn* expression.

Copper ions, in contrast, are much less efficient *Mto* inducers than cadmium ions. This is particularly striking in the case of D cells, which respond very poorly to copper induction. These results are difficult to compare with those reported for *Mtn*, which is highly copper-inducible (4). Both genes have been studied in completely different conditions—i.e., mainly cell lines in the case of *Mto* and mainly larvae in the case of *Mtn*. At first sight, it could appear that *Mto* and *Mtn* are characterized by strongly different inducing properties. They have both been cloned by differential screening. Copper-induced larvae were the starting material used in the case of *Mtn* (4) and cadmium-induced cells in the case of *Mto* (this work). It was possible, in each case, to select only a single type of *Drosophila* MT. Lower eukaryotes, such as *Neurospora* and *Saccharomyces*, possess copper-inducible MT genes (29, 30). In contrast, higher eukaryotes such as crustaceans (24), sea urchins (23), and vertebrates (1) possess MT genes inducible mainly by cadmium or zinc. *Drosophila*, thus, could possibly possess both types of MT genes. The following data show that this view of the *Drosophila* MT system is perhaps too simple. As demonstrated by Lastowski-Perry *et al.* (4), cadmium ions are also able to induce *Mtn* in larvae. In the case of *Mto*, copper ions are very poor inducers in D cells but are more efficient in adults and larvae. The study of both *Mtn* and *Mto* expression in cell lines and in the different developmental stages is thus obviously required to reach a correct interpretation of the *Drosophila* MT system.

We thank F. Michel, J. P. Muller, and M. L. Samson for helpful discussions and F. Peronnet for the λ Dm 1632 probe. The manuscript was much improved by the criticisms of F. Michel, M. L. Samson, and M. Samuels. This work was supported by Grant 83.C.0677 from the Ministère de l'Industrie et de la Recherche.

- Kägi, J. H. R. & Nordberg, M., eds. (1979) *Metallothionein* (Birkhäuser, Basel).
- Karin, M. (1985) *Cell* 41, 9–10.
- Debec, A., Mokdad, R. & Wegnez, M. (1985) *Biochem. Biophys. Res. Commun.* 127, 143–152.
- Lastowski-Perry, D., Otto, E. & Maroni, G. (1985) *J. Biol. Chem.* 260, 1527–1530.
- Florkin, M. (1966) *Aspects Moléculaires de l'Adaptation et de la Phylogénie* (Masson, Paris).
- Auffray, C. & Rougeon, F. (1980) *Eur. J. Biochem.* 107, 303–314.
- Terracol, R. (1986) *Genet. Res.* 48, 167–174.
- Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* 69, 1408–1412.
- Junakovic, N. (1980) *Nucleic Acids Res.* 8, 3611–3622.
- Laemmli, U. K. (1970) *Nature (London)* 227, 680–685.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
- Kurtz, D. T. & Nicodemus, C. F. (1981) *Gene* 13, 145–152.
- Mandel, M. & Higa, A. (1970) *J. Mol. Biol.* 53, 159–162.
- Beggs, J. D. (1978) *Nature (London)* 275, 104–109.
- Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* 72, 3961–3965.
- Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* 77, 5201–5205.
- Smith, G. E. & Summers, M. D. (1980) *Anal. Biochem.* 109, 123–129.

18. Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3683–3687.
19. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1523.
20. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
21. Ohi, S., Cardenosa, G., Pine, R. & Huang, P. C. (1981) *J. Biol. Chem.* **256**, 2180–2184.
22. Michel, F., Jacquier, A. & Dujon, B. (1982) *Biochimie* **64**, 867–881.
23. Nemer, M., Wilkinson, D. G., Travaglini, E. C., Sternberg, E. J. & Butt, T. R. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4992–4994.
24. Lerch, K., Ammer, D. & Olafson, R. W. (1982) *J. Biol. Chem.* **257**, 2420–2426.
25. IUPAC–IUB Commission on Biochemical Nomenclature (1968) *J. Biol. Chem.* **243**, 3557–3559.
26. Couderc, J. L., Sobrier, M. L., Giraud, G., Becker, J. L. & Dastugue, B. (1983) *J. Mol. Biol.* **164**, 419–430.
27. Maroni, G., Otto, E. & Lastowski-Perry, D. (1986) *Genetics* **112**, 493–504.
28. Searle, P. F., Davison, B. L., Stuart, G. W., Wilkie, T. M., Norstedt, G. & Palmiter, R. D. (1984) *Mol. Cell. Biol.* **4**, 1221–1230.
29. Münger, K., Germann, U. A. & Lerch, K. (1985) *EMBO J.* **4**, 2665–2668.
30. Fogel, S. & Welch, J. W. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5342–5346.