

# Biochemical characterization of the human placental transcription factor GCMA/1

Chenchen Lin, Meiyao Lin, and Hungwen Chen

**Abstract:** Glial cells missing (GCM) proteins are a novel family of zinc-containing transcription factors. Human GCMA/1 is primarily expressed in placental trophoblast cells and regulates *SYNCYTIN* gene expression, which mediates fusion of cytotrophoblasts to form the syncytiotrophoblast layer of the human placenta. To biochemically characterize the transcriptional activity of GCMA/1, we set up an in vitro transcription system for human GCMA/1 (hGCMA/1). Using G-free reporter constructs carrying multiple copies of wild-type or mutant GCMA-binding site (GBS) in front of a synthetic TATA box, we observed specific transcriptional activities of recombinant hGCMA/1 proteins prepared from a baculovirus – insect cell or *Escherichia coli* expression system. We further characterized GCMA/1-mediated transcriptional activation on the native syncytin promoter. Using G-free reporter constructs containing the native syncytin promoter, a TATA box downstream of the proximal GBS in the syncytin promoter was shown to be essential for the transcription activation directed by hGCMA/1. Therefore, our results demonstrate positive transcriptional activities of GCMA/1 in vitro and provide a better understanding of GCMA/1-mediated *SYNCYTIN* gene expression.

*Key words:* syncytin, transcription factor, GCMA/1, placenta.

**Résumé :** Les protéines « glial cell missing (GCM) » constituent une nouvelle famille de facteurs de transcription contenant du zinc. Le CGMA/1 humain est principalement exprimé dans les trophoblastes placentaires et régule l'expression du gène *SYNCYTIN* dont le produit est responsable de la fusion des cytotrophoblastes en syncytiotrophoblaste. Dans le but de caractériser son activité transcriptionnelle, nous avons élaboré un système de transcription in vitro pour le facteur GCMA/1 humain (hGCMA/1). Des constructions comportant une cassette sans résidus G et l'élément de réponse liant le GCMA (GBS) multimérisé, intact ou muté, clonés en amont d'une boîte TATA synthétique, nous ont permis d'observer une activité transcriptionnelle spécifique de GCMA/1 recombinant, préparé à partir de cellules d'insectes infectées au baculovirus, ou à partir de *Escherichia coli*. Nous avons davantage caractérisé l'activation transcriptionnelle dépendante de GCMA/1 sur le promoteur natif de la syncytine. Des constructions sans résidus G et comportant le promoteur natif de la syncytine nous ont permis de démontrer qu'une boîte TATA située en aval du site GBS proximal du promoteur de la syncytine est essentielle à l'activation de la transcription dirigée par GCMA/1. Ainsi, nos résultats démontrent que GCMA/1 exerce une activité transcriptionnelle positive in vitro et contribuent à une meilleure compréhension de l'expression du gène de la syncytine dépendante de GCMA/1.

*Mots clés :* Syncytine, facteur de transcription, GCMA/1, placenta.

[Traduit par la Rédaction]

## Introduction

The *glial cells missing (GCM)* gene was originally isolated from a *Drosophila melanogaster* mutant line that produces additional neurons at the expense of glial cells (Hosoya et al. 1995; Jones et al. 1995). Conversely, ectopic expression of *GCM* in flies generated excessive numbers of glial cells at

the expense of neurons (Hosoya et al. 1995; Jones et al. 1995). Two *GCM*-like genes (*GCMA/1* and *GCMb/2*) have been reported in fly, mouse, rat, and human and form a novel family of transcription factors, which all share sequence homology at the amino-terminal region (Hosoya et al. 1995; Jones et al. 1995; Alfonso and Jones 2002; Kammerer and Giangrande 2001; Kanemura et al. 1999; Kim et al. 1998). The amino-terminal region of GCM proteins constitutes the DNA-binding domain called the GCM motif. The optimal recognition sequence for the GCM motif is 5'-(A/G)CCC-(T/G)CAT-3' or its 5'-ATG(A/C)GGG(T/C)-3' complement (Akiyama et al. 1996; Schreiber et al. 1997). Mutation analysis further suggested that nucleotide residues 2, 3, 6, and 7 are important for recognition by *Drosophila* and murine GCMA/1 (Schreiber et al. 1998). Although sequence homology is less preserved outside the GCM motif, a transactivation domain was described in the carboxyl terminus of *Drosophila* GCMA/1 protein and two such domains in human and murine GCMA/1 and GCMb/2 proteins, respectively (Akiyama et al. 1996; Schreiber et al. 1997).

Received 13 December 2004. Revision received 3 February 2005. Accepted 4 February 2005. Published on the NRC Research Press Web site at <http://bc.b.nrc.ca> on 21 April 2005.

**C. Lin.** Graduate Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan.

**M. Lin.** Institute of Biological Chemistry, Academia Sinica, Nankang, Taipei 115, Taiwan.

**H. Chen.**<sup>1</sup> Graduate Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan, and Institute of Biological Chemistry, Academia Sinica, Nankang, Taipei 115, Taiwan.

<sup>1</sup>Corresponding author (e-mail: [hwchen@gate.sinica.edu.tw](mailto:hwchen@gate.sinica.edu.tw)).

*Drosophila GCMa/1* mRNA is transiently detected in glial precursors and immature glial cells except for mesectodermal midline glia during a short period of gliogenesis within the central nervous system (Hosoya et al. 1995; Jones et al. 1995). In glial cells, *GCMa/1* turns on *repo* via 11 GCM-binding sites in its promoter region (Akiyama et al. 1996). Repo is a homeodomain-containing transcription factor that is involved in the terminal differentiation of glial cells (Xiong et al. 1994; Campbell et al. 1994; Halter et al. 1995). In contrast, mouse *GCMa/1* mRNA is highly expressed in the labyrinthine trophoblast cells (Basyuk et al. 1999). Furthermore, Nait-Oumesmar et al. (2000) have immunolocalized *GCMa/1* proteins in the syncytio- and cytotrophoblasts of human placenta. *GCMa/1* is required for placental development, since genetic ablation of mouse *GCMa/1* leads to a failure of labyrinthine layer formation and no fusion of trophoblasts to syncytiotrophoblasts (Anson-Cartwright et al. 2000; Schreiber et al. 2000).

The human syncytin membrane protein is encoded by the *ENVELOPE* gene of the human endogenous retrovirus family W (HERV-W) and is specifically expressed in the syncytiotrophoblast layer of the human placenta (Mi et al. 2000). Previously, we have demonstrated that *GCMa/1* could modulate *SYNCYTIN* gene expression via two *GCMa/1*-binding sites (GBSs), dGBS and pGBS, in the 5'-flanking region upstream of the 5'-long terminal repeat (LTR) of HREV-W (Yu et al. 2002).

In vitro transcription reactions using nuclear extracts are particularly useful for studying the role of specific transcription factors in promoter regulation (Sawadogo and Roeder 1985; Gorski et al. 1986). Since human *GCMa/1* (h*GCMa/1*) is a transcription factor primarily expressed in the placenta, an in vitro transcription assay for *GCMa/1* will help to understand the molecular mechanism underlying *GCMa/1*-mediated transcriptional activation and placental gene expression. In this study, we established an in vitro transcription assay for *GCMa/1* using recombinant h*GCMa/1* proteins, HeLa nuclear extracts, and reporter constructs containing multiple copies of pGBS or the syncytin promoter region. In the reactions, the G-free cassette vector was used to generate reporter constructs, (pGBS)<sub>4</sub>TK and (pGBS)<sub>4</sub>TATA, carrying four tandem copies of the pGBS in front of the Herpes simplex virus (HSV) thymidine kinase (TK) promoter and a synthetic TATA box, respectively. We demonstrated that (pGBS)<sub>4</sub>TATA provided a higher signal to noise ratio than (pGBS)<sub>4</sub>TK, and recombinant h*GCMa/1* proteins prepared from both baculovirus – insect cell and *Escherichia coli* expression systems exhibited specific transcriptional activities. Furthermore, we studied *GCMa/1*-mediated transcriptional activation on the native syncytin promoter. Using G-free reporter constructs containing the native syncytin promoter, a TATA box downstream of the proximal GBS in the syncytin promoter was shown to be essential for the transcription activation directed by h*GCMa/1*. Therefore, the established in vitro transcription system offers a very useful tool to study the biochemical mechanism of *GCMa/1*-mediated gene expression.

## Materials and methods

### Plasmid constructs

The (pGBS)<sub>4</sub>TATA reporter plasmid was generated by

cloning four tandem copies of pGBS (5'-ATGAGGGC-3') into the G-free cassette vector, pLovTATA, which contains a synthetic TATA box derived from the HSV TK promoter and the chicken ovalbumin promoter (Klein-Hitpass et al. 1990). The HSV TK promoter was amplified by PCR using the pBLCAT2 vector as template and subcloned into (pGBS)<sub>4</sub>TATA to replace the synthetic TATA box, generating the (pGBS)<sub>4</sub>TK reporter plasmid. The G-free cassette does not contain G residues in its coding strand, which ensured specific detection of transcripts derived from the G-free cassette when in vitro transcription was performed in the absence of GTP or in the presence of 3'-O-methyl-GTP, which is a GTP analogue preventing transcriptional elongation. For syncytin promoter constructs, the 5'-LTR of HREV-W and its 5'-flanking region were PCR amplified using (25468/30953)pCAT or Δp(25468/30953)pCAT as the template (Yu et al. 2002). The pGBS site (nt 28026–28033 in the BAC clone 083M05) was deleted in Δp(25468/30953)pCAT. Genomic fragments corresponding to nt 27950–28244 and nt 27950–28314 in the BAC clone 083M05 were subcloned into the promoterless G-free cassette vector pC<sub>2</sub>AT (Sawadogo and Roeder 1985) to generate (27950–28244)C<sub>2</sub>AT, Δp(27950–28244)C<sub>2</sub>AT, (27950–28314)C<sub>2</sub>AT, and Δp(27950–28314)C<sub>2</sub>AT, respectively (Fig. 1A). In addition, the TATA box (nt 28286–28291) was deleted in the nt 27950–28314 genomic fragment to generate the Δt(27950–28314)C<sub>2</sub>AT construct. Plasmid AdML200 contains an adenovirus major late-promoter element (nt –400 to 10) inserted upstream of a 200-bp G-free cassette.

### Cell culture and nuclear extracts preparation

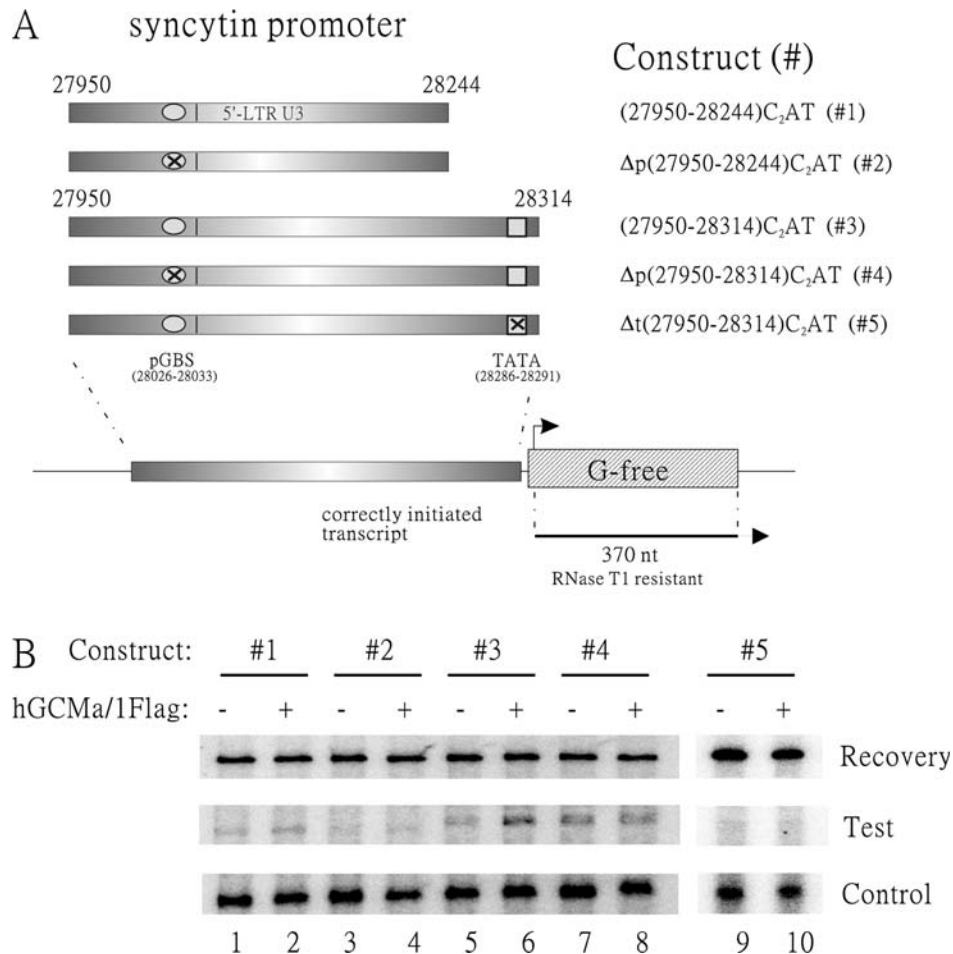
The human cervical cell line HeLa was obtained from the American Type Culture Collection (Manassas, Va.) and maintained at 37 °C in HEPES-buffered Dulbecco's modified minimal essential medium supplemented with 10% fetal bovine serum, streptomycin (100 μg/mL), and penicillin (100 units/mL). Sf-9 cells were maintained at 27 °C in Sf-900II SFM (Invitrogen, Carlsbad, Calif.) supplemented with amphotericin B (0.125 μg/mL), streptomycin (50 μg/mL), and penicillin (50 units/mL). Nuclear extracts of HeLa cells were prepared according to Dignam et al. (1983). Protein concentrations were measured using the BCA protein assay kit (Pierce, Rockford, Ill.).

### Preparation of recombinant *GCMa/1* protein

h*GCMa/1*Flag, a recombinant h*GCMa/1* containing a carboxyl terminal Flag tag, was prepared by the baculovirus – insect cell expression system as described by Yu et al. (2002). In brief, Sf-9 cells were infected with the h*GCMa/1*Flag recombinant baculovirus at a multiplicity of infection of 5. Cells were collected and lysed 48 h postinfection, and h*GCMa/1*Flag proteins were purified with anti-FLAG M2 agarose beads (Sigma, St. Louis, Mo.) according to the manufacturer's instructions. The eluted h*GCMa/1*Flag proteins were further purified with Mono Q chromatography (Amersham, Piscataway, N.J.). The resultant h*GCMa/1*Flag proteins were more than 95% pure as judged by SDS–PAGE.

The h*GCMa/1* cDNA fragment was subcloned into the pMAL-c2 vector (NEB, Beverly, Mass.) for preparation of the recombinant maltose-binding protein (MBP) – h*GCMa/1* fusion protein. In brief, BL21(DE3) harboring the MBP–h*GCMa/1*-

**Fig. 1.** Syncytin promoter analysis by the in vitro transcription assay. (A) Schematic representation of the syncytin promoter constructs. (B) In vitro transcriptional activation by hGCMA/1Flag on the syncytin promoter constructs. One hundred nanograms of reporter plasmid and 6.25 ng of AdML200 were incubated with or without 3 pmol of hGCMA/1Flag in the transcription reactions. The results shown here are from one of three similar experiments.



expressing plasmid was grown in LB-amp (100 μg/mL) until 0.8–1.0 OD<sub>600</sub> and then induced with IPTG (0.3 mmol/L) for 2.5 h at 37 °C. Affinity purification of MBP-hGCMA/1 proteins was performed with maltose agarose beads (NEB) according to the manufacturer's instructions. Purification of hGCMA/1 proteins from protease cleavage of MBP-hGCMA/1 was not successful because the cleaved hGCMA/1 proteins stuck to the agarose beads. Therefore, MBP-hGCMA/1 fusion proteins were used in this study. The eluted MBP-hGCMA/1 proteins were further purified with Mono Q chromatography. The purity of MBP-hGCMA/1 proteins was about 90% by SDS-PAGE.

#### Electrophoretic mobility shift assay (EMSA)

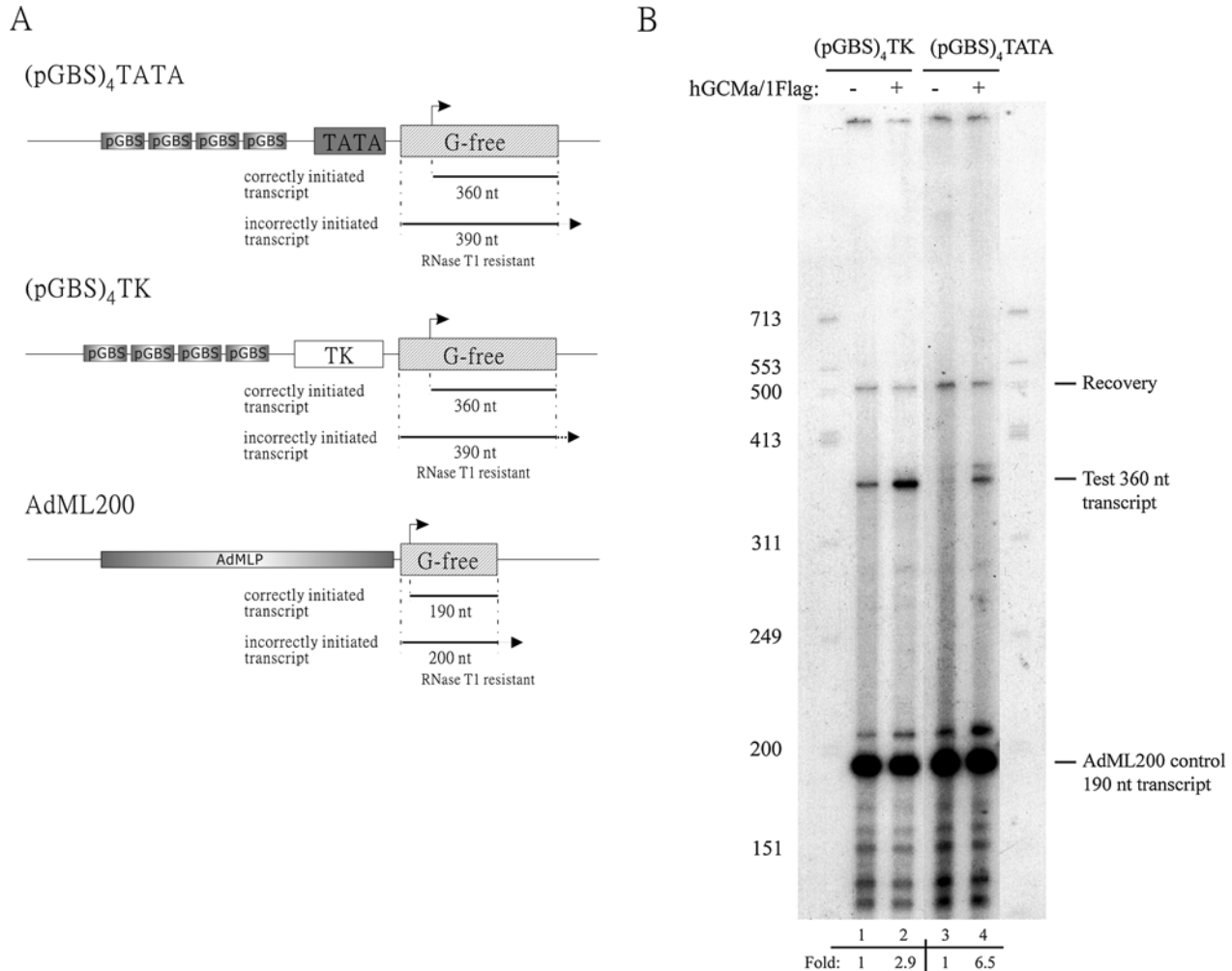
EMSA was performed as previously described (Yu et al. 2002). In brief, the end-labeled oligonucleotide probe (1 ng, (0.5–1) × 10<sup>5</sup> cpm) was incubated for 30 min on ice in binding reaction buffer (50 mmol Tris-HCl/L, pH 8.0, 100 mmol NaCl/L, 2 mmol MgCl<sub>2</sub>/L, 0.05 mmol ZnCl<sub>2</sub>/L, 4 mmol spermidine/L, 0.05% NP-40, 5 mmol DTT/L, and 10% glycerol) containing 0.25 μg of poly(dI-dC), 7.5 μg BSA, and 20 ng of hGCMA/1Flag or MBP-hGCMA/1 protein. After incubation, the reaction mixtures were analyzed by electrophoresis on 5% nondenaturing polyacrylamide gels in running buffer (25 mmol Tris-HCl/L, pH 8.5, 190 mmol glycine/L, and

1 mmol EDTA/L) at 4 °C. For competition experiments, the unlabeled oligonucleotides were used at 100-fold molar excess and incubated with hGCMA/1Flag or MBP-hGCMA/1 protein for 20 min on ice before adding radiolabeled probe. The pGBS oligonucleotide (5'-TTCTGGGATGAGGGCAAACG-3') containing the core pGBS sequence and short flanking sequences was synthesized for probe preparation. A mutant pGBS oligonucleotide, Mut (5'-TTCTGGGATGATAGCAAACG-3') containing mutated nucleotide residues important for recognition by GCMA/1, was also synthesized as a negative control.

#### In vitro transcription reaction

The transcription reaction was performed by preincubating the reporter plasmid and recombinant hGCMA/1 in assay buffer (10 mmol HEPES-KOH/L, pH 7.9, 3 mmol MgCl<sub>2</sub>/L, and 50 μmol ZnCl<sub>2</sub>/L) on ice for 15 min. The reaction was then supplemented with 12.5 or 6.25 ng (Fig. 1B) of AdML200, 0.5 μg of herring sperm DNA, 40 μg of HeLa nuclear extracts, 0.5 mmol ATP/L, 0.5 mmol CTP/L, 0.01 mmol UTP/L, 10 U of RNase T1, 2 mmol DTT/L, 0.4 mol creatine phosphate/L, 40 μCi of α-[<sup>32</sup>P]UTP (1 μCi = 37 kBq), 10 U of RNasin, and 1 mmol 3'-O-methyl-GTP/L (Fig. 2B) or 0.5 mmol 3'-O-methyl-GTP/L (Figs. 1 and 3) in

**Fig. 2.** Analysis of reporter constructs for in vitro transcription of hGCMA/1. (A) Schematic representation of the reporter constructs (pGBS)<sub>4</sub>TATA and (pGBS)<sub>4</sub>TK and the control plasmid AdML200. (B) (pGBS)<sub>4</sub>TATA and (pGBS)<sub>4</sub>TK exhibit differential reactivity to hGCMA/1Flag. One hundred nanograms of reporter plasmid and 12.5 ng of AdML200 were incubated with or without 3 pmol of hGCMA/1Flag in the transcription reactions. hGCMA/1 directed stronger transcriptional enhancement in (pGBS)<sub>4</sub>TATA (lanes 3 and 4) than in (pGBS)<sub>4</sub>TK (lanes 1 and 2). Relative transcriptional activity of hGCMA/1Flag was calculated as the ratio of the test transcript signal to the control transcript signal. Fold stands for fold increase of normalized transcription activation of the G-free reporter gene by hGCMA/1 relative to the reaction with reporter plasmid only (lanes 1 and 3). The results shown are from one of two similar experiments.



a final reaction volume of 25  $\mu$ L and incubated at 30  $^{\circ}$ C for 1 h. The reaction was stopped by addition of 100  $\mu$ L of stop buffer (25 mmol Tris-HCl/L, pH 7.5, 0.5% SDS, 10 mmol EDTA/L, 20  $\mu$ g of yeast tRNA, and 40  $\mu$ g of protease K) and incubated at 37  $^{\circ}$ C for 30 min. The reaction was then combined with 200  $\mu$ L of urea solution (10 mmol Tris-HCl/L, pH 7.5, 1 mmol EDTA/L, and 8 mol urea/L) and 1000 cpm recovery DNA and extracted with phenol-chloroform twice. The recovery DNA is a 501-bp DNA fragment derived from *Nde*I- and *Xba*I-digested pRc/CMV (Invitrogen) and <sup>32</sup>P labeled using a Klenow fill-in reaction. The reaction products were precipitated with ethanol and analyzed in 5% sequencing gels.

## Results

### Characterization of recombinant hGCMA/1 proteins

Recombinant hGCMA/1 proteins, hGCMA/1Flag and MBP-hGCMA/1, were prepared in insect and *E. coli* cells, respec-

tively (Fig. 4A). The biological activities of both recombinant proteins were first characterized by their binding to radioactive pGBS probes in EMSA. As shown in Fig. 4B, specific DNA-protein complexes were formed between hGCMA/1Flag and pGBS (compare lanes 3 and 4) and between MBP-hGCMA/1 and pGBS (compare lanes 6 and 7). As a negative control, the MBP protein did not bind to pGBS (lane 2). Furthermore, no significant competition was observed when Mut, a mutant pGBS containing mutated nucleotide residues important for recognition by GCMA/1, was added at a 100-fold molar ratio over the pGBS probe for the recombinant proteins (lanes 5 and 8). Therefore, both hGCMA/1Flag and MBP-hGCMA/1 were functional in terms of DNA-binding activity.

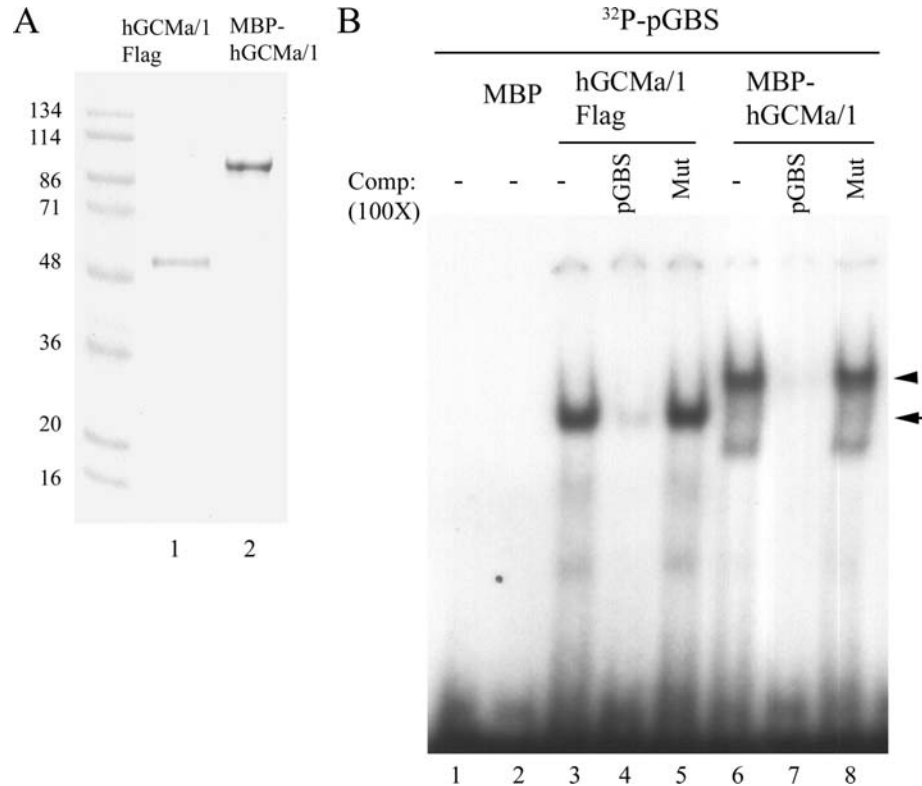
### Analysis of reporter constructs

To study in vitro transcription of GCMA/1, two G-free reporter constructs, (pGBS)<sub>4</sub>TK and (pGBS)<sub>4</sub>TATA, were generated and tested for their reaction to hGCMA/1Flag (Fig. 2A). Transcription activation stimulated by hGCMA/1Flag was





**Fig. 4.** Characterization of recombinant hGCMA/1 proteins. (A) Approximately 0.3  $\mu$ g of purified hGCMA/1Flag and 0.5  $\mu$ g of purified MBP-hGCMA/1 were analyzed by SDS-PAGE and detected by Coomassie Blue R-250 staining. Numbers on the left of the figure indicate the molecular mass (kDa) of protein markers. (B) EMSA of recombinant hGCMA/1 proteins and the radiolabeled pGBS probe. Twenty nanograms of hGCMA/1Flag or MBP-hGCMA/1 protein was incubated with radiolabeled probe and analyzed as described in the Materials and methods. Lane 1 contains probe only. Comp, presence or absence (-) of unlabeled pGBS or Mut oligonucleotide at 100-fold (100 $\times$ ) excess. The arrow and arrowhead indicate the hGCMA/1Flag-DNA complex and MBP-hGCMA/1-DNA complex, respectively.



observed with both reporter constructs (Fig. 2B, lanes 2 and 4). However, the basal activity of (pGBS)<sub>4</sub>TK was higher than that of (pGBS)<sub>4</sub>TATA in the absence of hGCMA/1Flag (Fig. 2B, compare lanes 1 and 3). We normalized hGCMA/1Flag transcription activation with that of the internal control AdML200. Accordingly, 2.9- and 6.5-fold increases in transcriptional activation of the G-free reporter gene by hGCMA/1Flag were observed for (pGBS)<sub>4</sub>TK and (pGBS)<sub>4</sub>TATA, respectively (Fig. 2B). Since the basal activity of the TK promoter may have compromised the detection sensitivity of *in vitro* transcription reactions, we used (pGBS)<sub>4</sub>TATA as the reporter construct in the following experiments.

#### In vitro transcription reaction of hGCMA/1

To test whether recombinant hGCMA/1 protein prepared from *E. coli* was also transcriptionally active, we tested the transcription activity of MBP-hGCMA/1 on (pGBS)<sub>4</sub>TATA. As shown in Fig. 3A, MBP-hGCMA/1 increased 5.1-fold reporter gene expression relative to the mock (compare lanes 1 and 2). Replacement of MBP-hGCMA/1 with MBP in the *in vitro* transcription reaction did not stimulate reporter gene expression (Fig. 3A, compare lanes 1 and 4). Furthermore, a higher activation by MBP-hGCMA/1 was observed on (pGBS)<sub>4</sub>TATA than on (Mut)<sub>5</sub>TATA, which contains five tandem copies of Mut (Fig. 3A, compare lanes 2 and 3). These results suggested that MBP-hGCMA/1 activates the reporter gene ex-

pression by selectively binding to the wild-type pGBS site on the reporter plasmid. Similarly, hGCMA/1Flag exhibited a higher specific transcription activation on (pGBS)<sub>4</sub>TATA than on (Mut)<sub>5</sub>TATA (data not shown).

We further tested the effects of increasing amounts of hGCMA/1Flag or (pGBS)<sub>4</sub>TATA on transcription activation of the G-free reporter gene. With a fixed amount of (pGBS)<sub>4</sub>TATA, addition of 0.6, 1.5, and 3 pmol of hGCMA/1Flag led to 4.3-, 5.2-, and 6.5-fold increases in transcription activation of the G-free reporter gene relative to the mock (Fig. 3B, lanes 1–4). Furthermore, transcription reactions with a fixed amount of hGCMA/1Flag and increasing amounts of (pGBS)<sub>4</sub>TATA also gradually increased transcription activation of the G-free reporter gene (Fig. 3C). The transcription reaction was dependent on RNA polymerase II because no transcripts were detected in the presence of  $\alpha$ -amanitin, an RNA polymerase II inhibitor (Fig. 3B, lane 5). Interestingly, doubling the pGBS sites in the reporter construct, ((pGBS)<sub>4</sub>)<sub>2</sub>TATA, further enhanced transcription activation of the G-free reporter gene by 9.3-fold relative to the mock (Fig. 3B, lanes 6 and 7).

#### Syncytin promoter analysis by the *in vitro* transcription assay

We next examined the transcriptional activity of GCMA/1 on syncytin promoter by the established *in vitro* transcription assay. Because different transcription initiation sites

(nt 28244 or nt 28314 in the bacterial artificial chromosome clone 083M05) of the *SYNCYTIN* gene have been reported (Blond et al. 1999; Cheng et al. 2004), two G-free reporter constructs, (27950–28244) $C_2AT$  and (27950–28314) $C_2AT$ , containing genomic fragments covering different reported transcription initiation sites were tested in *in vitro* transcription reactions (Fig. 1A). In the presence of hGCMa/1Flag, an increase of 6.2-fold in transcription activation of the G-free reporter gene was observed for (27950–28314) $C_2AT$  (Fig. 1B, lanes 5 and 6). Moreover, this activation was GCMa/1 dependent because deletion of the proximal GCMa/1-binding site (pGBS) in (27950–28314) $C_2AT$  completely abolished the activation (Fig. 1B, lanes 7 and 8). On the contrary, no significant transcription activation by hGCMa/1Flag was observed for (27950–28244) $C_2AT$  (Fig. 1B, lanes 1 and 2). We also tested the effect of the putative TATA box in (27950–28314) $C_2AT$  on the transcription activation directed by GCMa/1. As shown in Fig. 1B (lanes 9 and 10), deletion of the TATA box also completely abolished the GCMa/1-dependent transcription activation on (27950–28314) $C_2AT$ . Therefore, nt 28314 is a functional transcription initiation site for the *SYNCYTIN* gene and the TATA box is essential for transcription activation directed by GCMa/1.

## Discussion

GCMa/1 is a placenta-specific transcription factor required for placental development. To study the biochemical mechanism of GCMa/1-mediated gene expression, in this report, we established an *in vitro* transcription assay system for GCMa/1 using recombinant hGCMa/1 proteins, HeLa nuclear extracts, and G-free reporter constructs containing multiple copies of pGBS or the syncytin promoter region. The use of HeLa nuclear extracts provides two advantages in our assay system. First, HeLa cells are devoid of endogenous GCMa/1. Therefore, depletion of endogenous GCMa/1 is not necessary in this assay system. Second, HeLa nuclear extracts contain the basal transcription machinery, which is known to support transcription initiation by RNA polymerase II (Dignam et al. 1983). Therefore, the basal transcription machinery in HeLa nuclear extracts can be regulated by hGCMa/1Flag or MBP-hGCMa/1 to drive reporter expression via the pGBS site.

On establishing the *in vitro* transcription assay system for GCMa/1, we first demonstrated that recombinant hGCMa/1 in a form of either hGCMa/1Flag or MBP-hGCMa/1 was able to specifically bind pGBS. We further identified (pGBS) $_4$  TATA as a better reporter construct for a higher signal to noise ratio than (pGBS) $_4TK$ . Several lines of evidence supported that specific transcriptional activity of recombinant hGCMa/1 was detected in the *in vitro* transcription reactions. First, the mutant reporter construct (Mut) $_5TATA$  was barely responsive to MBP-hGCMa/1, and MBP was not able to activate the reporter gene expression on (pGBS) $_4TATA$ . Second,  $\alpha$ -amanitin, an RNA polymerase II inhibitor, completely blocked the transcription from (pGBS) $_4TATA$ . Third, dose-dependent responses of transcription activation of reporter gene expression were observed by increasing amounts of recombinant hGCMa/1 or the reporter template. To our knowledge, this

study is the first to demonstrating the transcriptional activity of GCMa/1 *in vitro*.

Recently, promoter analysis by transient expression experiments has revealed several ubiquitous factor-binding elements that are potentially involved in regulation of promoter activity of the *SYNCYTIN* gene (Lee et al. 2003; Cheng et al. 2004). Paradoxically, the study by Lee et al. (2003) showed that the putative TATA box in the 3'-LTR of HERV-W was required for its promoter activity, whereas in the study by Cheng et al. (2004), the putative TATA box in the 5'-LTR was dispensable. In addition, the transcription initiation site of the *SYNCYTIN* gene has been predicated to be downstream or upstream of the putative TATA box by 5'-RACE experiments (Blond et al. 1999) or mung bean nuclease protection analyses (Cheng et al. 2004). To reconcile these discrepancies, we tested the syncytin promoter constructs in *in vitro* transcription reactions in the present study. We demonstrated that recombinant hGCMa/1Flag proteins stimulated syncytin promoter activity via the proximal GCMa/1-binding site upstream of the 5'-LTR of HERV-W. In addition, the putative TATA box in the 5'-LTR is required for the transcription activation directed by hGCMa/1Flag. Previously, we have shown that ectopic expression of GCMa/1 by an adenoviral vector in placental cells is able to stimulate *SYNCYTIN* gene expression but not in nonplacental cells (Yu et al. 2002). This observation suggests that GCMa/1 is not sufficient for placenta-specific expression of the *SYNCYTIN* gene and may collaborate with other factors to regulate *SYNCYTIN* expression in placental cells. Identification of these factors is therefore necessary. The *in vitro* transcription assay established in this study may help to elucidate the roles of these factors in the transcriptional activity of GCMa/1 in the future.

The three-dimensional structure of the GCM domain – DNA complex has been recently elucidated by X-ray crystallography (Cohen et al. 2003). The GCM domain contains two zinc ions and can be divided into two subdomains. The large subdomain consists of an N-terminal extension, a five-stranded antiparallel  $\beta$ -sheet, and a short helix. The small subdomain contains a three-stranded  $\beta$ -pleated sheet and two helices. One of the two structural zinc ions tethers the two subdomains by coordination of amino acid residues in the two subdomains. GCMa/1 mutants carrying mutations of the zinc ion coordinated residues exhibit different DNA-binding activities. However, the transcriptional activity of these mutants was completely abolished. These observations suggested that interactions between the transactivation and DNA-binding domains of GCMa/1 are important for its biological function. In this study, MBP-hGCMa/1 was transcriptionally active in the transcription reaction, indicating that the MBP moiety does not interfere with the intramolecular interaction of hGCMa/1 in the fusion protein.

We noticed that the G-free reporter constructs used in this study are naked DNA templates lacking the chromatin structure found *in vivo*. However, since the DNA in a eukaryotic nucleus is packaged into a nucleosome array and the precise positioning of nucleosomes along the promoter region is critical for appropriate regulation of gene expression, a reconstituted chromatin template of natural syncytin promoter should be employed in future experiments to better mimic *in vivo* conditions.

## Acknowledgments

This work was supported by grants (to H.C.) from the National Science Council (grant 93-2311-B-001-065) and Academia Sinica of Taiwan. We also thank Dr. Yasuo Inoue for supporting this study.

## References

- Akiyama, Y., Hosoya, T., Poole, A.M., and Hotta, Y. 1996. The gcm-motif: a novel DNA-binding motif conserved in *Drosophila* and mammals. *Proc. Natl. Acad. Sci. U.S.A.* **93**: 14912–14916.
- Alfonso, T.B., and Jones, B.W. 2002. gcm2 promotes glial cell differentiation and is required with glial cells missing for macrophage development in *Drosophila*. *Dev. Biol.* **248**: 369–383.
- Anson-Cartwright, L., Dawson, K., Holmyard, D., Fisher, S.J., Lazzarini, R.A., and Cross, J.C. 2000. The glial cells missing-1 protein is essential for branching morphogenesis in the chorio-allantoic placenta. *Nat. Genet.* **25**: 311–314.
- Basyuk, E., Cross, J.C., Corbin, J., Nakayama, H., Hunter, P., Nait-Oumesmar, B., and Lazzarini, R.A. 1999. Murine Gcm1 gene is expressed in a subset of placental trophoblast cells. *Dev. Dyn.* **214**: 303–311.
- Blond, J.L., Beseme, F., Duret, L., Bouton, O., Bedin, F., Perron, H., Mandrand, B., and Mallet, F. 1999. Molecular characterization and placental expression of HERV-W, a new human endogenous retrovirus family. *J. Virol.* **73**: 1175–1185.
- Campbell, G., Goring, H., Lin, T., Spana, E., Andersson, S., Doe, C.Q., and Tomlinson, A. 1994. RK2, a glial-specific homeodomain protein required for embryonic nerve cord condensation and viability in *Drosophila*. *Development*, **120**: 2957–2966.
- Cheng, Y.H., Richardson, B.D., Hubert, M.A., and Handwerger, S. 2004. Isolation and characterization of the human syncytin gene promoter. *Biol. Reprod.* **70**: 694–701.
- Cohen, S.X., Moulin, M., Hashemhosseini, S., Kilian, K., Wegner, M., and Muller, C.W. 2003. Structure of the GCM domain – DNA complex: a DNA-binding domain with a novel fold and mode of target site recognition. *EMBO J.* **22**: 1835–1845.
- Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**: 1475–1489.
- Gorski, K., Carneiro, M., and Schibler, U. 1986. Tissue-specific in vitro transcription from the mouse albumin promoter. *Cell*, **47**: 767–776.
- Halter, D.A., Urban, J., Rickert, C., Ner, S.S., Ito, K., and Travers, A.A., Technau, G.M. 1995. The homeobox gene repo is required for the differentiation and maintenance of glia function in the embryonic nervous system of *Drosophila melanogaster*. *Development*, **121**: 317–332.
- Hosoya, T., Takizawa, K., Nitta, K., and Hotta, Y. 1995. glial cells missing: a binary switch between neuronal and glial determination in *Drosophila*. *Cell*, **82**: 1025–1036.
- Jones, B.W., Fetter, R.D., Tear, G., and Goodman, C.S. 1995. glial cells missing: a genetic switch that controls glial versus neuronal fate. *Cell*, **82**: 1013–1023.
- Kammerer, M., and Giangrande, A. 2001. Glide2, a second glial promoting factor in *Drosophila melanogaster*. *EMBO J.* **20**: 4664–4673.
- Kanemura, Y., Hiraga, S., Arita, N., Ohnishi, T., Izumoto, S., Mori, K. et al. 1999. Isolation and expression analysis of a novel human homologue of the *Drosophila* glial cells missing (gcm) gene. *FEBS Lett.* **442**: 151–156.
- Kim, J., Jones, B.W., Zock, C., Chen, Z., Wang, H., Goodman, C.S., and Anderson, D.J. 1998. Isolation and characterization of mammalian homologs of the *Drosophila* gene glial cells missing. *Proc. Natl. Acad. Sci. U.S.A.* **95**: 12364–12369.
- Klein-Hitpass, L., Tsai, S.Y., Weigel, N.L., Allan, G.F., Riley, D., Rodriguez, R. et al. 1990. The progesterone receptor stimulates cell-free transcription by enhancing the formation of a stable preinitiation complex. *Cell*, **60**: 247–257.
- Lee, W.J., Kwun, H.J., and Jang, K.L. 2003. Analysis of transcriptional regulatory sequences in the human endogenous retrovirus W long terminal repeat. *J. Gen. Virol.* **84**: 2229–2235.
- Mi, S., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L. et al. 2000. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature (London)*, **403**: 785–789.
- Nait-Oumesmar, B., Copperman, A.B., and Lazzarini, R.A. 2000. Placental expression and chromosomal localization of the human Gcm 1 gene. *J. Histochem. Cytochem.* **48**: 915–922.
- Sawadogo, M., and Roeder, R.G. 1985. Factors involved in specific transcription by human RNA polymerase II: analysis by a rapid and quantitative in vitro assay. *Proc. Natl. Acad. Sci. U.S.A.* **82**: 4394–4398.
- Schreiber, J., Sock, E., and Wegner, W. 1997. The regulator of early gliogenesis glial cells missing is a transcription factor with a novel type of DNA-binding domain. *Proc. Natl. Acad. Sci. U.S.A.* **94**: 4739–4744.
- Schreiber, J., Enderich, J., and Wegner, M. 1998. Structural requirements for DNA binding of GCM proteins. *Nucleic Acids Res.* **26**: 2337–2343.
- Schreiber, J., Riethmacher-Sonnenberg, E., Riethmacher, D., Tuerk, E.E., Enderich, J., Bosl, M.R., and Wegner, M. 2000. Placental failure in mice lacking the mammalian homolog of glial cells missing, GCMA. *Mol. Cell. Biol.* **20**: 2466–2477.
- Xiong, W.C., Okano, H., Patel, N.H., Blendy, J.A., and Montell, C. 1994. repo encodes a glial-specific homeo domain protein required in the *Drosophila* nervous system. *Genes Dev.* **8**: 981–994.
- Yu, C., Shen, K., Lin, M., Chen, P., Lin, C., Chang, G.D., and Chen, H. 2002. GCMA regulates the syncytin-mediated trophoblastic fusion. *J. Biol. Chem.* **277**: 50062–50068.