

Impacts of a new transcription factor family: mammalian GCM proteins in health and disease

Said Hashemolhosseini and Michael Wegner

Institut für Biochemie, Universität Erlangen-Nürnberg, 91054 Erlangen, Germany

GCM proteins constitute a small transcription factor family with a DNA-binding domain exhibiting a novel fold composed of two subdomains rigidly held together by coordination of one of two structural zinc cations. In all known cases, GCM proteins exert the role of master regulators: the prototypical family member determines gliogenesis in *Drosophila melanogaster*, whereas mammalian GCM proteins orchestrate divergent aspects of development and physiology in placenta, kidney, thymus, and parathyroid gland. Recent data point to an involvement of GCM proteins in different pathological contexts, such as preeclampsia, hyper- or hypoparathyroidism, and parathyroid gland tumors.

The story of *glial cells missing* (*gcm*) started with a screen for axonal pathfinding mutants in *Drosophila melanogaster* (Hosoya et al., 1995; Jones et al., 1995; Vincent et al., 1996). The identified prototypic GCM (also called Glide; glial cell deficient) turned out to be the earliest binary switch known to be necessary for glial cell development in the fruit fly. Surplus of GCM/Glide instructs neural precursor cells to differentiate into glial cells, whereas in the absence of GCM/Glide neural precursors predominantly mature to neurons. For details of GCM function in *Drosophila*, we refer the reader to several recent reviews (Jones, 2001; Van De Bor and Giangrande, 2002). GCM proteins encode transcription factors with a new type of zinc-containing DNA-binding domain, a so-called GCM domain, that recognizes the DNA sequence 5'-ATGCGGGT-3' or variations thereof (Akiyama et al., 1996; Schreiber et al., 1997, 1998; Cohen et al., 2002). The GCM domain represents the structural hallmark and the only region of significant homology of the different GCM family members. It is assembled from a large and a small subdomain, both rich in β -pleated sheets. One of the two zinc cations tethers the two subdomains together, while the other is located at the periphery of the small subdomain. A five-stranded β -sheet from the large domain protrudes into the major groove perpendicularly to the DNA

axis and thereby defines a novel use of β -sheets in DNA recognition (Cohen et al., 2003).

All GCM proteins known to date are listed in Fig. 1. In *Drosophila*, a second *gcm* gene was identified (Kammerer and Giangrande, 2001; Alfonso and Jones, 2002). This *gcm2/glide2* gene is the result of a recent gene duplication event and assists *gcm/glide* in its function during gliogenesis. In the sea urchin *Strongylocentrotus purpuratus*, SpGcm was identified as a Notch-dependent factor (Calestani et al., 2003). Interestingly, Notch signaling also accounts for *gcm/glide* activation in *Drosophila's* nervous system, and thus might be a common denominator in *gcm* regulation. A GCM protein has also been identified and characterized in *Gallus gallus* (Hashemolhosseini et al., 2004). Current sequencing efforts have revealed additional GCM proteins in the genomes of *Xenopus laevis*, *Fugu rubripes*, and *Danio rerio* (unpublished data).

GCM in mammals

In search for regulators of mammalian gliogenesis, two GCM proteins, named GCMa/Gcm1 and GCMb/Gcm2, were identified (Akiyama et al., 1996; Kim et al., 1998; Schreiber et al., 1998). No other functional *gcm* genes have been detected in the human and mouse genome. Initially, overexpression studies supported a function of mammalian GCM proteins in nervous system development. Thus, GCMa/Gcm1 disrupted neurogenesis or induced gliogenesis when ectopically expressed in the nervous system of *Drosophila* and mammals (Kim et al., 1998; Reifegerste et al., 1999; Nait-Oumesmar et al., 2002; Iwasaki et al., 2003). However, all studies reported to date failed to detect significant amounts of GCMa/Gcm1 or GCMb/Gcm2 in the mammalian nervous system. If at all detected in embryonic brain, expression rates were found orders of magnitude lower than in other tissues. In GCMa/Gcm1 knockout mice that express a β -galactosidase marker instead of GCMa/Gcm1, no β -galactosidase-positive cells were detected in the brain (Hashemolhosseini et al., 2002).

Expression of mammalian GCM proteins was highest in tissues other than the nervous system. During mouse development from embryonic day (E) 7.5 to E17.5 GCMa/

Address correspondence to Said Hashemolhosseini, Institut für Biochemie, Universität Erlangen-Nürnberg, Fahrstrasse 17, 91054 Erlangen, Germany. Tel.: 49 9131 85 24634. Fax: 49 9131 85 22484. email: sh@biochem.uni-erlangen.de

Key words: glial cells missing; glide; parathyroid; kidney; placenta

Abbreviations used in this paper: E, embryonic day; *gcm*, *glial cells missing*.

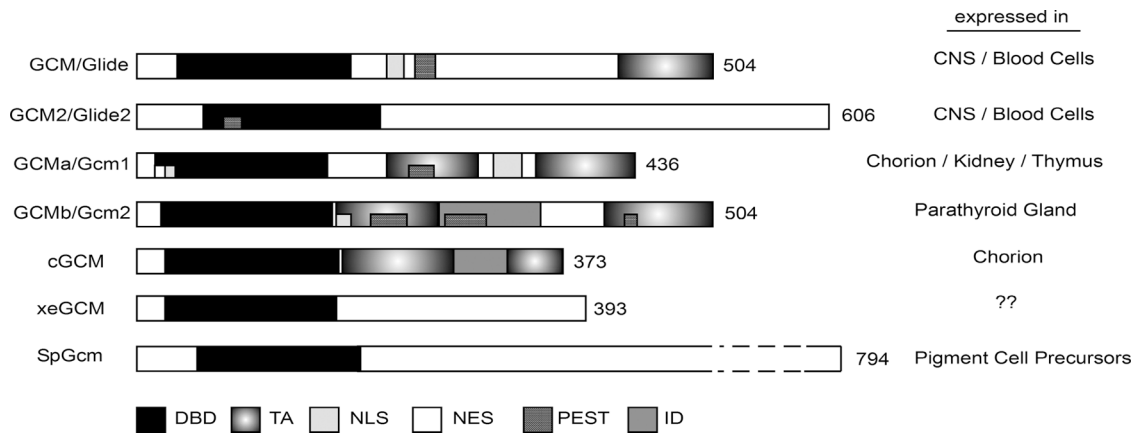


Figure 1. **Overview and domain topology of GCM proteins.** Numbers to the right indicate amino acid residues in each GCM protein. To avoid superimposing multiple domains in the same GCM protein, one of the two is shown at half height. At the right, the expression site of each GCM protein is listed. Note the presence of two transactivation domains for mammalian and chick GCM proteins. For GCM2/Glide2, xeGCM, and SpGcm, the GCM domain is the only characterized domain. At the bottom, differently patterned squares corresponding to different functional domains are shown (DBD, DNA-binding domain; TA, transactivation domain; NLS, nuclear localization signal; NES, nuclear export signal; PEST, proline-glutamine-serine-threonine rich motif; ID, inhibitory domain).

Gcm1 is expressed in the fetal part of the placenta and defines a unique subset of trophoblast cells (Basyuk et al., 1999). Knockout mouse models demonstrate embryonic lethality in the absence of GCMA/Gcm1 around E10 (Anson-Cartwright et al., 2000; Schreiber et al., 2000). GCMA/Gcm1 plays a pivotal role in placental labyrinth development where it is required for placental branching morphogenesis. After *GCMA/Gcm1* deletion, the allantois is still able to grow toward and contact the chorion, but subsequent steps of placental labyrinth formation are affected so that the failure to supply the embryo with sufficient amounts of nutrients and oxygen leads to embryonic death. Interestingly, the GCM protein identified from chicken is also found in the egg's chorion (Hashemolhosseini et al., 2004). Thus, its localization is reminiscent of the spatial expression profile of GCMA/Gcm1. Apart from the fetal placenta, GCMA/Gcm1 expression is selectively turned on at late embryonic stages in the thymus and perinatally in the kidney (Hashemolhosseini et al., 2002). In both of these organs, GCMA/Gcm1 is expressed throughout adulthood. In the thymic cortex a few clusters of ~30 GCMA/Gcm1-positive cells appear first at E16.5. During development and adult stages, the number of clusters increases with a concomitant decrease of the number of cells to two to three cells per cluster. In the kidney of mice, GCMA/Gcm1 expression was detected specifically in cells of the S3 segment of proximal tubules. Because renal GCMA/Gcm1 expression occurs after kidney development is almost completed, GCMA/Gcm1 is likely involved in kidney physiology.

Although not much is known about mechanistic aspects of GCMA/Gcm1 function in the kidney, several aspects of its regulation and function have been studied in the placenta. Integrin- $\alpha 4$ was identified by antibody-blocking experiments as an upstream component of the GCMA-signaling cascade in placental trophoblasts (Stecca et al., 2002). Once induced, GCM proteins must be able to efficiently translocate to the nucleus to perform their function as transcription factors. For that matter, all GCM proteins possess

nuclear localization signals. For some GCM proteins like GCMb/Gcm2 or the prototypical GCM/Glide, these conform to the classical bipartite nuclear localization motif. In others like GCMA/Gcm1 or chick GCM, nuclear localization is mediated by regions without resemblance to known nuclear localization signals (Hashemolhosseini et al., 2003). Nuclear import was counteracted in GCMA/Gcm1 by an amino terminal nuclear export activity. In fact, several GCM proteins including GCMA/Gcm1 have also been found in the cytosol (Bernardoni et al., 1999; Nait-Oumesmar et al., 2000). Thus, nucleocytoplasmic shuttling might be a conserved feature within this family of transcription factors and might offer control of their activity either through sequestration or cytoplasmic posttranslational modification.

The transactivation capacity of all mammalian GCM proteins can be attributed to two transactivation domains compared with one in the prototypical *Drosophila* GCM/Glide protein (Schreiber et al., 1998). Furthermore, GCM proteins contain PEST motifs within their primary structure that convey short half-lives and low protein stability (Tuerk et al., 2000) and influence target gene activation. For GCMA/Gcm1, two placental target genes are known (Yamada et al., 1999; Yu et al., 2002). These genes are human aromatase and human syncytin. In the case of the aromatase gene, GCMA/Gcm1 binds to a minimal enhancer, which is essential for placenta-specific transcription. In the case of syncytin, GCMA/Gcm1 induces the promoter activity of the long-terminal repeats of a syncytin-harboring retrovirus integrated into the human genome and specifically drives endogenous syncytin expression in human trophoblast lines. Syncytin in humans can mediate fusion of cytotrophoblasts into the syncytiotrophoblast layer, and syncytin expression is tightly regulated to maintain a functional syncytiotrophoblast layer. However, mice do not possess a retrovirally encoded syncytin, so that GCMA/Gcm1-dependent syncytin expression cannot explain syncytiotrophoblast formation in mice. This finding suggests additional placental GCMA/Gcm1-inducible factors for labyrinth formation in the mouse.

GCMa/Gcm1 is likely to activate its target genes in combination with interacting transcription factors. Recently, Pitx2 has been identified as the first protein interacting with GCMa/Gcm1 in placenta and kidney (unpublished data).

The expression of the second mammalian *GCM* gene, GCMb/Gcm2, is restricted to parathyroid glands (Kim et al., 1998). 30% of *GCMb/Gcm2*-deficient mice die shortly after birth owing to severe hypocalcaemia. The remaining were viable, fertile, and had milder hypocalcaemia associated with hyperphosphataemia and increased calcium elimination in the urine without evidence of renal participation, altogether features characteristic for hypoparathyroidism (Günther et al., 2000). In agreement, *GCMb/Gcm2*-deficient mice lack parathyroid glands but show normal parathyroid hormone serum levels. In the same study, GCMa/Gcm1-mediated parathyroid hormone production in the thymus was uncovered as an auxiliary regulatory mechanism for calcium homeostasis in the absence of parathyroid glands.

In humans and mice, parathyroid glands as well as the thymus arise from the pharyngeal pouches. GCMb/Gcm2 and Foxn1 are expressed in a complementary fashion in the common primordium, thereby dividing it into separate domains that prefigure parathyroid and thymus (Gordon et al., 2001). Transcription factors that determine cell fate in the common primordium and subsequent development of thymus and parathyroid glands have recently been identified. GCMb/Gcm2 expression is reduced in Pax1^{-/-} mutants by E11.5, and further reduced or absent in Hoxa3^{+/-}/Pax1^{-/-} double mutants, indicating that GCMb/Gcm2 is under control of these transcription factors (Su et al., 2001).

GCM in disease

GCM proteins so far turned out to exert crucial roles as master regulators in different developmental contexts. Therefore, it would be surprising if they had no significance in pathological circumstances. Concerning GCMa/Gcm1's action in the placenta, GCMa/Gcm1 is the first transcription factor capable of initiating syncytiotrophoblast formation. Accordingly, GCMa/Gcm1 might play a role in placental diseases such as preeclampsia and intrauterine growth restriction. These conditions complicate 5–10% of human pregnancies and result in fetal to neonatal mortality and morbidity (Anson-Cartwright et al., 2000). Indeed, decreased placental GCMa/Gcm1 expression has been observed in preeclampsia (Chen et al., 2004). GCMa/Gcm1 expression was reduced in a statistically significant number of preeclamptic placentae compared with gestational age-matched controls. In this work, changes in the expression of syncytin were not analyzed. One would assume that transcript levels of this GCMa/Gcm1 target gene are also diminished.

Because *GCMb/Gcm2* was identified as a master regulatory gene of parathyroid gland development, alterations of its levels or activity might contribute to pathological states of this organ. Homozygosity for a large intragenic deletion within the *GCMb/Gcm2* gene was identified as the cause for autosomal recessive hypoparathyroidism in a 5-wk-old patient suffering from generalized seizures (Ding et al., 2001). Recently, another case of familial hypoparathyroidism has been observed. In this case, a single amino acid change in the GCM domain of GCMb/Gcm2 that results in a transcrip-

tionally significantly less active GCMb/Gcm2 was identified as the underlying cause of the disease (unpublished data). Thus, loss of GCMb/Gcm2 function might be a more common cause of hypoparathyroidism. Unexpectedly and contrary to mice, embryonic and adult thymus of healthy humans were assayed negative for both parathyroid hormone and GCMa/Gcm1 (Maret et al., 2004).

Recent data also implicate GCMb/Gcm2 in tumorigenesis. These tumors are either located in the thymus or in the parathyroid gland. Intrathymic adenomas have been reported to secrete parathyroid hormone and produce GCMb/Gcm2, suggesting that these tumors originate from misplaced parathyroid cells (Maret et al., 2004). GCMb/Gcm2 might be a diagnostic marker for these intrathymic adenomas. The amount of GCMb/Gcm2 in cells of these intrathymic adenomas was not investigated. Human GCMb/Gcm2 was also detected in several sporadic adenomas of the parathyroid gland (Kanemura et al., 1999). In these parathyroid hormone-producing adenomas, GCMb/Gcm2 was present, but expressed at significantly lower levels than in surrounding normal parathyroid tissue arguing that there is no strict correlation between GCMb/Gcm2 and parathyroid hormone expression (Correa et al., 2002). It is tempting to speculate that constant high levels of GCMb/Gcm2 expression are incompatible with proliferation and de-differentiation in an adenoma.

Concluding remarks

Both mammalian GCM proteins play crucial roles at different steps of development. GCMa/Gcm1 may have additional functions in renal physiology. It is a prime candidate for a transcription factor that regulates expression of different renal transporter genes. Due to the early embryonic lethality of a constitutive *GCMa/Gcm1* knockout mouse, this has to be analyzed by generation and characterization of a conditional knockout mouse. Apart from their normal function, mammalian GCM proteins have recently been implicated in different human diseases such as preeclampsia, hyper- or hypoparathyroidism, and tumorigenesis. Further knowledge of the molecular action of mammalian GCM proteins through identification of more complete sets of target genes and binding partners might clarify mechanistic details of their contribution to disease.

S. Hashemolhosseini is supported by the Deutsche Forschungsgemeinschaft (grant SFB 473).

Submitted: 16 June 2004

Accepted: 3 August 2004

References

- Akiyama, Y., T. Hosoya, A.M. Poole, and Y. Hotta. 1996. The gcm-motif: a novel DNA-binding motif conserved in *Drosophila* and mammals. *Proc. Natl. Acad. Sci. USA*. 93:14912–14916.
- Alfonso, T.B., and B.W. Jones. 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., K. Dawson, D. Holmyard, S.J. Fisher, R.A. Lazzarini, and J.C. Cross. 2000. The glial cells missing-1 protein is essential for branching morphogenesis in the chorioallantoic placenta. *Nat. Genet.* 25:311–314.
- Basyuk, E., J.C. Cross, J. Corbin, H. Nakayama, P. Hunter, B. Nait-Oumesmar, and R.A. Lazzarini. 1999. Murine Gcm1 gene is expressed in a subset of pla-

- central trophoblast cells. *Dev. Dyn.* 214:303–311.
- Bernardoni, R., M. Kammerer, J.-L. Vonesch, and A. Giangrande. 1999. Gliogenesis depends on glide/gcm through asymmetric division of neuroglioblasts. *Dev. Biol.* 216:265–275.
- Calestani, C., J.P. Rast, and E.H. Davidson. 2003. Isolation of pigment cell specific genes in the sea urchin embryo by differential macroarray screening. *Development.* 130:4587–4596.
- Chen, C.P., C.Y. Chen, Y.C. Yang, T.H. Su, and H. Chen. 2004. Decreased placental GCM1 (glial cells missing) gene expression in pre-eclampsia. *Placenta.* 25:413–421.
- Cohen, S.X., M. Moulin, O. Schilling, W. Meyer-Klaucke, J. Schreiber, M. Wegner, and C.W. Müller. 2002. The GCM domain is a Zn-coordinating DNA-binding domain. *FEBS Lett.* 528:95–100.
- Cohen, S.X., M. Moulin, S. Hashemolhosseini, K. Kilian, M. Wegner, and C.W. Müller. 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.
- Correa, P., G. Akerstrom, and G. Westin. 2002. Underexpression of Gcm2, a master regulatory gene of parathyroid gland development, in adenomas of primary hyperparathyroidism. *Clin. Endocrinol. (Oxf).* 57:501–505.
- Ding, C., B. Buckingham, and M.A. Levine. 2001. Familial isolated hypoparathyroidism caused by a mutation in the gene for the transcription factor GCMB. *J. Clin. Invest.* 108:1215–1220.
- Gordon, J., A.R. Bennett, C.C. Blackburn, and N.R. Manley. 2001. Gcm2 and Foxn1 mark early parathyroid- and thymus-specific domains in the developing third pharyngeal pouch. *Mech. Dev.* 103:141–143.
- Günther, T., Z.-F. Chen, J. Kim, M. Priemel, J.M. Rueger, M. Amling, J.M. Moseley, T.J. Martin, D.J. Anderson, and G. Karsenty. 2000. Genetic ablation of parathyroid glands reveals another source of parathyroid hormone. *Nature.* 406:199–203.
- Hashemolhosseini, S., M. Hadjihannas, C.C. Stolt, C.S. Haas, K. Amann, and M. Wegner. 2002. Restricted expression of mouse GCMA/Gcm1 in kidney and thymus. *Mech. Dev.* 118:175–178.
- Hashemolhosseini, S., K. Kilian, E. Kardash, P. Lischka, T. Stamminger, and M. Wegner. 2003. Structural requirements for nuclear localization of GCMA/Gcm-1. *FEBS Lett.* 553:315–320.
- Hashemolhosseini, S., K. Schmidt, K. Kilian, E. Rodriguez, and M. Wegner. 2004. Conservation and variation of structure and function in a newly identified GCM homolog from chicken. *J. Mol. Biol.* 336:441–451.
- Hosoya, T., K. Takizawa, K. Nitta, and Y. Hotta. 1995. glial cells missing: a binary switch between neuronal and glial determination in *Drosophila*. *Cell.* 82:1025–1036.
- Iwasaki, Y., T. Hosoya, H. Takebayashi, Y. Ogawa, Y. Hotta, and K. Ikenaka. 2003. The potential to induce glial differentiation is conserved between *Drosophila* and mammalian glial cells missing genes. *Development.* 130:6027–6035.
- Jones, B.W. 2001. Glial cell development in the *Drosophila* embryo. *Bioessays.* 23:877–887.
- Jones, B.W., R.D. Fetter, G. Tear, and C.S. Goodman. 1995. glial cells missing: a genetic switch that controls glial versus neuronal fate. *Cell.* 82:1013–1023.
- Kammerer, M., and A. Giangrande. 2001. Glide2, a second glial promoting factor in *Drosophila melanogaster*. *EMBO J.* 20:4664–4673.
- Kanemura, Y., S. Hiraga, N. Arita, T. Ohnishi, S. Izumoto, K. Mori, H. Matsumura, M. Yamasaki, S. Fushiki, and T. Yoshimine. 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., B.W. Jones, C. Zock, Z. Chen, H. Wang, C.S. Goodman, and D.J. Anderson. 1998. Isolation and characterization of mammalian homologs of the *Drosophila* gene glial cells missing. *Proc. Natl. Acad. Sci. USA.* 95:12364–12369.
- Maret, A., I. Bourdeau, C. Ding, S.S. Kadkol, W.H. Westra, and M.A. Levine. 2004. Expression of GCMB by intrathyroid parathyroid hormone-secreting adenomas indicates their parathyroid cell origin. *J. Clin. Endocrinol. Metab.* 89:8–12.
- Nait-Oumesmar, B., A.B. Copperman, and R.A. Lazzarini. 2000. Placental expression and chromosomal localization of the human Gcm 1 gene. *J. Histochem. Cytochem.* 48:915–922.
- Nait-Oumesmar, B., B. Stecca, G. Fatterpekar, T. Naidich, J. Corbin, and R.A. Lazzarini. 2002. Ectopic expression of Gcm1 induces congenital spinal cord abnormalities. *Development.* 129:3957–3964.
- Reifegerste, R., J. Schreiber, S. Gulland, A. Ludemann, and M. Wegner. 1999. mGCMa is a murine transcription factor that overrides cell fate decisions in *Drosophila*. *Mech. Dev.* 82:141–150.
- Schreiber, J., E. Sock, and M. Wegner. 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. USA.* 94:4739–4744.
- Schreiber, J., J. Enderich, and M. Wegner. 1998. Structural requirements for DNA binding of GCM proteins. *Nucleic Acids Res.* 26:2337–2343.
- Schreiber, J., E. Riethmacher-Sonnenberg, D. Riethmacher, E.E. Tuerk, J. Enderich, M. Bösl, and M. Wegner. 2000. Placental failure in mice lacking the mammalian homolog of glial cells missing, GCMA. *Mol. Cell. Biol.* 20:2466–2474.
- Stecca, B., B. Nait-Oumesmar, K.A. Kelley, A.K. Voss, T. Thomas, and R.A. Lazzarini. 2002. Gcm1 expression defines three stages of chorio-allantoic interaction during placental development. *Mech. Dev.* 115:27–34.
- Su, D., S. Ellis, A. Napier, K. Lee, and N.R. Manley. 2001. Hoxa3 and pax1 regulate epithelial cell death and proliferation during thymus and parathyroid organogenesis. *Dev. Biol.* 236:316–329.
- Tuerk, E.E., J. Schreiber, and M. Wegner. 2000. Protein stability and domain topology determine the transcriptional activity of the mammalian glial cells missing homolog, GCMb. *J. Biol. Chem.* 275:4774–4782.
- Van De Bor, V., and A. Giangrande. 2002. glide/gcm: at the crossroads between neurons and glia. *Curr. Opin. Genet. Dev.* 12:465–472.
- Vincent, S., J.-L. Vonesch, and A. Giangrande. 1996. Glide directs glial fate commitment and cell fate switch between neurones and glia. *Development.* 122:131–139.
- Yamada, K., H. Ogawa, S. Honda, N. Harada, and T. Okazaki. 1999. A GCM motif protein is involved in placenta-specific expression of human aromatase gene. *J. Biol. Chem.* 274:32279–32286.
- Yu, C., K. Shen, M. Lin, P. Chen, C. Lin, G.D. Chang, and H. Chen. 2002. GCMA regulates the syncytin-mediated trophoblastic fusion. *J. Biol. Chem.* 277:50062–50068.