

Identification and characterization of metastasis-associated gene/protein 1 (MTA1)

Yasushi Toh · Garth L. Nicolson

Published online: 15 October 2014
© Springer Science+Business Media New York 2014

Abstract Metastasis is a complex series of sequential events involving several gene products and the regulated expression of several tumor cell genes. Using rat mammary adenocarcinoma cell lines of differing metastatic potentials and a differential complementary DNA (cDNA) hybridization method, our laboratory embarked in 1992 on a project to identify candidate metastasis-associated genes. Among the genes that were found to be abundantly overexpressed in highly metastatic rat cell lines compared to poorly metastatic cell lines, we identified a completely novel gene without any homologous or related genes in the database in 1994. The full-length cDNA of this gene was cloned, sequenced, and named *mta1* (metastasis-associated gene 1), and eventually, its human cDNA counterpart, *MTA1*, was also cloned and sequenced by our group. *MTA1* has now been identified as one of the members of a gene family (*MTA* gene family). The products of the *MTA* genes, the MTA proteins, are transcriptional co-regulators that function in histone deacetylation and nucleosome remodeling. In this review, we will briefly discuss the researches for the identification and characterization of the *mta1* gene, its human counterpart *MTA1*, and their protein products.

Keywords Metastasis-associated gene/protein family · MTA proteins · Chromatin remodeling · Histone deacetylation · Human cancers · Nucleosome remodeling and histone deacetylation complex (NuRD)

Y. Toh (✉)
Department of Gastroenterological Surgery, National Kyushu Cancer Center, 3-1-1 Notame, Minami-ku, Fukuoka 811-1395, Japan
e-mail: ytoh@nk-cc.go.jp

G. L. Nicolson
Department of Molecular Pathology, The Institute for Molecular Medicine, Huntington Beach, CA 92647, USA

1 Introduction

Metastasis is a complex series of events that involves several gene products, including those important for the invasion and detachment of neoplastic cells from the primary tumor, penetration into blood and lymphatics, arrest at distant sites, adhesion to endothelial cells, extravasation, induction of angiogenesis, evasion of host antitumor responses, and growth at metastatic sites [1, 2]. The regulated expression of several tumor cell genes is thought to be important in this process [3].

During an attempt to identify candidate metastasis-associated genes in rat mammary adenocarcinoma systems in 1994, we first identified *mta1* (metastasis-associated gene 1; rat homologue) complementary DNA (cDNA) as a completely novel gene [4]. Subsequently, we cloned the human homologue of *mta1*, *MTA1* [5], and investigated the expression of human *MTA1* messenger RNA (mRNA) in surgically resected human cancer tissues. We found significant positive correlations between the expression levels of *MTA1* mRNA and several clinicopathological factors related to malignant potential [6, 7]. In this brief review, we will discuss the researches for the first identification and characterization of *mta1/MTA1* genes and their encoded proteins (Mta1/MTA1), as the discoverers of the *mta1/MTA1* genes.

2 Identification of candidate metastasis-associated genes in the rat mammary adenocarcinoma metastatic system

In 1992, our laboratory embarked a project to identify candidate metastasis-associated genes at the University of Texas M. D. Anderson Cancer Center in Houston, Texas. At that time, several techniques had been used to search for genes involved in the metastatic cascade, including: somatic cell fusion, karyotypic analysis, transfection of isolated genes into recipient cells, and differential or subtraction cDNA hybridization.

Using these techniques, several genes have been identified as differentially expressed and possibly involved in the metastasis of mammary and breast cancers: *mts1*, *nm23*, *stromelysin-3*, among other genes.

We used the 13762NF rat mammary adenocarcinoma cell lines of differing metastatic potentials in experimental and spontaneous metastasis assays [8] to identify possible genes that are involved in the metastatic process. This model system has been well characterized and was known to be similar to human breast adenocarcinoma in many respects, including cytoskeletal organization, cell surface components, and pathologic mode of spread *in vivo* [8, 9]. In this system, gene expression was compared between the nonmetastatic cell line MTC.4 and the highly metastatic cell line MTLn3. MTC.4 was a subclone of the MTC cell line, derived from the primary tumor growing in the mammary fat pad and possessed no ability to metastasize from a primary implant site in the mammary fat pad or colonize tissues after intravenous injection. Line MTLn3 was derived from a spontaneous lung metastasis and was found to be highly spontaneously metastatic from the mammary fat pads of syngeneic rats as well as forming large numbers of lung colonies upon intravenous injection [8]. These cell lines were found to be phenotypically stable during prolonged passage *in vivo* or *in vitro* [8]. Such stability is important for differential hybridization analysis and identification of differentially expressed genes, and it is also essential for gene transfer experiments and metastasis assays.

Using the differential gene expression approach, multiple cDNAs were eventually isolated that were differentially expressed in either poorly or highly metastatic 13762NF lines. After tertiary screening, 10 separate cDNAs were finally identified [10]. Partial sequencing of those cDNAs, followed by a homology search with the GenBank/EMBL data banks, revealed that eight of them were known genes, including annexin I, elongation factor-1 α (EF-1 α), mitogen-activated protein kinase (MAPK), type IV collagenase (72kD) and cathepsin L. Most of these genes had been reported previously as implicated in metastasis of various cancer cells. Interestingly, EF-1 α had been independently identified in a similar differential hybridization approach to search for metastasis-associated genes using a *fos*-transfected metastatic cell line (*fos*-SR-3Y1-202) and a non-transfected, nonmetastatic version of the line [11]. Thus, we judged that this screen had worked well for the purpose of identifying candidate metastasis-associated genes in the 13762NF rat mammary adenocarcinoma system.

Two of 10 cDNAs identified above had no homologous nucleotide sequences in the databank, and one of these cDNAs was designated as clone 10.14. This cDNA clone contained about 2.2 kilobase pairs (kbp). Because its corresponding mRNA was ~3.0 kbp in RNA blotting analysis (to be described below), this was not a full-length cDNA clone. Since complete sequencing of the 10.14 cDNA revealed a

single open reading frame of 583 amino acid extending to the 5'-end of the sequence, and its corresponding amino acid sequence was also completely novel, we were determined to isolate a full-length cDNA of clone 10.14, and then further analyze this gene sequence. At this point, we named this candidate gene *mta1* (metastasis-associated gene 1).

3 Cloning and sequencing of a full-length *mta1* cDNA (10.14)

A full-length *mta1* cDNA (designated 10.14) was isolated by screening the cDNA library of subline MTC.4 cell line, followed by sequencing [4]. The sequence was 2756-bp long and contained a single open reading frame encoding a protein of 703 amino acid residues, which was named Mta1. A computer-assisted homology search was performed for the nucleotide and amino acid sequences at the National Center for Biotechnology and Information (NCBI) and no significant homology was found, indicating that *mta1* was completely a novel gene [4].

To confirm the differential expression of the *mta1* mRNA between the metastatic MTLn3 and nonmetastatic MTC.4 cell lines, we examined its steady-state mRNA levels by RNA blot analysis. This showed that both lines expressed *mta1* mRNA of ~3.0 kb in size and that the expression level was estimated to be fourfold higher in the highly metastatic MTLn3 cell line. Because Southern blotting showed that there were counterparts of the rat *mta1* gene in the human and mouse genomes, we further examined the expression of the human homologue (*MTA1*) of the *mta1* gene in two well-characterized human breast adenocarcinoma cell lines. As a result, these human cell lines also expressed *MTA1* mRNA, the homologue of the *mta1* gene, of approximately the same length as rat *mta1* mRNA.

For comparison, we used the human breast cancer cell line MCF-7, derived from the pleural effusion of a breast cancer patient, as a relatively noninvasive and nonmetastatic cell line in nude mice. Conversely, its malignant derivatives, sublines MCF7/LCC1 and MCF7/LCC2, are invasive and metastatic *in vivo* [12, 13]. The expression ratio of the *MTA1* mRNA in these cells was MCF-7:MCF7/LCC1:MCF7/LCC2=1:2:4. Using another cell line set, the expression of the *MTA1* mRNA in MDA-MB-231 (metastatic in nude mice) was calculated to be ~4 times as high as in MDA-MB-468 cells, which was nonmetastatic in nude mice [14]. Thus, the mRNA expression levels of the human homologue of the rat *mta1* gene also correlated with the metastatic potential found in two human breast cancer metastatic systems. These results suggested that the *mta1* and *MTA1* genes were good candidates for metastasis-associated genes.

The cDNA fragment containing the entire open reading frame of the rat *mta1* gene was inserted into a prokaryotic expression vector, and the protein derived from it was

expressed as a fusion protein with glutathione S-transferase (GST) in *Escherichia coli*. As a result, the fusion protein contained the expected molecular mass of ~108 kDa, including the full-length Mta1 protein of ~80 kDa. This indicated that an open reading frame does exist in the *mta1* cDNA clone. Using antibodies raised against the GST-Mta1 fusion protein or a synthetic oligopeptide (containing amino acid residues 329–343) we performed Western blot analysis with an MTLn3 cell lysate. Both antibodies recognized bands of ~80 kDa in size [4].

The expression levels of the Mta1 protein or its human homologue MTA1 protein were then examined in MTLn3, MTC.4, MCF-7, MCF7/LCC1, MCF7/LCC2, MDA-MB-468 and MDA-MB-231 cell lines, all of which were used previously for Northern blot analysis, as mentioned before. In these experiments, the Mta1/MTA1 protein expression levels correlated well with the metastatic potentials of the various cell lines, and the expression ratios of the protein were quite similar to the expression ratios obtained from the Northern blot analyses [4].

Using homology to the rat *mta1* cDNA, we cloned the human *MTA1* cDNA from a human melanoma A2058 cDNA library [5]. The human *MTA1* gene encoded a putative protein of 715 amino acid residues with a predicted molecular weight of ~82 kDa. The amino acid sequences of the rat and human proteins were compared and found to be 96 % identical and 98 % similar [5]. To assess the extent of evolutionary conservation of the *MTA1* gene, we analyzed genomic DNA of several species by Southern blot analysis and showed that the *MTA1* gene was conserved in human, rat, mouse, dog, cow, rabbit, and chicken [5].

4 Structural analysis of the Mta1/MTA1 protein and the clues to assume its function(s)

Hydropathy plot (Kyte-Doolittle) values for the predicted Mta1 protein did not show any apparent membrane-spanning or membrane-associated regions, nor was there an NH₂-terminal signal sequence. This protein was quite

hydrophilic, suggesting that it is not a cell surface protein, nor is it a secreted protein requiring a signal sequence [15]. Sequence analyses of the Mta1/MTA1 proteins revealed the existence of several common important protein sequence motifs [4, 5, 15–20]. These are illustrated in Fig. 1, which includes: (1) A bromo-adjacent homology (BAH) domain was found at the N-terminal of Mta1/MTA1 protein. This domain has been identified in a variety of proteins involved in transcriptional co-regulation and/or DNA binding and is thought to be involved in protein-protein interactions; (2) A SANT domain was located to the C-terminal site of the BAH domain of Mta1/MTA1 protein. The SANT domain is similar to the DNA binding domain of myb-related proteins and was identified in SWI3 (a yeast component of the SWI/SNF complex), along with ADA2 (a component of the histone deacetylase complex), N-CoR (a nuclear hormone co-repressor) and the TFIIB B subunit (a basal pol III transcription factor in yeast). The SANT domain has also been referred to as the WFY domain, since it has many aromatic amino acid residues at fixed positions; (3) An ELM domain (egl-27 and MTA1 homology) was located between the BAH and the SANT domains of Mta1/MTA1 protein. This has an unknown function, but the ELM-SANT domain of MTA1 has now been shown to be bound to HDAC1 [21]; (4) A zinc-finger motif (Cys-X₂-Cys-X₁₇-Cys-X₂-Cys) belonging to the type found in transcription factors that bind to the GATA sequence that are involved in hematopoiesis and heart development; (5) A leucine-zipper motif (Leu-X₆-Leu-X₆ IsoleuX₆-Leu-X₇-Leu) was found beginning at residue 251 of MTA1; (6) The Mta1/MTA1 protein contains proline-rich sequences at each carboxyl terminal extremity (rat: LPLRPPPPAP and human: LPPRPPPPAP). These amino acid sequences completely matched the consensus sequence for the src-homology3 (SH3) domain-binding motif. The SH3 domains are considered to be important in protein-protein interactions in signal transduction pathways, and they are known to associate with cytoskeletal components as well as their involvement in other protein-protein interactions. Another possible SH3-binding motif (PRPPKPDP) was also observed; (7) Three putative nuclear localization signal sequences were found at the

Fig. 1 Schematic representation of the structural domains or motifs of Mta1/MTA1 protein

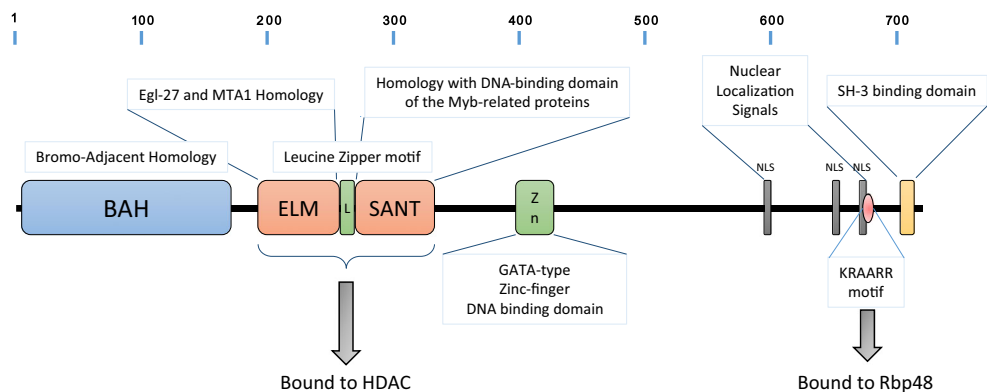




Fig. 2 Sequence alignment of MTA1, MTA2, and MTA3 from various eukaryotes, including human, mouse, rat, frog, and bovine. Known MTA1 domains, BAH, ELM and SANT, and zinc-finger motif are

indicated. The KRAARR motif is boxed in gray. This figure has been reproduced from Alqarni et al. [22] with permission from Dr. Mackay

carboxyl terminus of the Mta1/MTA1 protein; and (8) A KRAARR motif was found near the C-terminal end of MTA1 protein, which has been recently shown to bind to Rbp48 protein in the nucleosome remodeling and histone

deacetylation (NuRD) complex [22]. The first four domains or motifs are well conserved in all three primary members of the MTA family, including MTA1, MTA2, and MTA3 (Fig. 2). The findings on putative protein domains or motifs

strongly suggested that Mta1/MTA1 protein might possibly be involved in transcriptional control.

5 Conclusions

When we first reported on the amino acid sequence of the rat *mta1* cDNA and Mta1 protein in 1994, there were no similar or homologous nucleotide or protein sequences in the databases, suggesting that *mta1* was a completely new, novel gene [4]. By 1999, similar genes or genes with homologous amino acid sequences to Mta1/MTA1 protein appeared in the databases. However, the functions of MTA1 were still unknown at this time.

The first clues on the molecular and biochemical functions of MTA1 were obtained by four different groups from 1998 to 1999 [23–27]. In these studies, two disparate chromatin-modifying activities, ATP-dependent nucleosome remodeling activity and histone deacetylation, were discovered to be functionally and physically linked in the same protein complex. The original complex was named the NuRD (Nucleosome Remodeling and histone Deacetylation) complex, which has been shown to have transcription-repressing activity [24–27]. Furthermore, in order to determine the biologic function of MTA1, we had to wait for the epoch study by the Kumar laboratory where evidence was obtained that directly demonstrated the relationship between the MTA1 protein in the NuRD complex and invasion/metastasis of cancer cells [28].

After identification of the human *MTA1* cDNA, we first showed the clinicopathological significance of its overexpression in human cancer specimens, including esophageal, gastric, and colorectal cancers [6, 7]. Many clinicopathological correlative studies followed our studies and the conclusions obtained from those have been reinforced by additional experiments that show the biological relevance of MTA1 protein overexpression to carcinogenesis and cancer progression. All of these topics mentioned here will be introduced in the other chapters of this special issue.

Acknowledgments The authors acknowledge support from members of Department of Gastroenterological Surgery, National Kyushu Cancer Center, Japan, and foundation and private donations to the Institute for Molecular Medicine.

Conflict of interest The authors declare that they have no conflict of interest.

References

- Liotta, L. A. (1986). Tumor invasion and metastases—role of the extracellular matrix: Rhoads memorial award lecture. *Cancer Research*, 46(1), 1–7.

- Nicolson, G. L. (1988). Cancer metastasis: tumor cell and host organ properties important in metastasis to specific secondary sites. *Biochimica et Biophysica Acta*, 948(2), 175–224.
- Chambers, A. F., & Tuck, A. B. (1993). Ras-responsive genes and tumor metastasis. *Critical Review of Oncogenesis*, 4(2), 95–114.
- Toh, Y., Pencil, S. D., & Nicolson, G. L. (1994). A novel candidate metastasis-associated gene, *mta1*, differentially expressed in highly metastatic mammary adenocarcinoma cell lines. cDNA cloning, expression, and protein analyses. *Journal of Biological Chemistry*, 269(37), 22958–22963.
- Nawa, A., Nishimori, K., Lin, P., Maki, Y., Moue, K., Sawada, H., et al. (2000). Tumor metastasis-associated human MTA1 gene: its deduced protein sequence, localization, and association with breast cancer cell proliferation using antisense phosphorothioate oligonucleotides. *Journal of Cell Biochemistry*, 79(2), 202–212.
- Toh, Y., Oki, E., Oda, S., Tokunaga, E., Ohno, S., Maehara, Y., et al. (1997). Overexpression of the MTA1 gene in gastrointestinal carcinomas: correlation with invasion and metastasis. *International Journal of Cancer*, 74(4), 459–463.
- Toh, Y., Kuwano, H., Mori, M., Nicolson, G. L., & Sugimachi, K. (1999). Overexpression of metastasis-associated MTA1 mRNA in invasive oesophageal carcinomas. *British Journal of Cancer*, 79(11–12), 1723–1726.
- Neri, A., Welch, D., Kawaguchi, T., & Nicolson, G. L. (1982). Development and biologic properties of malignant cell sublines and clones of a spontaneously metastasizing rat mammary adenocarcinoma. *Journal of National Cancer Institute*, 68(3), 507–517.
- Nicolson, G. L., Gallick, G. E., Spohn, W. H., Lembo, T. M., & Tainsky, M. A. (1992). Transfection of activated c-H-rasEJ/psv2neo or psv2neo genes into rat mammary cells: rapid stimulation of clonal diversification of spontaneous metastatic and cell-surface properties. *Oncogene*, 7(6), 1127–1135.
- Pencil, S. D., Toh, Y., & Nicolson, G. L. (1993). Candidate metastasis-associated genes of the rat 13762NF mammary adenocarcinoma. *Breast Cancer Research and Treatment*, 25(2), 165–174.
- Taniguchi, S., Miyamoto, S., Sadano, H., & Kobayashi, H. (1991). Rat elongation factor 1 alpha: sequence of cDNA from a highly metastatic fos-transferred cell line. *Nucleic Acids Research*, 19(24), 6949.
- Thompson, E. W., Brunner, N., Torri, J., Johnson, M. D., Boulay, V., Wright, A., et al. (1993). The invasive and metastatic properties of hormone-independent but hormone-responsive variants of MCF-7 human breast cancer cells. *Clinical and Experimental Metastasis*, 11(1), 15–26.
- Brunner, N., Frandsen, T. L., Holst-Hansen, C., Bei, M., Thompson, E. W., Wakeling, A. E., et al. (1993). MCF7/LCC2: a 4-hydroxytamoxifen resistant human breast cancer variant that retains sensitivity to the steroidal antiestrogen ICI 162,780. *Cancer Research*, 53(14), 3229–3232.
- Zhang, R. D., Fidler, I. J., & Price, J. E. (1991). Relative malignant potential of human breast carcinoma cell lines established from pleural effusions and a brain metastasis. *Invasion and Metastasis*, 11(4), 204–215.
- Toh, Y., Pencil, S. D., & Nicolson, G. L. (1995). Analysis of the complete sequence of the novel metastasis-associated candidate gene, *mta1*, differentially expressed in mammary adenocarcinoma and breast cancer cell lines. *Gene*, 159(1), 97–104.
- Manavathi, B., Singh, K., & Kumar, R. (2007). MTA family of coregulators in nuclear receptor biology and pathology. *Nuclear Receptor Signaling*, 5, 1–8.
- Toh, Y., & Nicolson, G. L. (2009). The role of the MTA family and their encoded proteins in human cancers: molecular functions and clinical implications. *Clinical and Experimental Metastasis*, 26(3), 215–227.
- Toh, Y., & Nicolson, G. L. (2011). MTA1 of the MTA (metastasis-associated) gene family and its encoded proteins: molecular and

- regulatory functions and its role in human cancer progression. *Atlas of Genetics and Cytogenetics in Oncology and Haematology*, 15(3), 303–315.
19. Toh, Y., & Nicolson, G. L. (2013). Signaling pathways of MTA family proteins as regulators of cancer progression and metastasis. In R. R. Resende & H. Ulrich (Eds.), *Trends in stem cell proliferation and cancer research* (pp. 251–275). Dordrecht: Springer.
 20. Singh, R. R., & Kumar, R. (2007). MTA family of transcriptional metaregulators in mammary gland morphogenesis and breast cancer. *Journal of Mammary Gland Biology and Neoplasia*, 12(2–3), 115–125.
 21. Millard, C. J., Watson, P. J., Celardo, I., Gordiyenko, Y., Cowley, S. M., Robinson, C. V., et al. (2013). Class I HDACs share a common mechanism of regulation by inositol phosphates. *Molecular Cell*, 51(1), 57–67.
 22. Alqarni, S. S., Murthy, A., Zhang, W., Przewloka, M. R., Silva, A. P., Watson, A. A., et al. (2014). Insight into the architecture of the NuRD complex: structure of the RbAp48-MTA1 subcomplex. *Journal of Biological Chemistry*, 289(32), 21844–21855.
 23. Bowen, N. J., Fujita, N., Kajita, M., & Wade, P. A. (2004). Mi-2/NuRD: multiple complexes for many purposes. *Biochimica et Biophysica Acta*, 1677(1–3), 52–57.
 24. Tong, J. K., Hassig, C. A., Schnitzler, G. R., Kingston, R. E., & Schreiber, S. L. (1998). Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. *Nature*, 395(6705), 917–921.
 25. Xue, Y., Wong, J., Moreno, G. T., Young, M. K., Cote, J., & Wang, W. (1998). NuRD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Molecular Cell*, 2(6), 851–861.
 26. Zhang, Y., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Bird, A., & Reinberg, D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes and Development*, 13(15), 1924–1935.
 27. Wade, P. A., Geggion, A., Jones, P. L., Ballesta, R. E., Aubry, F., & Wolffe, A. P. (1999). Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nature Genetics*, 23(1), 62–66.
 28. Mazumdar, A., Wang, R. A., Mishra, S. K., Adam, L., Bagheri-Yarmand, R., Mandal, M., et al. (2001). Transcriptional repression of oestrogen receptor by metastasis-associated protein 1 corepressor. *Nature Cell Biology*, 3(1), 30–37.