Cloning of Complementary DNAs Encoding Structurally Related Homeoproteins from Preimplantation Mouse Embryos: Their Involvement in the Differentiation of Embryonic Stem Cells

Koichi Saito, Hajime Abe, Masato Nakazawa, Emiko Irokawa, Masafumi Watanabe, Yusuke Hosoi, Miki Soma, Kano Kasuga, Ikuo Kojima, and Masayuki Kobayashi

Department of Biotechnology, Faculty of Bioresource Sciences, Akita Prefectural University, Akita, Japan

ABSTRACT

During the preimplantation development of mouse embryos between the 4-cell to 8-cell stage and the morula stage, when the first irreversible segregation of cell fates proceeds into the pluripotent inner cell mass (progenitor cells to form the fetus) and the trophectoderm (to form the placenta) of blastocysts, pluripotency-maintaining and differentiation-inducing genes are expressed to coordinately regulate cell fates. Three structurally related cDNAs (Crxos1, Crxos1 sv2, and Crxos1 tv3) that exhibited concomitantly elevated expression during this critical period were identified by subtractive cDNA cloning. CRXOS1 contains two homeodomains, while CRXOS1 sv2 and CRXOS1 tv3 each contain one of the homeodomains included in CRXOS1. Crxos1, Crxos1 sv2, and Crxos1 tv3 were expressed differentially during in vitro embryonic stem (ES) cell differentiation. Even under differentiation-inducing conditions, forced expression of Crxos1 sv2 inhibited the differentiation of ES cells. In contrast, under conditions that promote self-renewal of ES cells, forced expression of Crxos1 induced differentiation. Forced expression of Crxos1 resulted in induction of Gata4 but in repression of T, probably indicating that Crxos1 promotes the differentiation of ES cells into primitive endoderm, while inhibiting differentiation into mesoderm. On the other hand, no apparent effects of forced expression of Crxos1 tv3 were observed. Taken together, it was concluded that these transcripts encoding homeoproteins are capable of regulating the maintenance and/or differentiation of mouse ES cells and likely regulate that of preimplantation embryos.

developmental biology, early development, embryo, ES cells

INTRODUCTION

Embryogenesis is regulated precisely by the coordinated expression of many transcription factors, and studies [1–3] on developmental processes have revealed a number of crucial transcription factors that regulate cell fates. Trophoderm (TE) is the first differentiated cell lineage to arise in mammalian embryogenesis and forms the placenta, a structure unique to mammalian development. In the early mouse embryo, the first irreversible segregation of cell fates occurs between the 4-cell to 8-cell stage and the morula stage, and individual blastomeres are considered to retain the potential to form all cell lineages until the 8-cell stage, as reviewed by Johnson and McConnell [4]. Subsequently, the blastomeres enhance cell-cell contacts to become a compacted morula in the morula stage, and the subsequent cell divisions increase the topological complexity of the morula to differentiate into two distinct cell phenotypes (i.e., TE and the inner cell mass [ICM]) at the 16-cell morula stage. Results of lineage-tracing experiments suggested that most outer cells of the morula can differentiate into TE, and the other inner cells become ICM of the blastocysts [5, 6]. Cells constituting ICM are capable of differentiating into all fetal and adult cell lineages and thus the founders of embryonic stem (ES) cells, which grow rapidly and infinitely, while maintaining pluripotency. Leukemia inhibitory factor (LIF) can maintain the self-renewal of mouse ES cells through the activation of STAT3 [7–9]. On the other hand, embryonal carcinoma (EC) cells derived from the epiblast maintain pluripotency in the absence of LIF.

Both POU5F1 (also known as OCT4), a member of the POU family homeoproteins [10], and NANOG, an NK-2 class homeoprotein [11, 12], are expressed in the ICM of blastocysts, ES cells, and EC cells, thereby maintaining the pluripotency of these cells. Particularly, POU5F1 in ES cells specifies the cell fate by blocking TE differentiation to either maintain pluripotency or differentiate the cells into primitive endoderm [13]. Meanwhile, NANOG blocks primitive endoderm differentiation and actively maintains the pluripotency of ES cells, as the forced expression of NANOG maintains the pluripotency and self-renewal of ES cells even under differentiation-inducing culture conditions [11, 12]. CDX2, a caudal-type homeoprotein, is essential for segregation of the ICM and TE lineages from the morula stage to the blastocyst stage by ensuring the repression of Pous5f1 and Nanog in both the outer cells of morulae and the TE in blastocysts [14–16]. In late blastocysts, GATA6 and GATA4 (zinc finger-containing transcription factors) are expressed in a portion of cells in the ICM to promote the segregation of the extraembryonic endoderm lineages such as visceral endoderm and parietal endoderm [17–19]. Therefore, it was concluded that these pluripotency-maintaining and differentiation-inducing genes are expressed to coordinately regulate cell fates during the preimplantation development of mammalian embryos. However, the mechanisms underlying the preimplantation development of mammalian embryos are not fully understood.

This study focused on morula-stage embryos obtained from mice as an excellent source to identify genes related to either the maintenance of pluripotency or the induction of differentiation. We present the cloning of three structurally related cDNAs that exhibited concomitantly elevated expression from the 4-cell to 8-cell stage to the morula stage in morula embryos.
by subtractive cDNA cloning, as well as the molecular and functional characterization, particularly using mouse ES cells.

MATERIALS AND METHODS

Media

M2 medium [20] was supplemented with 4 mg/ml of bovine serum albumin (BSA) (A4378; Sigma Chemical Company, St. Louis, MO), and Whitten medium [21] was supplemented with 3 mg/ml of BSA. Phenol red was not added to M2 medium or Whitten medium. Dulbecco modified Eagle medium (DMEM), Glasgow modified Eagle medium (GMEM), and M modified Eagle medium were purchased from Sigma Chemical Company.

Embryo Collection and Culture

Induction of superovulation and subsequent mating of 6- to 12-wk-old virgin female CD-1 mice (random bred; Swiss; Charles River Japan, Kanagawa, Japan) were performed as described previously [22, 23]. Two-cell embryos were flushed in M2 medium from excised oviducts 45 to 46 h after human chorionic gonadotropin injection and cultured in 1 ml of Whitten medium in the presence of 3 mg/ml of BSA in 4-well dishes (Nunc, Roskilde, Denmark). All animal procedures conformed to the Guidelines for the Care and Use of Laboratory Animals of Akita Prefectural University.

Culture of Established Cell Lines

Feeder-free mouse ES cell lines, EB3 [24] and MG1.19 [25], were routinely maintained on gelatin-coated plates in GMEM containing 10% fetal calf serum (FCS) (Gibco BRL-Life Technologies, Grand Island, NY) or M2 medium [26]. Mouse fibroblast NIH3T3 cells were cultured in DMEM supplemented with 10% calf serum (CS) (Sigma Chemical Company). Mouse embryonic fibroblasts and monkey kidney-derived COS7 cells were cultured in DMEM supplemented with 10% FCS.

cDNA Synthesis and Suppression Subtractive Hybridization

Total RNA was extracted from 4-cell to 8-cell embryos (~240 embryos), morulae (~250 embryos), and blastocysts (~230 embryos) using an RNaseasy Mini kit (Qiagen, Hilden, Germany) according to manufacturer’s instructions. Because of the small amounts of mRNA extractable from mouse embryos, mRNA was reverse transcribed to cDNA, followed by PCR amplification using a SMART cDNA Synthesis kit (Clontech, Palo Alto, CA). Protocols were performed in accord with the manufacturer’s instructions. Briefly, total RNA extracted from approximately 40 embryos was used for SMART cDNA amplification by PCR (24, 20, and 20 cycles for 4-cell to 8-cell embryos, morulae, and blastocysts, respectively). The PCR-based cDNA subtraction (suppression subtractive hybridization [SSH] [27]) was conducted between the 4-cell to 8-cell embryo-derived cDNA pool as a driver and the morula-derived cDNA pool as a tester using RsaI-digested cDNA fragments (300 ng each) and a PCR-Select cDNA Subtraction kit according to the manufacturer’s instructions (Clontech).

Cloning and Analysis of Subtracted cDNA

After PCR subtraction, the amplified products were cloned into pGEM-T Easy vector (Promega, Madison, WI) and used to transform Escherichia coli JM109. Independent colonies of transformed E. coli were used as a template for colony direct PCR to amplify the insert cDNA using the specifically designated primers in the PCR-Select cDNA Subtraction kit. An aliquot of the PCR products denatured by alkaline treatment (0.3 M NaOH) was arrayed onto a nylon membrane (Hybond-N+; GE Healthcare Bio-Sciences, Piscataway, NJ) in two replicates. For differential screening of subtracted cDNA clones, the membranes were hybridized with forward-subtracted probe (morula transcript-enriched subtracted cDNA) or reverse-subtracted probe (4-cell to 8-cell embryo transcript-enriched subtracted cDNA) labeled with digoxigenin using a PCR DIG Probe Synthesis kit (Roche Diagnostics, Indianapolis, IN). Chemiluminescent detection of hybridized probes was carried out using alkaline phosphatase-labeled anti-digoxigenin Fab fragments and Disodium 3-[4-methoxyispropyl] [2-(diethoxymethylene-3-phosphoryl)] oxirane [3,3,1.1′]-[ediane] 4-yl] phenyl phosphate (CSPD; Roche Diagnostics) as a substrate. The relative extent of differential expression of each clone was estimated by calculating the ratio of the chemiluminescent signal with the forward-subtracted probe to that with the reverse-subtracted probe. The clones for which the ratio was more than 2 were sequenced, and these sequences were then applied to a nucleotide-nucleotide BLAST search against the databases of GenBank (http://www.ncbi.nlm.nih.gov/Genbank/index.html) and the RIKEN FANTOM project (http://fantom.gsc.riken.jp/).

Synthesis of cDNAs for 5′- and 3′-Rapid Amplification of cDNA Ends

First-strand cDNAs for 5′-rapid amplification of cDNA ends (RACE) and 3′-RACE were synthesized from total RNA obtained from morulae (~50 embryos each) using a SMART RACE cDNA Amplification kit (Clontech) according to the manufacturer’s instructions. The gene-specific primers used are listed in Supplemental Table S1 (all Supplemental Data are available at www.biolreprod.org). The 5′-RACE and 3′-RACE products were cloned into pGEM-T Easy vector and used to transform E. coli.

Construction of Expression Vectors

Expression vectors for the genes of interest were constructed using a dicistronic mRNA-expressing pMK10<sub>pg</sub> vector [28]. This vector contains a promoter for expression (also known as β-actin) promoter with the cytomegalovirus immediate early enhancer and rabbit Hbb (also known as β-globin) intron II acceptor site (CAG promoter) [29]. The vector expresses a single mRNA concomitantly encoding a gene of interest, an internal ribosomal entry site (IRES) derived from the encephalomyocarditis virus, and the neomycin-resistance gene. In cells transfected with these vectors, a high degree of secondary structure of the IRES in the transcript is recognized by ribosomes, and the neomycin-resistance gene placed downstream of the IRES is translated independently. Such dicistronic vectors are advantageous to ensure the coexpression of the desired gene in cells expressing a selectable marker gene.

Two series of expression vectors for Crxos members were constructed using pMK10<sub>pg</sub>. One series was for the full-length cDNAs of Croxos (pMK10/ Croxos<sub>1</sub>_Full), Croxos<sub>1</sub> sv2 (pMK10/Croxos<sub>1</sub> sv2_Full), and Croxos<sub>1</sub> sv3 (pMK10/Croxos<sub>1</sub> sv3_Full), and the other series was for only the predicted coding sequence (CDS), with a Kozak sequence to ensure translational initiation from the IRES in these vectors, a high degree of secondary structure of the IRES in the transcript is recognized by ribosomes, and the neomycin-resistance gene placed downstream of the IRES is translated independently. Such dicistronic vectors are advantageous to ensure the coexpression of the desired gene in cells expressing a selectable marker gene.

An episomal expression vector, pMNPur, harboring the CAG promotor, IRES, puromycin-resistance gene, and polyoma ori [25] was generated as follows. The EcoRI-PstI fragment containing the puromycin-resistance gene excised from a pCAG-IP [24] was ligated to the EcoRI-PstI backbone of pMK10<sub>pg</sub>. Thereafter, an EcoRV-SalI fragment containing the polyoma ori from pCAG-IP was introduced into the NruI restriction site of the modified pMK10<sub>pg</sub>. After digestion with PstI and XhoI, followed by blunting treatment, the modified pMK10<sub>pg</sub> was self-ligated to give pMNPur. The Croxos<sub>1</sub> expression vector, pMNPur/Croxos<sub>1</sub>_CDS, was constructed by inserting the CDS for Croxos<sub>1</sub> sv3 into the cloning site of pMNPur. This vector was replicated episomally in transfected cells expressing polyoma large T antigen (such as MG1.19 cells) [25] and expressed Croxos<sub>1</sub> CDS at a high level.

Transfection of DNA and Selection of Stable Transfectants

NIH3T3, COS7, and ES cells were transfected with pMK10<sub>pg</sub>-derived plasmids by lipofection using the LipofectAMINE Reagent (Invitrogen) [30] with slight modifications. Eighteen hours before transfection, NIH3T3, COS7, and ES cells were plated into 6-well plates at densities of 3.0 × 10<sup>4</sup>, 2.0 × 10<sup>4</sup>, and 6.0 × 10<sup>4</sup> cells/well, respectively, and incubated with liposome-DNA complexes in 1 ml of serum-free DMEM for 5 h at 37°C under 5% CO<sub>2</sub>/air. Then, 1 ml of DMEM containing 20% CS for NIH3T3 cells, 20% FCS for COS7 cells, or DMEM containing 2× supplements (FCS, 2-mercaptoethanol, sodium pyruvate, and LIF) for ES cells was added.

For the isolation of stable ES (EB3) cell transfectants, after an additional
18-h incubation, the cells were diluted and divided into new culture dishes (100 mm) with fresh ES cell medium in the presence of 200 μg/ml of G418 (Sigma Chemical Company). Stable transfectants were selected by cultivation with selection medium for approximately 30 days and were then induced to differentiate by changing to −LIF medium or −LIF+RA medium containing 200 μg/ml of G418. After 5 days, colony formation was assessed, and colonies were stained with a nitro blue tetrazolium/5-bromo-4-chloro-3-iodophenyl phosphate (NBT/BCIP) solution for the detection of alkaline phosphatase activity or with Leishman solution [26]. The data were compared according to the percentage of ES cell colonies that indicated undifferentiated or differentiating phenotypes and were analyzed using chi-square test.

For the selection of colonies with high expression levels using an episomal expression system [18, 25], MG1.19 ES cells expressing polyoma large T antigen were supertransfected with the episomal expression vectors. Transfectants were replated and cultured in the presence of 6 μg/ml of puromycin (Sigma Chemical Company) for 2 wk.

**RT-PCR Analysis**

Total RNA fractions were extracted from ES cells using guanidium thiocyanate-phenol [31]. cDNA was synthesized using RT (ReverTra Ace; Toyobo, Tokyo, Japan), oligo-dT20 primer, and total RNA as the template. The expression levels of mRNAs were evaluated by PCR with cDNA as a template. Primer sets are given in Supplemental Table S2. Hydroxymethylbilane synthase (Hmbs) and glyceraldehyde-3-phosphate dehydrogenase (Gapdh) were analyzed as housekeeping genes.

**Preparation of Polyclonal Antiserum Against Bacterially Expressed Protein**

The cDNA fragments corresponding to the majority of the CDS for CRXOS1 sv2 (amino acids 1–116) and the full-length CDS for CRXOS1 tv3 (amino acids 1–142) were inserted into the cloning sites of pGEX-6P-3. The expression vector was transfected into E. coli DH5α, and expression of the foreign gene as a GST fusion protein was induced by the addition of 1 mM isopropyl β-D-thiogalactoside (IPTG). After a 4-h incubation, the cells were harvested and lyzed by sonication at 4°C. Soluble proteins were subjected to glutathione-sepharose column chromatography (GE Healthcare Bio-Sciences) for purification of the GST fusion proteins. Polyclonal antiserum was obtained by immunizing rabbits with each GST fusion protein.

**Western Blotting**

Cellular proteins of NIH3T3 transfected with *Crx* expression vectors on SDS polyacrylamide gels were transferred electrophoretically onto a polyvinylidene fluoride (PVDF) membrane (Immobilon, Millipore, Bedford, MA) and then incubated with antiserum (1:400) against CRXOS1 sv2 or CRXOS1 tv3. Cellular β-actin and *S. japonicum* GST expressed from the *Crx* gene were detected with a mouse anti-human β-actin monoclonal antibody (1:2500; Sigma Chemical Company) and rabbit anti-*S. japonicum* GST polyclonal antibodies (1:10,000; Sigma Chemical Company), respectively. Subsequently, the membrane was reacted with horseradish peroxidase-conjugated goat anti-rabbit IgG antibodies (1:10,000; Chemicon International, Temecula, CA) or horseradish peroxidase-conjugated goat anti-mouse IgG antibodies (1:10,000; GE Healthcare Bio-Sciences) as secondary antibodies. The enzyme reaction was detected using ECL Western Blotting Reagent (GE Healthcare Bio-Sciences) and an LAS-1000 Plus Image Analyzer (Fujifilm, Tokyo, Japan).

To increase the sensitivity and to reduce the background signals of Western blotting for the detection of CRXOS proteins, cellular proteins of preimplantation embryos and ES cells separated by SDS-PAGE, followed by transfer onto the PVDF membrane, were incubated with PVDF Blocking Reagent (Toyobo). Subsequently, the membrane was incubated with anti-CRXOS1 sv2 (1:20,000) or anti-CRXOS1 tv3 (1:20,000) in Immunoreaction Enhancer Solution 1 (Can Get Signal kit; Toyobo). The membrane was incubated with horseradish peroxidase-conjugated goat anti-rabbit IgG antibodies (1:10,000) as secondary antibodies in Immunoreaction Enhancer Solution 2 in the kit. The enzyme reaction was detected using ECL Plus Western Blotting Reagent (GE Healthcare Bio-Sciences) and an LAS-4000 Image Analyzer (Fujifilm).

**RESULTS**

**Application of SSH to Identify Genes Differentially Expressed in Mouse 4-Cell to 8-Cell Embryos and Morulae**

The SSH method was adopted to identify genes whose expression was up-regulated from the 4-cell to 8-cell stage to the morula stage. We constructed a morula-enriched cDNA library subtracted with the 4-cell to 8-cell-stage cDNAs by SSH. Of 288 clones selected randomly by differential expression screening using the digoxigenin-labeled morula-enriched cDNAs as a mixture of probes, 29 candidates (10.1%) were sequenced and subsequently subjected to nucleotide-nucleotide BLAST searches (Supplemental Table S3). The cloning procedure was verified by the presence of clones with high homology to *Nanog* [11, 12] (clone B598) and *Dppa5* [32] (clones B640, B666, and B712), whose expression levels are known to be increased from the 4-cell to 8-cell stage to the morula stage.

The expression patterns of the selected clones B598, B640, B704, and B722 (highly homologous to *Nanog*, *Dppa5*, *Tex20*, and *Dppa3*, respectively) and clone B606 (*Crxos1* tv3, a member of cloned cDNAs and transiently designated as *Egam1C* as detailed herein), together with *Pou5f1*, were confirmed by PCR using cDNAs derived from the 4-cell to 8-cell-stage embryos, morulae, and blastocysts (Fig. 1A). As a result, the expression levels of all these six genes increased from the 4-cell to 8-cell stage to the morula stage.

**Cloning of Three Full-Length cDNAs Encoding Structurally Related Homeoproteins**

**Cloning of *Crxos1* tv3 (**Egam1C**). BLAST searches on the cDNA fragment (272 base pair [bp]) inserted in clone B606 revealed a high homology to the cDNA for *Crxos1* tv3, *Crx* opposite strand 1 transcript variant 3 (XM_988342). *Crxos1* [33] was designated as expressing from the opposite strand of the mouse *Crx* gene, which has an important role in eye development and function [34], and is located 15 kilobase away from the *Crx* gene. The deduced amino acid sequence from the cDNA of *Crxos1* tv3 indicated that the gene encodes a homeoprotein with unknown function. We isolated the full-length cDNA for *Crxos1* tv3 from the morula stage to find the structural differences between the open reading frames (ORFs) derived from both preimplantation embryo and eye. As shown in Figure 1B, the 5′-end and 3′-end sequences of the cDNA were determined by 5′-RACE and 3′-RACE using specific primers *Egam1C_5′RACE* and *Egam1C_3′RACE*, respectively (Supplemental Table S1). Based on the 5′-end and 3′-end sequences, full-length cDNA for *Crxos1* tv3 was amplified by gene-specific primers *Egam1C_fullF* and *Egam1C_fullR* (Supplemental Table S2), followed by determination of the complete nucleotide sequence. Sequence analysis revealed an ORF encoding a protein of 142-amino acid residues that contained a homeodomain between amino acid residues 19 and 78, which was identified and annotated by the SMART server (http://smart.embl-heidelberg.de). The ORF completely matched the corresponding part of the *Crxos1* tv3 cDNA. Because the gene was found from mRNAs expressed at the morula stage and to indicate clearly the structural feature of encoded protein as described later in our study, we transiently designated the cDNA as *Egam* (expressing gene at morula stage) *IC* (581 bp, AB472694) and indicated it with the approved, but confusable, gene symbol.

**Cloning of *Crxos1* (Egam1) and *Crxos1* sv2 (Egam1N).** The nucleotide sequence of the *Crxos1* tv3 (**Egam1C**) was further applied to BLAST searches against the GenBank database, and a high similarity was found within the 3′-end of *Crxos1* (NM_001033638 [33]). The isolation of the *Crxos1* cDNA from morulae by 3′-RACE (Fig. 1B) using the primer *Egam1_3′RACE* resulted in the cloning of two types of structurally related cDNAs. One cloned cDNA contained an ORF completely matching that of the *Crxos1* cDNA. The nucleotide sequence for the 5′-end of the cDNA was
determined by 5'-RACE using specific primer Egam1_5'-RACE (Supplemental Table S1); therefore, the full-length cDNA for Crxos1 was amplified by primers Egam1fullF and Egam1fullR (Supplemental Table S2). The complete nucleotide sequence of Crxos1 cDNA was determined and transiently designated as Egaml (908 bp, AB472692). The ORF of Crxos1 (Egaml) is predicted to encode a protein of 246-amino acid residues and contains two putative homeodomains at amino acid residues 13 to 72 and 123 to 182. The other cloned cDNA demonstrated an ORF that was predicted to encode a putative homeoprotein of 120-amino acid residues. The nucleotide sequence of the 5'-end of the cDNA was determined by 5'-RACE (Fig. 1B) using the primer Egam1_5'-RACE. The full-length cDNA for Crxos1 sv2 (1069 bp) was amplified by primers Egam1fullF and Egam1fullR (Supplemental Table S2), and the complete nucleotide sequence was also determined. Approximately half of the ORF completely matched the 5'-region of Crxos1 and completely matched the entire ORF of Crxos1 splicing variant clone 2 (Crxos1 sv2, AY593966). This cDNA was transiently designated as EgamlN (1069 bp, AB472693).

Expression of Crxos1 and Crxos1 sv2 in Preimplantation Embryos and the Structural Relationship of the Encoded Homeoproteins

The expression of Crxos1 (Egaml) and Crxos1 sv2 (EgamlN) was examined in preimplantation embryos. As shown in Figure 2A, RT-PCR analysis revealed that the expression of Crxos1 and Crxos1 sv2 in 4-cell to 8-cell embryos, morulae, and blastocysts increased from the 4-cell to 8-cell stage to the morula stage in a manner similar to that of clone B606 (Fig. 1A, equivalent to Crxos1 sv3). On the basis of clone B606, three structurally related cDNAs (Crxos1 sv2, and Crxos1 sv3) with plausible homeoboxes were cloned. ENTREZ Gene (http://www.ncbi.nlm.nih.gov/sites/entrez) analysis indicated that mRNAs for Crxos1 (Egaml), Crxos1 sv2 (EgamlN), and Crxos1 sv3 (EgamlC) were transcribed from the Crxos1 gene (gene ID 546024) on chromosome 7.

As shown in Figure 2B, their protein products, CRXOS1 sv2 (EGAM1N) and CRXOS1 sv3 (EGAM1C), were distinct from each other with different homeodomains, while CRXOS1 sv2 (EGAM1N) and CRXOS1 sv3 (EGAM1C), were distinct from each other with different homeodomains. CRXOS1 sv2 (EGAM1N) and CRXOS1 sv3 (EGAM1C) were similar to the engrailed type (35% amino acid identity to mouse ENGRAILED2, P09066).
Expression of CRXOS Proteins in Somatic Cells Transfected with the Expression Vectors for Full-Length cDNA, Preimplantation Embryos, and Undifferentiated and Differentiating ES Cells

The amino acid sequences of CRXOS1 (EGAM1), CRXOS1 sv2 (EGAM1N), and CRXOS1 tv3 (EGAM1C) proteins were determined by computational analysis (AB472692, AB472693, and AB472694, respectively). To identify the translation products, two types of expression vectors for the full-length cDNA and for the CDS, preceded by the Kozak sequence, which ensures the translational initiation from the predicted initiator AUG, were generated for Crxos1 (Egam1), Crxos1 sv2 (Egam1N), and Crxos1 tv3 (Egam1C) and then transiently transfected into mouse fibroblast NIH3T3 cells. The transfected cells were lysed and analyzed by Western blotting for detection of the expressed proteins using polyclonal antisera (Fig. 3A). As a result, it became clear that the CRXOS proteins were expressed from both types of vectors as the expected protein sizes of 29, 14, and 17 kDa, respectively. The transfection efficiencies of these six vectors seemed equivalent because the expression levels of GST protein from the cotransfected GST-expressing vector exhibited approximately the same levels. Furthermore, the intensities of the expressed proteins demonstrated that these proteins were efficiently expressed from both the full-length cDNAs and the predicted CDS. Therefore, it was concluded that all three members of the CRXOS proteins were expressed from their corresponding mRNAs in murine cells.

While the expression levels of Crxos mRNAs in blastocysts were lower than those in morulae, the cell number per embryo was several-fold higher in blastocysts (>100 cells/blastocyst [37]) than in morulae. It was estimated that the total amounts of CRXOS proteins were higher in blastocysts than in morulae; therefore, the expressions of these proteins were analyzed in blastocysts. Using approximately 640 blastocysts, CRXOS1 (EGAM1) and CRXOS sv2 (EGAM1N), but not CRXOS1 tv3 (EGAM1C, data not shown), were detected by Western blot analysis (Fig. 3B). It is likely that the low avidity of anti-CRXOS1 tv3, indicated by comparing the intensities of the CRXOS1 signals (29 kDa) detected by anti-CRXOS1 sv2 and anti-CRXOS1 tv3 (Fig. 3A), resulted in failed detection of CRXOS1 tv3 in preimplantation embryos. Expression levels of CRXOS proteins were also analyzed in a pluripotent mouse ES cell line, EB3. As shown in Figure 3B, the expression of CRXOS1 (EGAM1) was detected at a low level in undifferentiated ES cells, while it was almost undetectable in differentiating cells in the absence of LIF (−LIF) and increased in differentiating cells in the presence of RA (−LIF+RA). On the other hand, the expression of CRXOS1 sv2 (EGAM1N) was low in ES cells cultured in the absence of LIF (−LIF) but higher in both undifferentiated and RA-treated differentiating ES cells. Furthermore, CRXOS1 tv3 (EGAM1C) was detected throughout the experiment, and the expression level was enhanced slightly by the addition of RA (−LIF+RA).
mRNAs for Crxos Were Expressed in Undifferentiated and Differentiating ES Cells and EC Cells but Only in a Limited Number of Adult Tissues

After the study using cDNAs isolated from preimplantation mouse embryos, the expression patterns of Crxos1 (Egam1), Crxos1 sv2 (Egam1N), and Crxos1 tv3 (Egam1C) were examined during in vitro differentiation of pluripotent cell lines such as mouse ES (EB3) cells and mouse EC cells derived from the ICM and epiblast of early embryos, respectively. Because preliminary Northern blot analysis indicated that no signal for Crxos mRNAs was detected in both ES cells and EC cells and adult tissues (data not shown), RT-PCR was used to detect their transcripts.

The expression patterns of Crxos mRNAs (Fig. 4A) were positively correlated to those of the translated proteins in ES cells, as shown in Figure 3B. The transcript of Crxos1 (Egam1) was barely detected by RT-PCR in undifferentiated ES cells and undetected in ES cells expressing the mesoderm marker T and a parietal endoderm marker, tissue plasminogen activator (Plat), in the absence of LIF. On the other hand, Crxos1 (Egam1) mRNA was observed in differentiating ES cells expressing Plat, a visceral endoderm marker α-fetoprotein (Afp), and a neuroectoderm marker Ascl1 (previously known as Mash1) in the presence of RA. The expression level of Crxos1 sv2 (Egam1N) was very low in ES cells cultured without LIF but high in both undifferentiated and RA-induced differentiating ES cells. The expression pattern of Crxos1 tv3 (Egam1C) mRNA remained constant in EC cells throughout the experiment.
These results demonstrated that the expression of Crxos1 (Egam1) mRNA was induced in RA-treated differentiating ES cells and EC cells. The expression of Crxos1 sv2 (Egam1N) mRNA was observed in undifferentiated ES cells and EC cells and RA-treated differentiating ES cells but decreased to a large extent in differentiating ES cells and EC cells as a result of the removal of LIF or the addition of RA.

The expression level of Crxos mRNAs was analyzed by RT-PCR in 16 adult murine tissues (Fig. 4C). The concomitant expression of Crxos mRNAs was detected in tissues such as eye and thymus. The expression of Crxos1 sv2 (Egam1N) and Crxos1 tv3 (Egam1C) was detected in brain and ovary, respectively, and in testis in common but not in bone containing marrow cells. No detectable expression of Crxos mRNAs was seen in NIH3T3 cells or cultured embryonic fibroblasts (data not shown).

**Forced Expression of CRXOS1 and CRXOS1 sv2 Affects the Differentiation of ES Cells**

It became obvious that the expression of CRXOS proteins varied independently during the course of differentiation in ES cells. To explore the contribution of Crxos1 (Egam1), Crxos1 sv2 (Egam1N), and Crxos1 tv3 (Egam1C) to the maintenance of the undifferentiated state of cells or the induction of cell differentiation, cDNAs for individual Crxos members were overexpressed constitutively in mouse ES cells by the method reported previously [24, 36, 38] using dicistronic mRNA-expressing vectors. After transfection of the individual expression vectors (pMK10/Crxos1_CDS, pMK10/Crxos1 sv2_CDS, and pMK10/Crxos1 tv3_CDS), EB3 ES cell colonies were grown in the presence of LIF (+/LIF) and G418 for 15 days. Portions of the transfected ES cells were cultured with +/LIF medium containing G418 for 10 days and subsequently were induced to differentiate by cultivation with medium in the absence of LIF (−/LIF) or in the presence of RA (+/LIF+RA) for 5 days.

During the +/LIF culture (+/LIF) and 3 days after the induction of cell differentiation (−/LIF and −/LIF+RA), the forced expression of CRXOS proteins was detected in transfectants (Fig. 5A). These forced expression levels were 14- to 28-fold, 3.3- to 5.7-fold, and 3.2- to 19-fold those of the control (empty vector) for CRXOS1, CRXOS1 sv2, and CRXOS1 tv3, respectively. The levels of forced expression were almost maintained even 5 days after the induction of differentiation (data not shown).
As shown in Figure 5B, even for cultivation in +LIF medium, transfection with the Crxos1 (Egam1)-expressing vector significantly increased the formation rate of differentiated flat cells harboring no alkaline phosphatase activity (4%, \( P < 0.001 \)), while that of the control vector remained only 0.5%. In addition, transfection of an expression vector for Pou5f1 resulted in significant formation of differentiated ES cell colonies (14%, \( P < 0.001 \)) in the same manner as observed previously [13]. In contrast, in the absence of LIF (−LIF), the rates of undifferentiated tightly aggregated colonies with strong alkaline phosphatase activity increased after transfection with either the Crxos1 sv2 (Egam1N, ∼40%, \( P < 0.001 \)) or Nanog (∼50%, \( P < 0.001 \)) expression vectors. Furthermore, similar results were observed for Crxos1 sv2 and Nanog expression when differentiation was induced by RA treatment (−LIF+RA). However, the effects of forced expression of Crxos1 tv3 (Egam1C) remained unclear.
To determine the effects of forced CRXOS expression in ES cells, the expression levels of marker genes were analyzed by RT-PCR (Fig. 5C). In the absence of LIF (−LIF), semiquantitative analysis revealed that the expression levels of Nanog (0.89-fold), particularly T (0.38-fold), which encodes the earliest marker of mesoderm differentiation [39, 40], were below the levels of the control (empty vector) in Crxos1 (Egam1) transfectants. In transfectants with the Crxos1 sv2 (Egam1N)-expressing vector, the expression levels of Pou5f1 and Nanog were 1.37-fold and 1.90-fold higher, respectively, than those of the control. Although the expression level of Nanog (1.46-fold) was high in Crxos1 tv3 (Egam1C) transfectants, that of Pou5f1 (0.88-fold) was less than that of the control. On the other hand, no apparent changes were observed in the expression levels of Gata4, Gata6, Afp, Plat, or Ascl1 or in +LIF or −LIF+RA culture (data not shown). It is likely that the changes in the ratios of colonies were low in both +LIF and −LIF+RA cultures (Fig. 5B); therefore, the changes in gene expression were hardly detected.

Taken together, these data indicate that the forced expression of CRXOS1 (EGAM1) induced the differentiation of ES cells. In contrast, the forced expression of CRXOS1 sv2 (EGAM1N) suppressed differentiation of ES cells.

Forced Expression of CRXOS1 Probably Induces the Differentiation of ES Cells into Primitive Endoderm

To clarify the fate of ES cells by the overexpression of CRXOS1 (EGAM1), MG1.19 cells (a mouse ES cell line) were supertransfected with pMN1 Pur/Crxos1_CDS, an episomal expression vector for Crxos1 (Fig. 6). This vector replicated episomally in MG.1.19 cells expressing polyoma large T antigen [25] and expressed CRXOS1 protein at a high level (Fig. 6A). As shown in Figure 6B, forced expression of CRXOS1 resulted in induction of the expression of Gata4, a marker gene for primitive endoderm, but resulted in the suppression of T. No apparent changes were observed in the expression levels of Gata6, Afp, Plat, or Ascl1 (data not shown).

FIG. 6. Changes in the expression of differentiation marker genes in ES cells highly expressing CRXOS1 (EGAM1). A) MG1.19 cells, a mouse ES cell line, were transfected with either an episomal expression vector pMN1 Pur/Crxos1_CDS (Crxos1) or empty vector (Empty). Transfectants were cultured in the presence of LIF and puromycin for 2 wk. Cell lysates (1.2 × 105 cells/lane for CRXOS1 and 1.2 × 105 cells/lane for ACTB) were separated by SDS-PAGE, and CRXOS1 and ACTB were subsequently detected by Western blotting. B) After transfection, MG1.19 cells were cultured in the presence of LIF (+LIF), as indicated in A. Then, the expression of differentiation marker genes was analyzed by RT-PCR. The numbers under each panel represent the relative densities of amplified cDNA signals normalized with the density of Gapdh. The sizes of the PCR products indicated are as follows: Gata4 (229 bp), T (834 bp), and Gapdh (1083 bp).

DISCUSSION

In mouse embryos, the first irreversible segregation of cell fates proceeds between the 4-cell to 8-cell stage and the morula stage, and pluripotency-maintaining and differentiation-inducing genes are thought to be expressed to establish two types of cell lineages (i.e., ICM and TE) in blastocysts, thereby giving rise to lineages for the embryo and placenta, respectively. Mouse ES cells are derived from ICM of 3.5-day-old blastocysts and maintained as a pluripotent self-renewing population by cultivation in the presence of LIF. When LIF is removed, the cells undergo apparently spontaneous in vitro differentiation, with commitment to a range of embryological lineages. Therefore, the ES cell culture system is useful for the analysis and characterization of molecular pathways or the functions of specific genes involved in pluripotency and differentiation. In fact, using this system for functional expression cloning [12] and subtractive cDNA cloning [32], genes have been identified encoding transcription factors such as Nanog and Dppa5, both of which are associated with pluripotency and show up-regulated expression from the 4-cell to 8-cell stage to the morula stage.

Assuming that the preimplantation embryos, especially morulae, provide an excellent source for finding both pluripotency-maintaining and differentiation-inducing genes, SSH was applied to identify three structurally related cDNAs,
Crxosl in differentiating ES cells and EC cells may indicate that Crxosl sv2 (Egam1N) is associated with the maintenance of pluripotency, while Crxosl (Egam1) is associated with differentiation.

Overexpression of Nanog resulted in LIF-independent self-renewal of ES cells [11, 12], while Pou5f1 overexpression led to differentiation of the cells into a primitive endoderm lineage [13]. It is well known that alkaline phosphatase activity is associated with the undifferentiated state of pluripotent cells; therefore, in situ staining of alkaline phosphatase activity is generally used to identify undifferentiated ES cells, as reported elsewhere [24, 36]. Western blot analysis clearly indicated that CRXOS1 (EGAM1) and CRXOS1 sv2 (EGAM1N) were expressed in blastocysts. Considering the estimated cell number (~640 blastocysts (~6.4 x 10^6 cells/lane)) in Figure 3B, it is postulated that the contents of CRXOS1 (EGAM1) and CRXOS1 sv2 (EGAM1N) in blastocysts are much higher than those of ES cells. The forced expression levels of CRXOS1 (EGAM1) and CRXOS1 sv2 (EGAM1N) were regarded as comparable to those in blastocysts (i.e., 0.9- to 1.5-fold and 0.4- to 1.5-fold those of blastocysts, respectively). In this condition, CRXOS1 sv2 maintained tightly aggregated ES cell colonies exhibiting strong alkaline phosphatase activity under differentiation-inducing conditions, and CRXOS1 induced the formation of flat differentiated ES cell colonies with no alkaline phosphatase activity under conditions that promote the self-renewal of ES cells, suggesting that CRXOS1 sv2 inhibits and CRXOS1 induces the differentiation of ES cells. In the absence of LIF, the molecular analysis revealed that a decrease in the expression level of Nanog was accelerated in Crxosl transfectants, while Nanog and Pou5f1 in Crxosl tv2 transfectants were suppressed, which probably resulted from their maintained expression levels. These results support the possible function of these proteins. However, a portion of drug-resistant colonies (4%-40%) was affected in the differentiating state, as shown in Figure 5B, although the expression of each CRXOS protein was ensured by the concomitant expression of a drug-resistant gene using dicistronic mRNA-expressing vectors. It is well known that the expression of foreign genes is potently silenced in ES cells [42-44]. We believe that the effects of Crxosl expression vectors were exerted in the differentiating state of drug-resistant ES cells only when sufficient expression levels of CRXOS proteins were obtained. Although statements about the effect of forced expression in a mixture of antibiotic-resistant cells should be judged with caution, the effect of CRXOS1 (EGAM1) or CRXOS1 sv2 (EGAM1N) expression was detected in only one of the culture conditions, including the undifferentiated state-maintaining conditions (+LIF) and differentiation-inducing conditions (–LIF and –LIF+RA), suggesting that the effect of forced expression of each protein was strong and unilateral. However, the effect of forced CRXOS1 tv3 (EGAM1C) expression was probably weak because no obvious effect was seen throughout the experiments. It is also possible that an apparently decreased effect or no effect in the respective cDNA transfectants was due to squelching, as is often observed in high expressers [45]. Further work will be required to address the detailed function of CRXOS proteins by establishing the transfected ES cell clones with these cDNAs.

In the case of limited cell populations expressing CRXOS1 (EGAM1) in morula and blastocysts, the expression level in the cells should be higher than that in EB3 transfectants. Therefore, the Crxosl expression vector was supertransfected into MG1.19 ES cells to express CRXOS1 at a high level to clarify the fate of ES cells. The expression level of CRXOS1 was estimated as more than 50-fold higher than that of EB3 transfectants by Western blotting (data not shown). The overexpression of CRXOS1 induced the expression of Gata4, despite the presence of LIF, and suppressed the expression of T, which indicated a possibility that CRXOS1 promotes the differentiation of ES cells into primitive endoderm. A recent study [46] reported that ES cell cultures normally contain a population of cells expressing T, designated as early mesoderm-specific progenitors, and the present study demonstrated the mutually exclusive expression of Crxosl and T in differentiating ES cells. Therefore, it is likely that the overexpression of Crxosl promotes the differentiation of ES cells into primitive endoderm, while it suppress the number of mesoderm-specific progenitors or the spontaneous differentiation of ES cells into mesoderm. These results indicate that Crxosl (Egam1) and Crxosl sv2 (Egam1N) are capable of regulating the differentiation of ES cells in positive and negative fashions, respectively. Direct analysis such as knockout or knockdown experiments against the Crxosl gene will be required to elucidate the exact functions of Crxosl mRNAs and its encoded proteins in preimplantation embryos.

Western blot analysis also demonstrated that all the CRXOS proteins were observed in NIH3T3 cells by forced expression of the respective full-length cDNAs. Therefore, Crxosl mRNAs can be translated into the corresponding proteins not only in ES cells and preimplantation embryos but also in somatic cells. The expression of Crxosl mRNAs in limited tissues (gonads, eye, brain, and thymus) seems indicative of additional roles in the regulation of differentiation and maintenance of several types of somatic cells or germ cells. In fact, as reported by Alfano et al. [33], Crxosl (Egam1) is expressed in the retina of adult mice, and forced expression of Crxosl in the retina reduced the expression of Crxosl.

Many members of homeotic gene families are key players and act as transcriptional master regulators during embryonic development and adult cell differentiation [47, 48]. The homeobox encodes the 60-amino acid homeodomain, which represents the DNA-binding domain of much larger homeo-proteins [41]. The presence of two different homeodomains in CRXOS1 (EGAM1) and one each in CRXOS1 sv2 (EGAM1N) and CRXOS1 tv3 (EGAM1C) suggests that these proteins may act as transcriptional regulators and govern the DNA-binding activity. Outside of the homeodomains, no other conserved motifs were found in these three proteins. Furthermore, the present results suggest that the homeodomain in CRXOS1 sv2 (EGAM1N) and either one or both in CRXOS1 (EGAM1) are responsible for the regulation of target genes involving the inhibition and promotion, respectively, of ES cell differentiation. On the other hand, no obvious effect of Crxosl tv3 (Egam1C) overexpression was seen on the differentiation of ES cells. Further study is required to address the function of Crxosl tv3 in ES cells.

In conclusion, the present study isolated cDNAs encoding structurally related homeoproteins CRXOS1 (EGAM1), CRXOS1 sv2 (EGAM1N), and CRXOS1 tv3 (EGAM1C) from preimplantation mouse embryos. It was demonstrated that these proteins are capable of regulating the maintenance and/or differentiation of mouse ES cells and likely regulate that of preimplantation embryos.

ACKNOWLEDGMENTS

We thank Dr. Hitoshi Niwa (RIKEN Center for Developmental Biology, Kobe, Japan) for providing EB3 cells, human LIF-expressing vector, pCAG-IP vector, and mouse Pou5f1 cDNA. We thank Dr. Max Gassmann (University of Zurich, Zurich, Switzerland) for providing pMGD20neo vector and MG1.19 cells. We are grateful to Drs. Nobuhiro Ban and Daisuke Watanabe (Akita University School of Medicine, Akita, Japan) for
REFERENCES


