FEBS Letters 536 (2003) 12–18 FEBS 26931

Promoter analysis of a medaka fish intestinal guanylyl cyclase gene

Mina Nakauchi, Norio Suzuki*

Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan

Received 23 October 2002; revised 11 December 2002; accepted 11 December 2002

First published online 14 January 2003

Edited by Ned Mantei

Abstract **We characterized the promoter activity of a medaka fish intestinal guanylyl cyclase gene, *OlGC6*, by assay of enzyme activity in response to various promoter–luciferase fusion gene constructs introduced into CACO-2 cells and medaka fish embryos. A transient transfection assay of the various fusion gene constructs showed that the nucleotides between −98 and −89 in the 5′-flanking region of the *OlGC6* gene are essential for transcription of the *OlGC6* gene in CACO-2, and that the *OlGC6* gene fragment between −98 and +50 is sufficient to drive gene expression in the medaka fish intestine. An electrophoretic mobility shift assay and ultraviolet (UV) cross-linking experiments demonstrated that a nuclear protein from CACO-2 cells and the adult medaka fish intestinal cells binds specifically to the AGACCTTTGC nucleotides in the regulatory element.**

Key words: Guanylyl cyclase; Intestine; Medaka fish; *cis*-Regulatory element; CACO-2 cell

1. Introduction

The intestine of vertebrates expresses a specific gene for membrane guanylyl cyclase (membrane GC). In mammals, such a membrane GC gene, *GC-C*, is expressed at high levels throughout the epithelium of the intestine, from the duodenum to the colon, and the gene products are detected abundantly in both the crypts and villi of the small intestine, as well as in the crypts and surface epithelium of the colon [1–3]. Mammalian GC-C has been shown to be the receptor for *Escherichia coli* heat-stable enterotoxin (STa) [4], and for the endogenous peptides guanylin and uroguanylin [5,6]. In response to the binding of the ligand to the extracellular domain of GC-C, the intracellular cyclase catalytic domain converts GTP to cGMP, subsequently activating cGMP-dependent protein kinase II (PKG-II). It has been reported that PKG-II phosphorylates the cystic fibrosis transmembrane conductance regulator, increases chloride secretion, and inhibits electroneutral sodium absorption in the gastrointestinal tract [7]. The pathological consequence in the case of STa is severe secretory diarrhea [8]. Targeted disruption of the *GC-C* gene induces no deleterious effects in mice and, in fact, confers resistance to STa-induced diarrhea [9,10]. However, all vertebrates contain an intestine-specific membrane *GC* gene, implying that it serves some important and undefined physiological roles.

In a previous study, we isolated a cDNA clone for a membrane GC, *OlGC6*, from a medaka fish intestine cDNA library, and determined the complete nucleotide sequences of the cDNA and genomic DNA [11]. The 5′-flanking region of the *OlGC6* gene was found to be not well-conserved with that of the mammalian *GC-C* gene, the intestinal expression of which requires hepatocyte nuclear factor 4α (HNF-4α) or Cdx2 [12–14]. This finding suggests that the transcriptional regulatory mechanisms of the *OlGC6* and *GC-C* genes differ. In this paper, we report that the nucleotides between −98 and −89 in the 5′-flanking regions are important to transcription of the *OlGC6* gene in CACO-2 cells, and that the *OlGC6* gene fragment between −98 and +50 is sufficient to drive intestinal gene expression in the adult medaka fish. We also report that a specific nuclear protein from CACO-2 cells or from the medaka fish intestines binds to the AGACCTTTGC nucleotides in the regulatory element.

2. Materials and methods

2.1. Southern blot hybridization

A genomic Southern hybridization experiment was carried out using genomic DNA from a male individual of *Oryzias latipes* strain Hd-rR and a *OlGC6* cDNA fragment (nucleotides +184 to +225). Extraction of genomic DNA, blotting, labeling of the probe, and hybridization were performed as described previously [15]. Blots were washed three times with 2×SSC, 0.1% sodium dodecyl sulfate (SDS), at 50°C for 15–20 min. Imaging of the radioactive signals was performed using a FUJIX Bio-Imaging Analyzer BAS2000 (Fuji Photo Film, Japan).

2.2. Primer extension analysis

Total RNA was prepared from the intestines of the adult medaka fishes (orange-red variety of *O. latipes*) by the acid guanidium thiocyanate/phenol/chloroform extraction method [16]. Poly(A)$^+$ RNA was isolated using Oligotex-dT30⟨Super⟩ (Roche Molecular Biochemicals, Switzerland), according to the manufacturer's protocol. A primer extension experiment with 5 μg of poly(A)$^+$ RNA was carried out essentially by means of a previously described procedure [15].

2.3. Construction of the fusion genes

Genomic DNA fragments (nucleotides −1757 to +50, −1001 to +50, −491 to +50, −299 to +50, −201 to +50, −98 to +50, −88 to +50, −47 to +50, −18 to +50) were amplified using polymerase chain reaction (PCR). The PCR was conducted in a 50-μl reaction mixture containing 100 pg of template plasmid clone [11] (the 5′-flanking region of the *OlGC6* gene), 1×Ex Taq buffer, 0.2 mM dNTP mixture, 10 pmol paired primers, and 2.5 units Ex Taq (TaKaRa, Japan). The reaction was performed under the following conditions: 1 min at 94°C and 30 cycles of 30 s at 96°C, 30 s at 60°C, and 3 min at 72°C. The downstream primer used was 5′-TCAGAAAGCTTCTCCCTTTGTGCTCTGC-3′ and the upstream primers were 5′-AAGGGGGTACCATCTTCATACGGCGTC-3′, 5′-CTGGAGGGTACCCCCCCTCAG-3′, 5′-GGCCACGGTACCTTTCTCTCCTGG-3′, 5′-TGTGAGGGTACCCTGCTGTTCAG-3′, 5′-CCACATGGTACCTGAAGCTCTTATC-3′, 5′-GCAGCCGGTACCGCACACACG-3′, 5′-ACA-

*Corresponding author. Fax: (81)-11-706 4461.
E-mail address: norio-s@sci.hokudai.ac.jp (N. Suzuki).

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)00002-4

GCCGGTACCTGTGTCTTTAAAGG-3′, and 5′-ATTGGTACCC-TGACAAGGAAAGGACAACTC-3′ (GenBank™ accession number AB007192). For preparation of the mutant construct, the first PCR was performed as described by Ho et al. [17] using two sets of primers: 5′-AAGGGGGTACCATCTTCATACGGCGTC-3′, 5′-GTGC-AAAGGTAGTTGTGTGCGCTG-3′ for the upstream fragment; 5′-GCACACAACTACCTTTGCACACCC-3′, 5′-TCAGAAAGCTTC-TCCCTTTGTGCTCTGC-3′ for the downstream fragment. The second PCR was performed employing the first PCR fragment as a template using the following primers: 5′-AAGGGGGTACCATCTT-CATACGGCGTC-3′ and 5′-TCAGAAAGCTTCTCCCTTTGTGC-TCTGC-3′. The PCR product was digested with *Hin*dIII and *Kpn*I and ligated into the *Hin*dIII–*Kpn*I sites in a pGL3-enhancer luciferase vector (Promega, Madison, WI, USA). Plasmid was purified using the Qiagen Lambda Midi Kit (Qiagen, Germany). The correctness of the nucleotide sequence of each plasmid construct was examined by restriction enzyme digestion and partial sequencing.

2.4. Cell culture and DNA transfections

CACO-2 cells and COS1 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum (FBS) (HyClone®, Logan, UT, USA) and 1×penicillin–streptomycin–glutamine (Gibco™ Invitrogen Corporation, Netherlands) under a humidified 5% CO_2 atmosphere. For the reporter enzyme assay, 10^6 CACO-2 cells were plated per well in a six-well plate, or 10^6 COS1 cells were plated in a 100×20 mm dish and cultured for 24 h before transfection. A plasmid construct consisting of the *OlGC6* promoter region and luciferase gene was cotransfected with pSV-β-galactosidase (20 μg) (Promega) into CACO-2 cells or COS1 cells by the calcium phosphate coprecipitation method. After a 48-h culture of CACO-2 cells or a 24-h culture of COS1 cells, cells were lysed in Reporter Lysis Buffer (Promega), and the luciferase activity was assayed using a Luciferase Assay System (Promega) according to the manufacturer's protocol. β-Galactosidase activity was also assayed in order to normalize it in regard to variations in transfection efficiency.

2.5. Production and identification of transgenic medaka fish

The plasmid constructs 2kbpGC6-Luc, 100bpGC6-Luc, 50bpGC6-Luc, and 20bpGC6-Luc containing −1757 to +50, −98 to +50, −47 to +50, and −18 to +50 of the *OlGC6* gene fragments, respectively, upstream of the luciferase gene (see Section 2.3) were microinjected into a fertilized medaka fish embryo as described previously [18] and allowed to develop to term. Fishes were kept in indoor tanks under artificial reproductive conditions (10 h dark, 14 h light, 27°C) and fed on OtohimeB2 flakes (Nisshinseihun/Nisshinshiryou, Japan). A medaka fish carrying a transgene was identified for the presence of the luciferase gene by PCR of genomic DNA from the caudal fin using the following specific primers (GenBank™ accession number U47295): the upstream primer 5′-GCTGGAGAGCAACTGCATAA-3′ and the downstream primer 5′-CCTTCTTGGCCTTTATGAGG-3′. A F_0 medaka fish possessing a transgene was mated with a non-transgenic littermate to establish a line from each founder. Offspring was analyzed by both PCR and Southern blot hybridization for transmission of the transgene.

2.6. RNA analysis by reverse transcription (RT)-PCR

Total RNA was prepared from 20 medaka fish embryos (stages 35–39) and several adult medaka fish tissues by the acid guanidium thiocyanate/phenol/chloroform extraction method [16]. The first-strand cDNA was synthesized from total RNA of embryos (2.0 μg) or total RNA of each tissue (500 ng) using Super Script™ First-Strand Synthesis System (Invitrogen™ Life Technologies) according to the protocol of the manufacturer. PCRs were then carried out with specific primers for the luciferase gene (GenBank™ accession number U47295), the *OlGC6* gene (GenBank™ accession number AB007192), or the *OlCA1* gene (the cytoplasmic actin gene of the medaka fish, GenBank™ accession number D89627): *luc*-F, 5′-GCT-GGAGAGCAACTGCATAA-3′ and *luc*-R, 5′-CCTTCTTGGCCTT-TATGAGG-3′ (1570-bp product); *OlGC6*-F, 5′-ACACTGATCCA-GAGCTGACA-3′, and *OlGC6*-R, 5′-CCTTCTTCAGCATTGATA-GC-3′ (226-bp product); *OlCA1*-F, 5′-GGGTCTTCA-TGACGGGC-3′, and *OlCA1*-R, 5′-CAAGTCGGAACACATGTGCA-3′ (100-bp product) and 2 μl of the reverse-transcribed cDNA solution. The amplification conditions were as follows: 96°C, 30 s; 56°C, 30 s; and 72°C, 2.5 min performed 35 cycles. PCR products were analyzed by electrophoresis on a 1.5% agarose gel for the *OlGC6* gene, the *OlCA1* gene and on a 0.7% agarose gel for the luciferase gene.

2.7. In situ hybridization

A digoxigenin (DIG)-labeled probe was prepared by transcription from pBluescript II (KS+) containing the *OlGC6* cDNA fragment (nucleotides +1 to +314) at *Not*I and *Xho*I sites, and the luciferase gene (from pGL3 vector) at *Hin*dIII and *Xba*I sites using a DIG-RNA Labeling kit (Roche Molecular Biochemicals) according to the manufacturer's instructions.

The frozen tissue was sliced into a 10-μm section on a cryostat microtome (Leica, CM1900). The section was mounted on a slide coated with silane and dried. The section was fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS) and washed with PBS twice, then incubated in a solution containing 10 μg/ml proteinase K (Roche Molecular Biochemicals) for 4 min at room temperature. After washing twice with PBS for 5 min, the section was acetylated with 0.25% acetic anhydride in 1.5% triethanolamine for 10 min at room temperature. After washing with PBS for 5 min, the section was dehydrated through ethanol serial washes and then dried.

The section was prehybridized in a solution consisting of 50% formamide, 600 mM NaCl, 1×Denhardt's solution, 0.25% SDS, 1 mM ethylenediamine tetraacetic acid (EDTA), 10% dextran sulfate, 200 μg/ml yeast tRNA, and 10 mM Tris–HCl, pH 7.6 at 55°C for 1 h. After prehybridization, the section was transferred to fresh hybridization buffer containing 40 ng/ml DIG-labeled RNA probe. Hybridization was continued for at least 16 h at 55°C. The hybridized section was rinsed in 5×SSC at 55°C and then washed in a solution containing 5×SSC and 50% formamide at 55°C. The section was then treated with a solution containing 1 mM EDTA, 0.5 M NaCl, and 10 mM Tris–HCl, pH 7.5 with RNase A (Roche Molecular Biochemicals) for 30 min at 37°C, followed by washing in 2×SSC at 55°C for 20 min and in 0.2×SSC at 55°C for 20 min. After this step, the section was washed for 5 min at room temperature with DIG 1 buffer (150 mM NaCl and 100 mM Tris–HCl, pH 7.5) and then incubated in 0.5% bovine serum albumin (BSA) and 0.1% Tween 20 in DIG 1 buffer for 1 h at room temperature, followed by incubation in 1:1000 diluted anti-digoxigenin Fab fragment conjugated with alkaline phosphatase (DIG Nucleic Acid Detection kit), in DIG 1 buffer containing 0.1% Tween 20 for 1 h at room temperature. Unbound antibody conjugate was removed by washing twice with DIG 1 buffer containing 0.1% Tween 20 for 15 min each. Finally, the section was preincubated in DIG 3 buffer (100 mM NaCl, 50 mM $MgCl_2$, and 100 mM Tris–HCl, pH 9.5) and then incubated for 3 days in the dark in the same buffer containing the substrates, nitroblue tetrazolium (NBT, 450 μg/ml), and 5-bromo-4-chloro-3 indolyl phosphate (BCIP, 175 μg/ml) (Roche Molecular Biochemicals). Color development was stopped by incubation in DIG 3 buffer.

2.8. Preparation of nuclear protein extract

Preparation of nuclear extract from CACO-2 cells (3×10^8) was carried out exactly according to the protocol described by Dignam et al. [19]. That from the adult medaka fish intestines or other tissues was carried out according to the method described by Bres and Eales [20]. Protein concentration was determined by the method of Lowry et al., according to Schacterle and Pollack's modification [21,22].

2.9. Electromobility shift assay (EMSA)

Oligonucleotides corresponding to a part of the *OlGC6* promoter (nucleotides −299 to −201) were amplified by PCR, with PCR products being filled in the 5′-adenosine protruding end by Klenow (TaKaRa) and [α-^{32}P]dTTP (Amersham Pharmacia Biotech, Piscataway, NJ, USA). The labeled probe was purified using a CHROMA SPIN+STE-30 column (Clontech, Palo Alto, CA, USA). Oligonucleotides corresponding to a part of the *OlGC6* promoter in both orientations (nucleotides −108 to −79, wild type, designated as E1, or mutations; see Fig. 5A) were synthesized so as to leave a 5′-guanine protruding upon annealing. The annealed double-stranded oligonucleotide was filled in at the protruding 5′-end by Klenow (TaKaRa) and [α-^{32}P]dCTP (Amersham Pharmacia Biotech). The labeled probe was purified by a CHROMA SPIN+STE-10 column (Clontech). A protein–DNA binding reaction was performed in a buffer containing 4% glycerol, 50 mM NaCl, 2.5 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 1 μg of poly(dI-dC)–poly(dI-dC) (Amersham Pharmacia Bio-

tech), 10 mM Tris–HCl (pH 7.5), and the labeled probe. The reaction was initiated by the addition of nuclear extract and was incubated for 30 min at room temperature in order to proceed with the formation of protein–DNA complex. The reaction mixture was then subjected to electrophoresis with a 4% polyacrylamide gel in 0.5×TBE running buffer (1 mM EDTA and 45 mM Tris–borate, pH 8.3). The gel was dried before visualization of the radiolabeled complex by autoradiography. For the competition assays, unlabeled competitor was added to the reaction mixture before addition of the nuclear extract.

2.10. UV cross-linking analysis

A protein–DNA binding reaction was performed in a 1.5-ml tube under the same conditions as those used for the EMSA. Mineral oil was added to the tube to prevent evaporation, and the reaction mixture was then irradiated for 60 min on ice using an ultraviolet lamp with a 254-nm emission. The cross-linked proteins were separated by electrophoresis on a denaturing 12% polyacrylamide gel. The gel was dried and visualized for the radiolabeled complex by autoradiography.

2.11. Database search

Transcription factor binding sequences within the regulatory element were searched using the TFSEARCH program from GenomeNet (http://www.rwcp.or.jp/papia/) [23].

3. Results

3.1. Southern hybridization

A Southern blot hybridization experiment demonstrated that only a single band was detected in each lane, suggesting that only a single copy of the *OlGC6* gene is present on the medaka fish genome (data not shown).

3.2. Characterization of the OlGC6 gene promoter in cultured cell

The transcription initiation site of the *OlGC6* gene, determined by the primer extension analysis, was 213 bp upstream of the putative start codon and was designated as +1 (data not shown). A putative TATA box exists in the position between −30 and −23 and there are several consensus binding sequences for the Cdx2 homeobox protein and GATA family proteins, which are known to be important for intestine-specific gene expression [24,25].

To understand the mechanisms of transcription of the *OlGC6* gene, we analyzed the promoter/enhancer activity of the gene by assaying the luciferase activity of the various luciferase fusion gene constructs and introduced the fusion gene constructs into CACO-2 cells (intestine-derived cell line) or COS1 cells (kidney-derived cell line). With CACO-2 cells, the luciferase activity did not change until deletion down past −299 (Fig. 1). Further deletion down to −18 (deletion of the TATA box) resulted in almost no luciferase activity (Fig. 1). The luciferase activity of a −201/+50 construct was approximately half that of the constructs −299/+50, −491/+50, −1001/+50, and −1757/+50 (Fig. 1). The luciferase activity of −88/+50 and −47/+50 constructs was approximately half that of the constructs −201/+50 and −98/+50, respectively (Fig. 1). Introduction of the 3-bp mutation into the region at −91, −90, and −89 in the construct −1757mutCAG/+50 reduced the luciferase activity down to the same level as that of a −47/+50 construct. These results suggest that there are three important regions (the first, between −299 and −201; the second, between −98 and −88; the third, between −47 and −18) for *OlGC6* gene transcription. On the other hand, although the luciferase activity was measurable with COS1 cells, it remained much lower than that

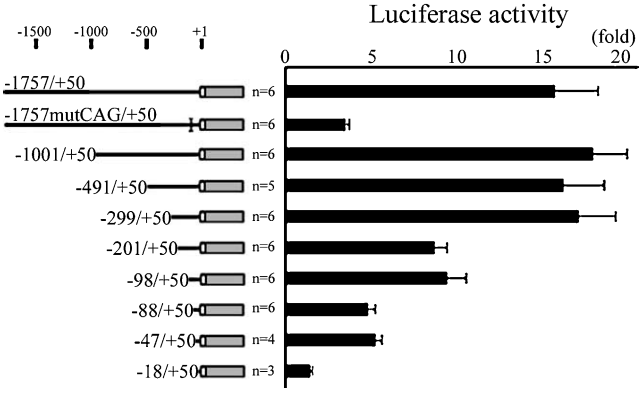

Fig. 1. Summary of *cis*-regulatory element analysis of the *OlGC6* gene using various *OlGC6*–luciferase fusion gene constructs. The structures of the fusion genes were indicated at the left of the figure. White boxes show the 5′-leader sequence of *OlGC6* gene and shaded boxes denote the luciferase open reading frame. An untranscribed 5′-flanking region of the *OlGC6* gene is indicated by a horizontal line. Numbers at the top indicate nucleotide positions relative to the transcription initiation site (+1). The luciferase fusion gene constructs were cotransfected with β-galactosidase control plasmid. On the right side of the figure, black boxes with a bar indicate the luciferase activity in CACO-2 cells. Data represent the luciferase-to-β-galactosidase ratio and are expressed by fold activation relative to the empty luciferase construct as 1. '*n*' denotes the number of independent transfections for each construct.

seen with CACO-2 cells (data not shown). When deleted down to −201, the luciferase activity was almost the same level as that of the empty construct, suggesting that the nucleotides between −98 and −89 are essential to transcription of the *OlGC6* gene only in CACO-2 cells.

3.3. Characterization of the OlGC6 gene promoter in medaka fish

To further understand the regulatory mechanisms of the *OlGC6* transcription in vivo, we generate transgenic medaka fish by microinjecting the fusion gene constructs into the medaka fish embryos. Three transgenic founders were generated for 2kbpGC6-Luc, six founders for 100bpGC6-Luc, four founders for 50bpGC6-Luc, and three founders for 20bpGC6-Luc. Each founder was mated with non-transgenic littermates to establish a line, and the embryos of each F_1 were analyzed to examine the transmission of the transgene. As a result, one line was obtained for 2kbpGC6-Luc, three lines for 100bpGC6-Luc, four lines for 50bpGC6-Luc, and two lines for 20bpGC6-Luc. Two of four lines for 50bpGC6-Luc were lost, because F_1 individuals died before the beginning of the experiments. Transgene expression was detected by RT-PCR analysis. In our previous study using RT-PCR analysis, we demonstrated that the *OlGC6* gene is expressed in the embryo and in the adult medaka fish, mainly in the intestine and to a much lesser extent in the brain, eye, kidney, liver, spleen, and gonad, but not in the gill and heart [11]. In the embryos (stages 35–39), the lines carrying 2kbpGC6-Luc (line 11), and 100bpGC6-Luc (lines 10, 12, and 14) all expressed the luciferase gene. In contrast, the lines with 50bpGC6-Luc (lines 7 and 9) or 20bpGC6-Luc (lines 1 and 2) did not express the transgene (Fig. 2A). In the 100bpGC6-Luc, line 12 medaka fish, the luciferase gene was expressed only in the intestine, although the *OlGC6* gene tran-

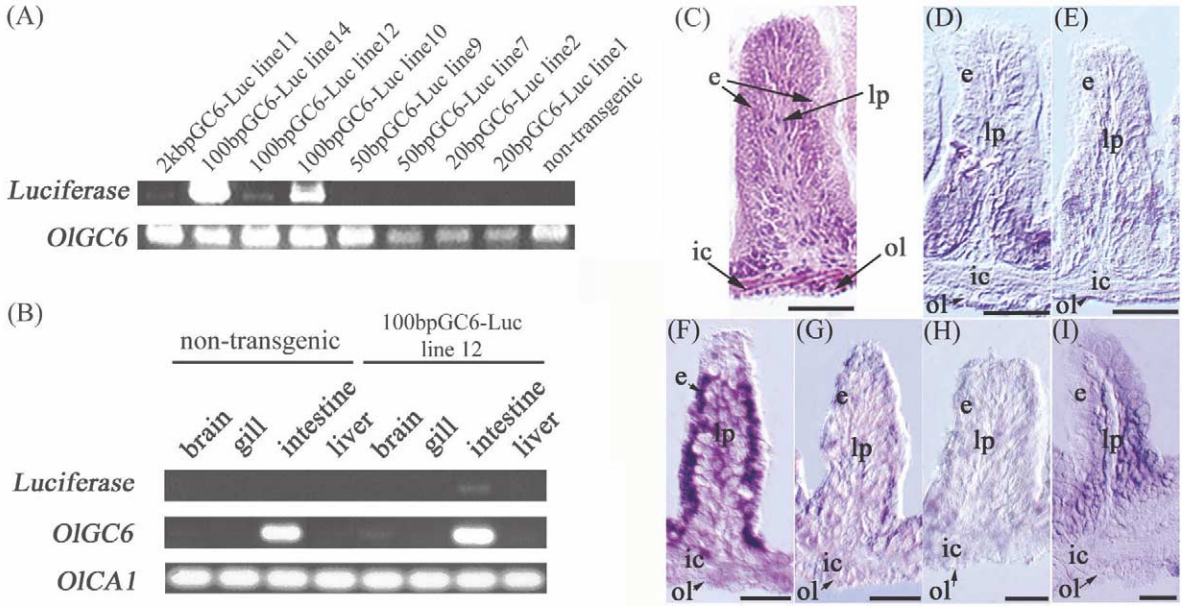

Fig. 2. Transgene expression in the embryos and in the adult intestine. A: Total RNA (2 μg) extracted from F_1 embryos of transgenic lines was reverse-transcribed and PCR amplified using the luciferase or *OlGC6* gene-specific primers. B: Total RNA (500 ng) extracted from the brain, gill, intestine, and liver of a non-transgenic medaka or a 100bpGC6-Luc, line 12 medaka fish was reverse-transcribed and PCR amplified using a specific primer set for the luciferase, *OlGC6* or *OlCA1* gene, which was used as a positive control. C: Hematoxylin-eosin staining of the section from the adult medaka fish intestine. In situ hybridization analysis of the adult medaka fish intestinal section from a non-transgenic medaka fish hybridized using a DIG-labeled *OlGC6* antisense probe (D) and sense probe (E). In situ hybridization analysis of an adult medaka fish intestinal section from a 100bpGC6-Luc line 12 medaka fish (F), or a non-transgenic medaka fish (H) hybridized using a DIG-labeled luciferase antisense probe. G shows absence of staining in the 100bpGC6-Luc line 12 medaka fish when hybridized with a DIG-labeled luciferase sense probe. I shows a 100bpGC6-Luc line 12 medaka fish hybridized with the *OlGC6* antisense probe. ic, inner circular muscle layer; ol, outer longitudinal muscle layer; lp, lamina propria; e, epithelium; scale bar, 50 μm.

script was detected mainly in the intestine and faintly in the brain and liver, but not in the gill (Fig. 2B). The luciferase gene was not expressed in any tissues of the non-transgenic medaka fish, although the *OlGC6* gene was expressed in the same tissues as in the transgenic medaka fish (Fig. 2B). The endogenous *OlGC6* gene was expressed in the epithelial cells in the basement of the intestinal folds (Fig. 2D). The medaka fish is stomachless and the intestine is continuous with the pharynx through a short esophagus [26,27]. Medaka fish intestinal epithelium contains most of the cell types (enterocytes, goblet cells, and enteroendocrine cells) observed in the small intestine of other vertebrates, but lacks crypts and Paneth cells. An intestinal stem cell has not been identified in the teleost intestine, but it has been reported that the intestinal epithelium undergoes continuous renewal and the differentiated enterocytes proliferate in the base of the intestinal fold [26,27]. To localize the cells expressing the luciferase reporter gene, in situ hybridization analysis was performed on sections of the intestine of the 100bpGC6-Luc, lines 10 and 12 or wild type as a control. A DIG-labeled luciferase antisense probe hybridized to the epithelial cells from the basement to the top of the intestinal folds (Fig. 2F). Hybridization with the luciferase sense probe control showed no detectable background staining (Fig. 2G). There was no detectable hybridization of the luciferase antisense probe to the section from wild-type medaka fish (Fig. 2H). As a positive control, the endogenous *OlGC6* transcript was also localized to the epithelial cells in the basement of the intestinal folds (Fig. 2I). These results suggest that the *OlGC6* upstream region between −98 and +50 is sufficient to drive intestinal expression in vivo.

3.4. *Complex formation of the 5′-regulatory region of the* OlGC6 *gene with CACO-2 or medaka intestinal nuclear extracts*

The ability of the first and second regions to form complex(es) with nuclear proteins was determined by incubating

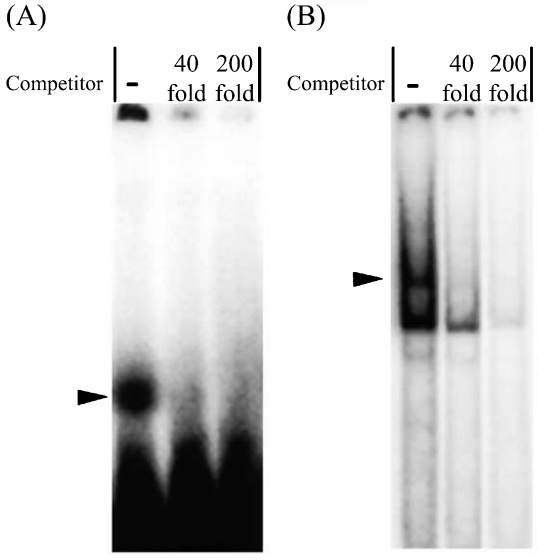

Fig. 3. Electrophoretic mobility shift assay with the E1 probe using nuclear extracts prepared from adult medaka fish intestinal cells (A); and CACO-2 cells (B). Complex competition was achieved by addition of unlabeled E1 (in 40-fold excess or 200-fold excess). Arrowheads indicate specific complexes.

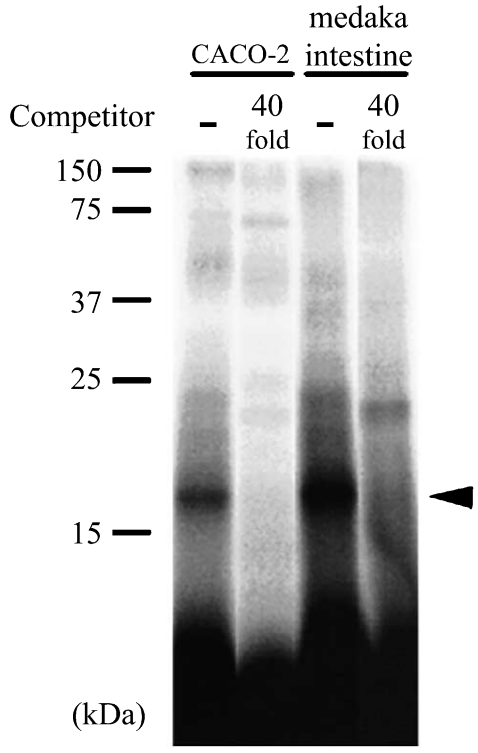

Fig. 4. UV cross-linking analysis using nuclear extracts prepared from CACO-2 cells and the adult medaka fish intestinal cells. Complex competition was achieved by addition of a 40-fold excess amount of unlabeled E1.

an oligonucleotide probe with nuclear extracts prepared from CACO-2 cells or the medaka fish intestine. Several complexes were formed when the labeled first region was incubated with the nuclear extracts prepared from CACO-2 cells, but no complex was formed when incubated with the nuclear extracts from the medaka fish intestine (data not shown). However, two complexes were formed when the E1 probe (containing the second region) was incubated with the nuclear proteins from CACO-2 cells, and a complex was formed when the probe was incubated with the nuclear proteins from the medaka fish intestine (Fig. 3). The complex-forming ability of E1 with the nuclear protein from the medaka fish intestine was abolished by addition of increasing amounts of unlabeled E1 (Fig. 3A). Although two complexes were formed using E1 and the CACO-2 nuclear protein, complex formation (represented by the upper band in Fig. 3B) was inhibited by addition of increasing amounts of unlabeled E1, although the other complex was formed even in the presence of 200-fold amounts of unlabeled E1. Cross-linking and subsequent SDS–polyacrylamide gel electrophoresis (PAGE) experiments indicated that the apparent molecular mass of the complex of E1 and binding protein from CACO-2 cells and the medaka fish intestine was 17 kDa (Fig. 4). However, the mobility of both complexes in non-denaturing gels in the EMSA seems to be slightly different, probably due to the complex formation by E1 and the same nuclear binding protein associated with additional different protein factor or subunit.

3.5. Nucleotides required for protein binding

The nucleotides in E1 responsible for protein binding were identified by EMSA, examining the competition of unlabeled competitors for labeled wild-type E1 binding to CACO-2 cells or the medaka fish intestinal nuclear extracts (Fig. 5). In each assay, competitors 5, 7, and 12 did not compete with labeled wild-type E1, suggesting that the nucleotides AG in the second region are important for binding to the nuclear protein. However, since nucleotides are located at the 3′-end of the second region we also introduced mutations into the further downstream region (in competitors 13, 14, 15, 16, and 17) and examined their competitive activity by the EMSA with CACO-2 cells or medaka fish intestinal nuclear extracts

(A)

Competitor	Sequence
wild type *OlGC6*	5'-G AGCCTGCAGC GCACACACAG ACCTTTGCAC C-3'
1	5'-G AGCCTGCAGC **TA**ACACACAG ACCTTTGCAC C-3'
2	5'-G AGCCTGCAGC GC**CA**ACACAG ACCTTTGCAC C-3'
3	5'-G AGCCTGCAGC GCAC**CA**ACAG ACCTTTGCAC C-3'
4	5'-G AGCCTGCAGC GCACAC**CA**AG ACCTTTGCAC C-3'
5	5'-G AGCCTGCAGC GCACACAC**CT** ACCTTTGCAC C-3'
6	5'-G AGCCTGCAGC **TAC**CACACAG ACCTTTGCAC C-3'
7	5'-G AGCCTGCAGC GCACACA**ACT** ACCTTTGCAC C-3'
8	5'-G AGCCTGCAGC GCA**A**ACACAG ACCTTTGCAC C-3'
9	5'-G AGCCTGCAGC GCAC**C**CACAG ACCTTTGCAC C-3'
10	5'-G AGCCTGCAGC GCACA**A**ACAG ACCTTTGCAC C-3'
11	5'-G AGCCTGCAGC GCACAC**C**CAG ACCTTTGCAC C-3'
12	5'-G AGCCTGCAGC **CCCCCCCCCC** ACCTTTGCAC C-3'
13	5'-G AGCCTGCAGC GCACACACAG **CA**CTTTGCAC C-3'
14	5'-G AGCCTGCAGC GCACACACAG AC**AG**TTGCAC C-3'
15	5'-G AGCCTGCAGC GCACACACAG ACCT**GG**GCAC C-3'
16	5'-G AGCCTGCAGC GCACACACAG ACCTTT**TA**AC C-3'
17	5'-G AGCCTGCAGC GCACACACAG ACCTTTGC**CA** C-3'

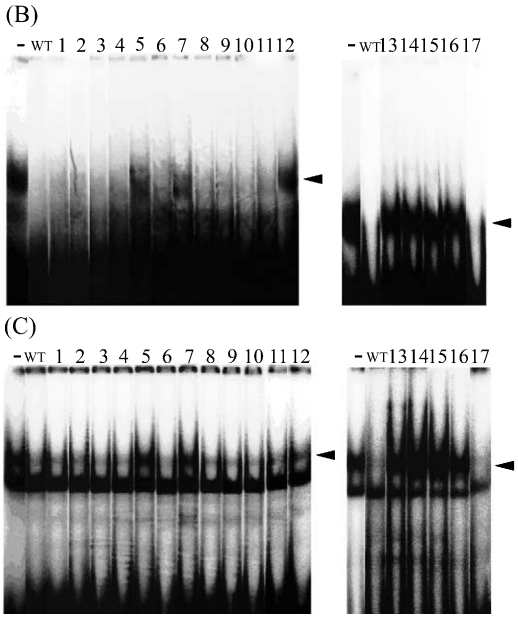

Fig. 5. Competitive EMSA of labeled E1 and unlabeled mutated competitors. A: Competitors for the EMSA. The mutated nucleotides are indicated by bold letters with underline. B: Medaka fish intestinal nuclear extracts were combined with labeled E1 in the presence or absence of a 20-fold excess of the indicated unlabeled competitor. C: CACO-2 nuclear extracts were combined with labeled E1 in the presence or absence of a 40-fold excess of the indicated unlabeled competitor.

(Fig. 5). In each assay, competitors 13, 14, 15, and 16 did not compete with labeled wild-type E1, indicating that the AGACCTTTGC nucleotides are critical for binding to the nuclear protein from CACO-2 cells and the medaka fish intestine.

3.6. Specific complex formation of E1 with a protein in the intestinal nuclear extracts

Several complexes were formed when labeled E1 was incubated with brain, testis, or liver nuclear extracts, and these could be competed by excess unlabeled E1 (data not shown). However, all of these complexes exhibited mobility patterns different from that of the intestine-specific complex (data not shown).

4. Discussion

The 5′-flanking region of the *OlGC6* gene contains several putative consensus binding sequences for transcriptional factors such as GATA protein family members and Cdx2, all of which are involved in the development of intestine [24,25]. In order to understand the mechanisms of transcriptional regulation of the *OlGC6* gene, we analyzed the *cis*-regulatory region of the *OlGC6* gene by transfecting the luciferase fusion gene into mammalian cells and by generating transgenic lines of the medaka fish *O. latipes*. From in vitro transient transfection assays, we identified three regions influencing transcription of the *OlGC6* gene by a stepwise deletion of the 5′-flanking region. The first region was detected between −299 and −201, a region that is important for transcription of the *OlGC6* gene in CACO-2 cells. However, the labeled first region did not form a complex with nuclear proteins from medaka fish intestinal cells in the EMSA. The third region between −47 and −18 contained a putative TATA box, and deletion of the region resulted in a large decrease in luciferase activity, suggesting that the *OlGC6* gene is a TATA-dependent gene in vitro. The second region, between −98 and −89 (GCACACACAG), was important for *OlGC6* transcription in the CACO-2 cells. The E1 probe containing the second region formed specific complexes with a nuclear protein from CACO-2 cells and a specific complex with a medaka fish intestinal nuclear protein (Fig. 3). Furthermore, the complex formed by E1 and the medaka intestinal nuclear protein was intestine specific, suggesting that the second region is important for the transcription of the *OlGC6* gene in the medaka fish intestine. This may be supported by the results obtained from in situ hybridization experiments using 100bpGC6-Luc transgenic medaka fish lines, demonstrating that the luciferase gene in the medaka fish genome was expressed in the adult transgenic medaka fish intestine, although the expression pattern was slightly different from the endogenous *OlGC6* transcript; i.e. the luciferase gene expression was expanded to the top of the intestinal folds (Fig. 2F). RT-PCR analysis also demonstrated that the transgene was detected in the 2kbpGC6-Luc and 100bpGC6-Luc transgenic lines, but not in the 50bpGC6-Luc and 20bpGC6-Luc transgenic lines (Fig. 2A). As shown in Fig. 2B, the expression of the luciferase gene was detected only in the intestine by RT-PCR when the 100bpGC6-Luc, line 12 transgenic medaka fish was examined. From these results, we conclude that the 5′-flanking region of the *OlGC6* gene between −98 and +50 is sufficient to drive the intestine-specific gene expression. On the other hand, our results demonstrating that a nuclear protein from CACO-2 cells or the medaka intestinal cells recognizes the same sequence (AGACCTTTGC) for binding (Fig. 5), and that introduction of mutation into the sequence (the nucleotides AG) resulted in reduction of the promoter activity (Fig. 1), suggest that the binding protein(s) contribute to the transcription of the *OlGC6* gene. Considering that a nuclear protein both from the mammalian and medaka fish intestinal cells binds to the same region, the sequence is clearly important to intestine-specific transcription of the intestinal membrane GC and has been conserved during the evolution of vertebrates. A database search revealed that there is no highly conserved consensus binding sequence in the second region for a known transcriptional factor, suggesting that the second region is a novel regulatory element.

Acknowledgements: We would like to thank Dr. Hiroshi Hori, Nagoya University, for providing the Hd-rR strain and staff members of the Center for Advanced Science and Technology, Hokkaido University, for the use of laboratory facilities. We appreciate Dr. Takehiro Kusakabe, Himeji Institute of Technology, and Dr. Minoru Tanaka and Dr. Atsushi Kimura, Hokkaido University, for their technical advices regarding the construction of fugion genes, the EMSA, and cell culture. We are also grateful to Mrs. Hiroko Kuboshita for taking care of medaka fishes. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (number 11236202) and a grant from The National Project on Protein Structure and Functional Analyses.

References

[1] Cohen, M.B., Mann, E.A., Lau, C., Henning, S.J. and Giannella, R.A. (1992) Biochem. Biophys. Res. Commun. 186, 483–490.
[2] Krause, G., Bayerl, A., Heim, J.M., Singh, S. and Gerzer, R. (1994) Gut 35, 1250–1257.
[3] Swenson, E.S., Mann, E.A., Jump, M.L., Witte, D.P. and Giannella, R.A. (1996) Biochem. Biophys. Res. Commun. 225, 1009–1014.
[4] Schulz, S., Green, C.K., Yuen, P.S. and Garbers, D.L. (1990) Cell 63, 941–948.
[5] Currie, M.G., Fok, K.F., Kato, J., Moore, R.J., Hamra, F.K., Duffin, K.L. and Smith, C.E. (1992) Proc. Natl. Acad. Sci. USA 89, 947–951.
[6] Hamra, F.K., Forte, L.R., Eber, S.L., Pidhorodeckyj, N.V., Krause, W.J., Freeman, R.H., Chin, D.T., Tompkins, J.A., Fok, K.F., Smith, C.E., Duffin, K.L., Seigel, N.R. and Currie, M.G. (1993) Proc. Natl. Acad. Sci. USA 90, 10464–10468.
[7] Vaandrager, A.B. and De Jonge, H.R. (1994) Adv. Pharmacol. 26, 253–283.
[8] Knoop, F.C. and Owens, M. (1992) J. Pharmacol. Toxicol. Methods 28, 67–72.
[9] Schulz, S., Lopez, M.J., Kuhn, M. and Garbers, D.L. (1997) J. Clin. Invest. 100, 1590–1595.
[10] Mann, E.A., Jump, M.L., Wu, J., Yee, E. and Giannella, R.A. (1997) Biochem. Biophys. Res. Commun. 239, 463–466.
[11] Mantoku, T., Muramatsu, R., Nakauchi, M., Yamagami, S., Kusakabe, T. and Suzuki, N. (1999) J. Biochem. 125, 476–486.
[12] Swenson, E.S., Mann, E.A., Jump, M.L. and Giannella, R.A. (1999) Am. J. Physiol. 276, G728–G736.
[13] Park, J., Schulz, S. and Waldman, S.A. (2000) Gastroenterology 119, 89–96.
[14] Di Guglielmo, M.D., Park, J., Schulz, S. and Waldman, S.A. (2001) FEBS Lett. 507, 128–132.
[15] Yamagami, S., Suzuki, K. and Suzuki, N. (2001) J. Biochem. 130, 39–50.
[16] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[17] Ho, S.N., Hunt, H.D., Horton, R.M., Pullenm, J.K. and Pease, L.R. (1989) Gene 77, 51–59.
[18] Mikami, T., Kusakabe, T. and Suzuki, N. (1999) J. Biol. Chem. 274, 18567–18573.

[19] Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) Nucleic Acids Res. 11, 1475–1489.
[20] Bres, O. and Eales, J.G. (1988) Gen. Comp. Endocrinol. 69, 71–79.
[21] Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. (1951) J. Biol. Chem. 193, 265–279.
[22] Schacterle, G.R. and Pollack, R.L. (1973) Anal. Biochem. 51, 654–655.
[23] Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A.E., Kel, O.V., Ignatieva, E.V., Ananko, E.A., Podkolodnaya, O.A., Kolpakov, F.A., Podkolodny, N.L. and Kolchanov, N.A. (1998) Nucleic Acids Res. 26, 362–367.
[24] Suh, E., Chen, L., Taylor, J. and Traber, P.G. (1994) Mol. Cell. Biol. 14, 7340–7351.
[25] Molkentin, J.D. (2000) J. Biol. Chem. 275, 38949–38952.
[26] Harder, W. (1975) Anatomy of Fishes, Eschweizerbartsche Verlagsbuchhandlung, Stuttgart.
[27] Rombout, J.H., Stroband, H.W. and Taverne-Thiele, J.J. (1984) Cell Tissue Res. 236, 207–216.