Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 8145–8149, September 1993
Developmental Biology

Csx: A murine homeobox-containing gene specifically expressed in the developing heart

(cardiac development/transcription factor/tissue-specific gene expression/embryonic stem cell/evolutionary conservation)

ISSEI KOMURO AND SEIGO IZUMO*

Molecular Medicine and Cardiovascular Divisions, Beth Israel Hospital, and Department of Medicine and Program in Cell and Developmental Biology, Harvard Medical School, Boston, MA 02215

Communicated by Louis M. Kunkel, May 25, 1993 (received for review February 4, 1993)

ABSTRACT **The molecular control of the differentiation process depends in part on lineage-restricted transcription factors that regulate expression of tissue-specific genes. Although significant progress has been made in molecular understanding of skeletal muscle differentiation, no information is available concerning the genes involved in development of the heart, the first organ to form in vertebrate embryos. Many vertebrate homeobox-containing genes have been shown to be expressed in broad regions of the mouse embryo, but no expression of a homeobox gene has been found in the most anterior region of the early embryo, the heart primordium. We report here on the cloning of a murine homeobox cDNA, *Csx* (cardiac-specific homeobox). The *Csx* homeodomain sequence is divergent from those of the *Hox* class genes but is related to that of *Drosophila msh-2* (*NK-4*), which plays a key role in *Drosophila* heart formation. *Csx* is conserved in evolution and *Csx* homologs exist in all vertebrates examined. Transcripts of *Csx* are detected from the presomite stage (7.5 days postcoitum), when mesoderm differentiates into promyocardium. *Csx* expression is restricted in the myocardial cells from 8.5 days postcoitum through adult. *Csx* is not expressed in skeletal or smooth muscle or any other tissues examined. Expression of *Csx* precedes that of cardiac-specific genes in embryonic stem cells differentiating into beating myocardial cells *in vitro*. Although physiological function of *Csx* is yet to be determined, the temporal and spacial pattern of *Csx* expression raises a possibility that *Csx* may play a critical role in the differentiation of cardiac cells.**

Concerted activation of regulatory genes plays a fundamental role in determining the temporal and spacial patterns of embryonic development. Lineage-restricted transcription factors that regulate tissue-specific genes are especially important for tissue differentiation. The isolation and extensive characterization of the MyoD gene family have brought significant progress to the molecular understanding of skeletal muscle differentiation (1). Although many genes expressed in skeletal muscle are also expressed in cardiac muscle, the MyoD gene family is not expressed in the heart. To date, no cardiac-specific helix–loop–helix-type gene (like the MyoD family) has been isolated, despite intensive efforts by many laboratories. This raises the possibility that there might be very divergent MyoD-like genes expressed in the heart or, alternatively, that differentiation of cardiac muscle may be controlled by transcription factors other than the helix–loop–helix type.

Much progress has been made in unraveling the regulatory events of differentiation process of *Drosophila* (2). Among the genes that govern development of *Drosophila*, the sequential activation of homeotic and segmentation genes controls the identity, polarity, and number of body segments (3). Many such genes, including the Antennapedia (*Antp*), Engrailed (*En*), and Paired (*Prd*) families, contain a characteristic 180-bp sequence motif called the homeobox (4). Homeodomain-containing proteins act as sequence-specific transcription factors that transregulate the expression of other genes (5). Many homologs of invertebrate homeodomain proteins have been isolated in mammals, including mice and humans (4). The best-studied vertebrate homeobox-containing genes are *Antp*-like *Hox* genes, which exist in four major clusters in the mouse genome (6). Each cluster exhibits intriguing similarities with the complement of genes within the fly *Antp* and *Bithorax* clusters, not only in homeodomain sequences but also in the temporal order of activation, anterior boundary of expression during embryogenesis, and possible role in segmentation (4). Extensive data of ectopic expression and null mutation of *Hox* genes suggest that mammalian *Hox*-type homeobox genes might function similarly to the invertebrate homologs during development (7–9).

A homeobox gene whose expression is restricted to specific cell lineages would be of particular interest as a candidate for a "cell-type commitment" gene. Some members of the POU gene family, a class of homeobox-containing genes, are expressed in developing central nervous system and show restricted expression patterns in the adult brain (10). One of the POU family genes, *Pit-1*, which is expressed only in the anterior pituitary gland, is necessary for formation of pituitary cells and activates pituitary-specific genes such as the growth hormone and prolactin genes (11).

Recently, the *Drosophila* homeobox-containing gene *msh-2* (*NK-4*) has been shown to be expressed in the developing dorsal vessel, an insect equivalent of the vertebrate heart (12). Mutations in *msh-2* (*NK-4*) gene do not affect mesoderm invagination or dorsal spreading but result in loss of heart formation in embryo (12). This suggests that *msh-2* (*NK-4*) plays a critical role in *Drosophila* heart development. Because the genes that play key roles in cell differentiation are likely to be conserved in evolution, we searched for a mammalian homolog of *msh-2* (*NK-4*), which may play a critical role in vertebrate heart development. Here, we report on the isolation and characterization of a murine gene whose homeodomain sequence† has similarity to that of *msh-2* (*NK-4*) but differs significantly from that of any *Hox* class genes. This gene is expressed from the time of heart differentiation and its cardiac-specific expression continues from early embryo through adult. To indicate this restricted expression, this cDNA is denoted *Csx*, for cardiac-specific homeobox. Expression of *Csx* precedes that of cardiac-

Abbreviations: ES, embryonic stem; RT-PCR, reverse transcription PCR; p.c., postcoitum; αMCH, α-myosin heavy chain; EB, embryoid body.

*To whom reprint requests should be addressed at: Molecular Medicine Division, Beth Israel Hospital, 330 Brookline Avenue, Boston, MA 02215.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. L20300).

specific genes in embryonic stem (ES) cells differentiating into beating myocardial cells *in vitro*.

MATERIALS AND METHODS

cDNA Library Screening and DNA Sequencing. The homeodomain of a mouse *NK-3* homolog (denoted *Kbx*) was first isolated by polymerase chain reaction (PCR) using degenerate primers corresponding to *Drosophila NK-3* homeodomain (amino acids 1–7 and 48–54) from an 8.5-day mouse embryo cDNA library (13). PCR was performed by denaturing at 94°C for 1 min, annealing at 42°C for 30 sec, and elongation at 72°C for 1 min. A mouse genomic library constructed in cosmids (gift from E. Geissler, Beth Israel Hospital) was screened using mouse *Kbx* cDNA. A genomic clone that contains two homeobox genes (*Kbx* and *Imx*) was isolated. Using the *Imx* homeodomain as a probe, *Csx* was isolated from a mouse cardiac cDNA library (Stratagene) under high stringency conditions.

Southern and Northern Blot Analysis. Each lane of Southern blot contained 5 μg of *Eco*RI-digested DNA. Hybridizations were performed in 6× SSC, 5× Denhardt's, 0.5% SDS, and 50 μg of salmon sperm DNA per ml at 65°C. The blot was washed three times under low stringency in 2× SSC at 65°C. Ten or 20 μg of poly(A)$^+$ RNA was size-separated on 1% agarose/formaldehyde gels and then transferred to a nylon membrane (Hybond N, Amersham). Hybridization was performed at 42°C in a buffer containing 40% formamide, followed by serial washings with a final wash in 0.1× SSC/0.1% SDS at 60°C. The cDNA probe used for Southern and Northern blot analysis was a 5′ part of the *Csx* cDNA (nucleotides 1–305).

***In Situ* Hybridization Analysis.** *In situ* hybridizations were carried out as described (14). The oligonucleotides correspond to the sense and antisense strands of the *Csx* cDNA nucleotides 18–65. The antisense strand DNA of nucleotides 307–354 was also synthesized. After hybridization, sections of 18-day-old embryo were exposed to Hyperfilm β-Max x-ray film (Amersham) for 7 days; other embryo sections were dipped in NTB-2 nuclear track emulsion (Kodak) and slides were developed in D19 (Kodak) for 3 min after exposure for 3 weeks. Sections were counterstained with hematoxylin/eosin and mounted in Permount (Fisher).

Reverse Transcription-PCR (RT-PCR) Analysis. Reverse transcription was carried out using random hexamers on total RNA extracted from five embryos at each stage or from an adult heart. PCR was performed for 35 cycles, with each cycle consisting of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min. The primers used were nucleotides 229–252 and 427–449 of *Csx* cDNA.

ES Cell Culture and *in Vitro* Differentiation of ES Cells. The culture and differentiation of ES cells were carried out according to ref. 15. In brief, cells of the ES cell line ES-D3 were propagated in high-glucose Dulbecco's modified Eagle's medium (DMEM), 0.1 mM 2-mercaptoethanol, and 15% fetal calf serum (FCS) with mitomycin C-treated STO cells as a feeder layer. For differentiation into embryoid body (EB), ES cells were removed from the feeder layer and transferred to 100-mm plastic bacterial Petri dishes containing 10 ml of maintenance medium (DMEM and 15% FCS), which was changed every other day.

RNase Protection Analysis. The *Csx*-containing plasmid was linearized by *Pst* I and [^{32}P]UTP-labeled antisense RNA probe was generated using T3 RNA polymerase (Stratagene). For the α-myosin heavy chain (αMHC) complementary RNA probe, 205-bp carboxyl-terminal coding and 23-bp 3′ noncoding regions of the mouse αMHC cDNA were isolated by PCR using the oligonucleotide primers 5′-ACAAGCTGCAGCTGAAGGTG-3′ and 5′-GCGAGGGTCTGCTGGAGAGGTTATTCCTCG-3′ (15). The PCR fragment was subcloned into pBluescript and linearized for *in vitro* transcription. RNase protection assay was performed using 40 μg of total RNA as described.

RESULTS

Isolation of Cardiac-Specific Homeobox cDNA, *Csx*. We first attempted to isolate a mouse homolog of *msh-2* (*NK-4*) by low stringency hybridization from an 8.5-day embryonic library. We isolated a homeobox-containing cDNA, denoted *Gtx* (14). The third helix of *Gtx* homeodomain has >80% identity with that of *msh-2* (*NK-4*). However, overall homology of *Gtx* homeodomain with that of *msh-2* (*NK-4*) is only 43% and *Gtx* is not expressed in the heart (14). This makes *Gtx* highly unlikely to be a mouse *msh-2* (*NK-4*) homolog. In *Drosophila*, *msh-2* (*NK-4*) gene was known to be localized tandemly with *NK-3* on the same chromosome (16). We first isolated a potential mouse *NK-3* homolog (*Kbx*) by PCR and subsequent screening of the 8.5-day embryonic cDNA library. The homeodomain sequence of *Kbx* has a 77% identity to that of *Drosophila NK-3* (unpublished results). To isolate a homeobox gene that may be linked to *Kbx*, a mouse genomic library constructed in cosmids was screened with the *Kbx* cDNA probe. We isolated a genomic clone that contained two homeoboxes (*Kbx* and *Imx*) separated by ≈5 kb (I.K., H. Inagaki, and S.I., unpublished results). The *Imx* homeodomain is 67% homologous to *Drosophila NK-4*. Since

a

```
1      TCCCTGGATGACAGGAGCGACGGGCAGTTCTGCGTCACCCAGTCTAGAAGCGGTGATCGC
61     CATTTCCTGTTTGTACATCCTGAACCTGGAGCAGCAGCAGCGTAGCTGGCGTCTGGGGAC
121    CTGTCTGCGCGCCTCGAGGCCACCCTGGCCCCTGCCTCCTGCATGCTGGCCGCCTTCAAG
181    CCCGAGGCCTACTCTGGCCCCGAGGCGGCAAGGTCCGGCCTGGCAGAGCTGCGCGCGGAG

241    ATGGGCCCCGCGCCTTCGCCCCCCAAGTGCTCTCCTGCTTTCCCAGCCGCCCCCACATTT
       MetGlyProAlaProSerProProLysCysSerProAlaPheProAlaAlaProThrPhe

301    TACCCGGGAGCCTACGGTGACCCTGACCCAGCCAAAGACCCTCGGGCGGATAAAAAGAG
       TyrProGlyAlaTyrGlyAspProAspProAlaLysAspProArgAlaAspLysLysGlu

361    CTGTGCGCGCTGCAGAAGGCAGTGGAGCTGGACAAAGCCGAGACGGATGGCGCCGAGAGA
       LeuCysAlaLeuGlnLysAlaValGluLeuAspLysAlaGluThrAspGlyAlaGluArg

421    CCACGCGCACGGCGGCGACGGAAGCCACGCGTGCTCTTCTCGCAGGCGCAGGTCTACGAG
       ProArgAlaArgArgArgArgLysProArgValLeuPheSerGlnAlaGlnValTyrGlu

481    CTGGAGCGGCGCTTCAAGCAACAGCGGTACCTGTCGGCGCCAGAGCGCGACCAGCTGGCC
       LeuGluArgArgPheLysGlnGlnArgTyrLeuSerAlaProGluArgAspGlnLeuAla

541    AGCGTGCTGAAGCTCACGTCCACGCAGGTCAAGATCTGGTTCCAGAACCGTCGCTACAAG
       SerValLeuLysLeuThrSerThrGlnValLysIleTrpPheGlnAsnArgArgTyrLys

601    TGCAAGCGACAGCGCAGGACCAGACTCTGGAGCTTCTGGGGCCGCCGCCGCCGCCCGCGC
       CysLysArgGlnArgArgThrArgLeuTrpSerPheTrpGlyArgArgArgArgProArg

661    GCAGGATCGCGGTGCCCGTGCTGGTGCGCGACGGGAAGCCCTGCCTGGGGGACCCCGCGG
       AlaGlySerArgCysProCysTrpCysAlaThrGlySerProAlaTrpGlyThrProArg

721    CCTACACTCCCGCCTACGGTGGGTCTCAATGCCTATGGCTACAACGCCTACCCCTACCCC
       ProThrLeuProProThrValGlyLeuAsnAlaTyrGlyTyrAsnAlaTyrProTyrPro

781    AGCTACGGCGGCGCGGCCTGCAGTCCCGGCTACAGCTGCGCCGCCTACCCCGCTGCGCCC
       SerTyrGlyGlyAlaAlaCysSerProGlyTyrSerCysAlaAlaTyrProAlaAlaPro

841    CCCGCCGCGCAGCCCCCGCCGCCTCCGCCAACAGCAACTTCGTGAACTTTGGCGTCGGGG
       ProAlaAlaGlnProProProProProProThrAlaThrSerEnd

901    ACTTGAACACCGTGCAGAGTCCCGGGATGCCGCAGGGCAATTCGGGCGTCTCCACGCTGC
961    ACGGCATCCGAGCCTGGTAGGGAn
```

b

```
                  helix 1              helix 2            helix 3
             --------------        ------------    ---------------------

Csx   RRKPRVLFSQ AQVYELERRF KQQRYLSAPE RDQLASVLKL TSTQVKIWFQ NRRYKCKRQR
NK-1  P-RA-TA-TY E-LVS--NK- -TT----VC- -LN--LS-S- -E-------- ---T-W-K-N
NK-2  K--R----TK --T------- R--------- -EH---LIR- -P-------- -H---T--AQ
NK-3  KKRS-AA--H ---F------ A------G-- -SEM-KS-R- -E-------- -----T--KQ
NK-4  K--------- -------C-- RLKK--TGA- -EII-QK-N- SA-------- -----S--GD
(msh2)

Antp  -KRG-QTYTR Y-TL---KE- HFN---TRRR -IEI-HA-C- -ER-I----- ---M-W-KEN
bcd   P-RT-TT-TS S-IA---QH- L-G---T--R LAD-SAK-A- GTA------K ---RRHKI-S
en    EKR--TA--S E-LAR-K-E- NEN---TERR -Q--S-E-G- NEA-I----- -K-A-I-KST
```

FIG. 1. Nucleotide sequence and deduced amino acid sequence of *Csx*. (*a*) Nucleotide sequence and amino acid sequence of *Csx* cDNA. The predicted amino acid sequences are presented under the nucleotide sequences. The homeodomain is underlined. (*b*) Comparison of *Csx* homeodomain with other homeodomains. Dashes indicate amino acid identity with *Csx*. The three helix motifs of the homeodomain are overlined.

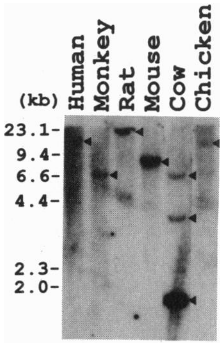

FIG. 2. Southern blot analysis of *Eco*RI-digested genomic DNA from human, monkey, rat, mouse, cow, and chicken. The probe was a 5′ portion of *Csx* (nucleotides 1–305). The hybridizing bands are indicated with arrowheads.

Imx transcript is expressed in several tissues, including the heart, a mouse cardiac cDNA library was screened using the *Imx* homeodomain as a probe. A cDNA clone, denoted *Csx*, was isolated. Interestingly, the homeodomain sequence of *Csx* was significantly different from that of *Imx* genomic clone (77% identity), indicating that we isolated a related, but distinct, gene product. The *Csx* cDNA sequence has an open reading frame that is predicted to encode a polypeptide of 214 amino acids (Fig. 1*a*). An in-frame stop codon is found 138 nucleotides upstream of the putative initiation codon. The predicted amino acid sequence of the *Csx* homeodomain shows a very limited homology (between 30% and 50%) to most homeodomain sequences except for those of the *NK* family (16); it is most similar to that of the *NK-2* family (17) and next to *msh-2* (*NK-4*) (Fig. 1*b*).

Southern Blot Analysis. To examine whether other vertebrate genomes contain *Csx*-related genes, Southern blot analysis of *Eco*RI-digested genomic DNA from several species was performed using a probe that does not contain the homeodomain of *Csx* (Fig. 2). The result showed a single band in mouse DNA and one to three bands in DNA from humans, monkeys, rats, cows, and chickens. Furthermore, we have recently isolated a *Xenopus* homolog of *Csx* (unpublished results). These results suggest that *Csx* is highly conserved among many vertebrates, including amphibian, avian, and primate species.

Expression of *Csx* in the Heart. To examine the tissue distribution of *Csx* transcripts, poly(A)$^+$ RNA prepared from various tissues of adult mice was analyzed by Northern blot hybridization (Fig. 3*a*). In adult mice, *Csx* is expressed only in the heart. *Csx* expression was not detected in other adult tissues such as brain, intestine, kidney, liver, lung, ovary, placenta, skeletal muscle, skin, spinal cord, spleen, or uterus (Fig. 3*a*). To examine developmental regulation of *Csx* expression, poly(A)$^+$ RNA was isolated from the hearts of 15-day embryonic and 2-day neonatal mice for Northern blot analysis. *Csx* was expressed abundantly in the embryonic and neonatal heart as well as in the adult heart (Fig. 3*b*). Next, we performed RT-PCR to examine the onset of *Csx* expression at the earlier stage. *Csx* transcript was expressed from 7.5-day p.c. head-fold presomite stage but was not detectable in 5.5- and 6.5-day embryos (Fig. 3*c*).

To localize *Csx* expression sites in the embryo, *in situ* hybridization analysis was performed (Fig. 4). In 8.5-day embryos, *Csx* was already expressed at high levels in the myocardial layer of the heart (Fig. 4 *a* and *b*). In 9- and 10-day embryos, *Csx* was abundantly expressed all over the atrial and ventricular myocardium (Fig. 4 *c–h*). No signals were observed in other areas, including endocardium, aorta, brain, thyroid, somites, stomach, gut, spinal cord, or tail bud. These *in situ* hybridization data suggest that *Csx* is expressed only in the developing myocardium but not in skeletal or smooth muscle or any other tissues. Cardiocyte-specific expression of *Csx* was also observed in later embryonic stages, such as 18-day embryos (Fig. 4*j*). There were no specific signals in the heart when a sense probe was used for hybridization (Fig. 4*k*).

Expression of *Csx* in Differentiated ES Cells. We examined *Csx* expression in an *in vitro* cardiac differentiation system using the mouse ES cells. When ES cells are cultured in suspension state without the feeder layer or the leukemia inhibitory factor, they aggregate, develop into EB, and express cardiac-specific genes such as αMHC (15). As shown in Fig. 5, *Csx* expression was not observed in undifferentiated ES cells. However, when ES cells were cultured in suspension for 5 days, expression of *Csx* was observed. The expression of a cardiac-specific gene, the cardiac αMHC, was observed from 8 days after suspension culture, which was 3 days later than the initiation of *Csx* expression.

DISCUSSION

The results presented here demonstrate that *Csx* is a cardiac-specific homeobox cDNA and has sequence similarity to the *Drosophila NK* family homeodomain. *Drosophila* genes belonging to the *NK* family have homeodomains divergent from the *Antp, En, Prd*, or Even-skipped classes and seem to specify cell fates in specific tissues. Expression of the *S59* (*NK-1*) gene is restricted to the somitic mesoderm, subsets of the central nervous system, and a small region of the midgut (18). A rat homeobox gene, *TTF-1*, is closely related to the *Drosophila NK-2* and is expressed in the thyroid and lung anlage and in restricted neuroblast populations (19). Two additional *NK-2*-related genes are expressed in restricted

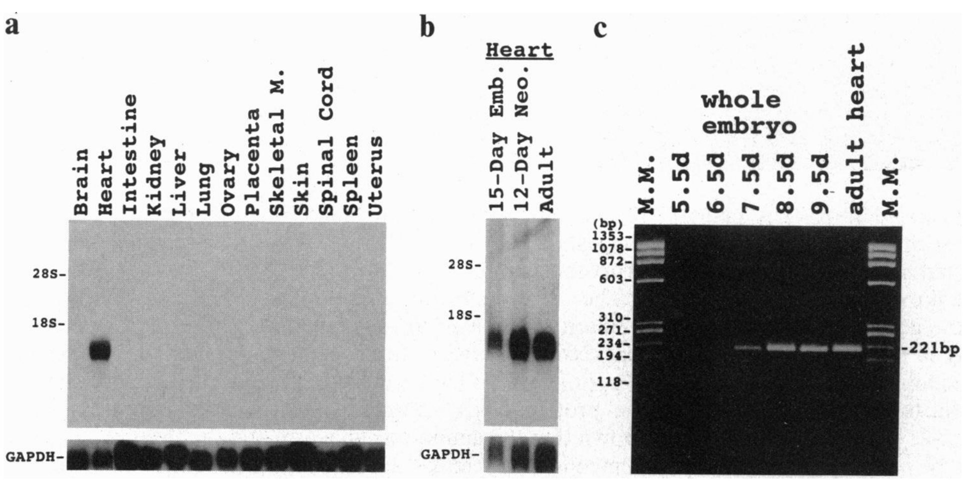

FIG. 3. Northern blot and RT-PCR analysis of *Csx* mRNA. (*a*) Twenty micrograms of poly(A)$^+$ RNA isolated from various adult tissues was loaded to each lane and hybridized with the 5′ part (nucleotides 1–305) of *Csx* cDNA. (*b*) Poly(A)$^+$ RNAs isolated from hearts of 15-day embryo (10 μg), 2-day neonatal (20 μg), and adult mice (20 μg) were analyzed by Northern blot. Hybridizations with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) are shown as internal controls. (*c*) Total RNA was isolated from five whole embryos [5.5–9.5 days postcoitum (p.c.)] and from an adult heart. PCR was performed for 35 cycles as described in the text. M.M., molecular marker of ϕX174 DNA digested with *Hae* III. The expected size of a PCR product is 221 bp.

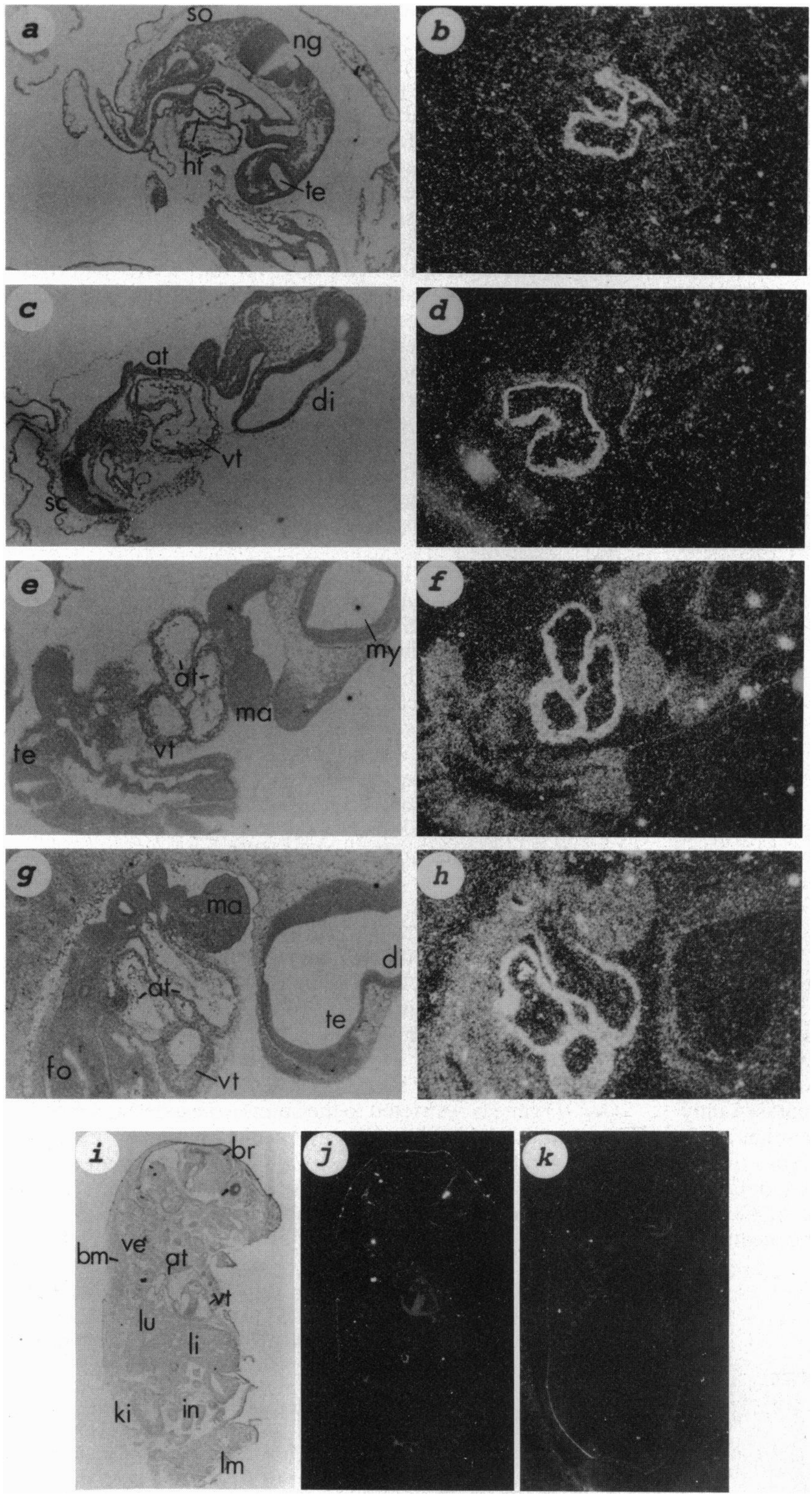

FIG. 4. Expression pattern of *Csx* during heart development as observed by *in situ* hybridization. Each bright-field picture (*a*, *c*, *e*, *g*, and *i*) corresponds to each dark-field picture (*b*, *d*, *f*, *h*, and *j*). (*a* and *b*) Saggital sections of an 8.5-day p.c. fetus hybridized with a *Csx*-specific antisense oligonucleotide probe. Specific signals are observed in the developing heart (ht). (*c–f*) Saggital (*c* and *d*) and frontal (*e* and *f*) sections of a 9-day p.c. fetus. Distinctive labeling of the atrium (at) and the ventricle (vt) is observed with a *Csx* probe. There are no signals in other areas, including endocardium. (*g* and *h*) Saggital sections of a 10-day p.c. fetal middle portion. Expression is observed only in the myocardium of the atrium (at) and the ventricle (vt). (*i–k*) Saggital section of an 18-day embryo. The labeling is observed only in the heart with the antisense probe (*j*) but not with the sense probe (*k*). Signals in the skin and the edge of vertebrate columns are likely to represent nonspecific "edge effects," because similar signals were seen using the sense probe (*k*) and other unrelated probes (not shown). at, Atrium; bm, back muscle; br, brain; di, diencephalon; fo, foregut; ht, heart; in, intestine; ki, kidney; li, liver; lm, leg muscle; lu, lung; ma, mandibular arch; my, myeloceol; ng, neural groove; sc, spinal cord; so, somite; te, teloceol; ve, vertebral column; vt, ventricle. (*a–h*, ×25; *i–k*, ×2.3.)

regions of the mouse forebrain (17). *msh-2* (*NK-4*) is expressed in all mesoderm cells in the segmental parts of the embryo during germband elongation, but soon afterwards its expression becomes restricted to the dorsal mesoderm and the heart primordium (12). Beyond the germband-extended stage, *msh-2* (*NK-4*) is expressed only in the heart (12). The mutant having deletion in the chromosomal region including the *msh-2* (*NK-4*) locus does not form the heart (12). Furthermore, forced expression of *msh-2* cDNA under the control of the heat shock promoter in the *msh-2* null mutants rescued formation of the heart precursor cells (20). This suggests that *msh-2* (*NK-4*) plays a critical role in *Drosophila* heart formation.

The cardiac-specific expression of *Csx* and its sequence similarity to *msh-2* (*NK-4*) homeodomain raises a possibility that *Csx* may play a critical role in mouse heart development, like *msh-2* (*NK-4*) in *Drosophila*. Recent structural studies on the protein–DNA complex of three homeodomains have shown that the amino-terminal arm and the third helix of the homeodomains make specific contacts with bases in the

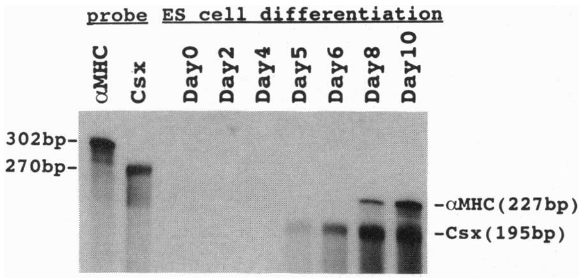

FIG. 5. Expression of *Csx* in ES cells during *in vitro* differentiation. Total RNA was extracted from ES cells that were differentiated in suspension culture for the indicated days. RNase protection assay was performed using *Csx* and αMHC antisense riboprobes. The probe lengths were 270 nucleotides (*Csx*) and 302 nucleotides (αMHC). The protected fragments of the expected sizes (*Csx*, 195 bp; αMHC, 227 bp) were observed in RNA samples from 5-day-old EBs for *Csx* and from 8-day-old EBs for αMHC.

minor and major grooves of DNA, respectively (21–23). Intriguingly, *Csx* and *msh-2* (*NK-4*) have identical residues in the homeodomain that are predicted to contact DNA (21–23). It would be interesting to determine whether the same target genes might be regulated by *Csx* and *msh-2* (*NK-4*) during heart development.

Csx is a highly conserved gene in the vertebrates. *Csx*-like genes exist in mammalian, avian, and amphibian species. Very recently we have isolated a *Xenopus* homolog of *Csx*. The homeodomain sequence of *Xenopus Csx* is 93% identical to that of the murine *Csx* (unpublished data). Interestingly, the *Xenopus Csx* gene is expressed only in the heart, like the murine *Csx* and *Drosophila msh-2*. These data suggest a potential evolutionary conservation in the molecular control of cardiogenesis in the animal kingdom, as has been suggested for neurogenesis (24).

Csx is specifically expressed in the heart from 8.5-day embryos through adult. No expression was recognized in other tissues by Northern blot and *in situ* hybridization analysis. More than 50 homeobox-containing genes have been isolated in mammals so far as is; however, there are few homeobox genes that are expressed in the rostral part of the mouse embryo (4, 25). There has been no report of homeobox-containing genes that are expressed in the developing myocardium, the most anterior structure in early mouse embryos. Expression of *Csx* starts from 7.5 days p.c., though we have not localized the site of expression at this stage. At around 7.25–7.5 days p.c., the intraembryonic mesoderm splits to form the intraembryonic coeloma. The dorsal lining of somatic mesoderm forms a squamous mesothelial epithelium and the ventral splanchnic mesoderm differentiates into the cuboidal epithelium, the cardiogenic plate, or promyocardium (26). In 8- to 8.5-day embryos, when the mesoderm cells increase rapidly and begin to differentiate into the heart, notochord, and somites, the embryo body is very coiled and the heart is formed more anterior to the forebrain (27). At this stage, *Csx* was already expressed at high levels in the heart. At ≈9 days p.c., there is a common ventricle and atrium, and the dorsal aorta are fused. Between days 9 and 10 p.c., blood circulation in the visceral yolk sac begins as the heart starts to beat (27). In 9- and 10-day embryos, *Csx* was abundantly expressed all over the atrial and ventricular myocardium. Expression of *Hox* genes is generally not restricted in one tissue but are observed in broad regions of the body (4, 25). A mouse homeobox-containing gene, *Hox7*, is expressed in the neural fold, cephalic neural crest, and developing limb buds as well as in the developing valves of the embryonic heart (28). In contrast, *Csx* is expressed very early in the myocardial layer but not in the endocardial cushion. The highly cell-type-specific expression of *Csx* during all developmental stages raises the possibility that *Csx* may play an important role in cardiomyocyte differentiation and maintenance of cardiac muscle phenotype.

Expression of *Csx* was also observed in differentiating ES cells before that of the cardiac-specific gene. When ES cells are differentiated *in vitro*, they develop into EB, express cardiac-specific genes, and start to contract spontaneously from ≈8 days after suspension culture (15). This indicates that during *in vitro* differentiation of ES cells, many, if not all, aspects of cardiogenesis are recapitulated. In the *in vitro* differentiation system of ES cells, cardiac-specific genes such as αMHC, myosin light chain 2, and atrial natriuretic factor are expressed from 8 days (15, 29), whereas *Csx* expression starts from 5 days after suspension culture. The αMHC transcript is expressed as early as 8 days p.c. in the mouse embryo *in vivo* (30), and *Csx* expression is detectable from 7.5 days p.c. *in vivo*. Taken together, *Csx* expression seems to precede that of many cardiac-specific genes *in vitro* and *in vivo*. Although at present we do not know the physiological function of *Csx*, the expression pattern, developmental kinetics, and evolutionary conservation of *Csx* suggest that it may play important roles in cardiac differentiation in vertebrates.

We thank Rolf Bodmer for discussion and communication of unpublished results; Lothar Jahn and Martin Schalling for their help and advice on *in situ* analysis; Yuh Nung Jan, Robert D. Rosenberg, and William Grossman for encouragement; and Jon Goldman for editorial assistance. This work was supported in part by the Jacob D. Indursky Memorial Fund. I.K. is supported by the Medical Foundation and the Bayer/Japan Heart Foundation Scholarship. S.I. is an Established Investigator of the American Heart Association.

1. Olson, E. N. (1993) *Circ. Res.* **72,** 1–6.
2. St. Johnston, O. & Nusslein-Volhard, C. (1992) *Cell* **68,** 201–219.
3. Akam, M. (1987) *Development* **101,** 1–22.
4. McGinnis, W. & Krumlauf, R. (1992) *Cell* **68,** 283–302.
5. Hayashi, S. & Scott, M. (1990) *Cell* **63,** 883–894.
6. Graham, A., Papalopulu, N. & Krumlauf, R. (1989) *Cell* **57,** 367–378.
7. Kessel, M., Balling, R. & Gruss, P. (1990) *Cell* **61,** 301–308.
8. Chisaka, O. & Capecchi, M. (1991) *Nature (London)* **350,** 473–479.
9. Lufkin, T., Dierich, A., LeMeur, M., Mark, M. & Chambon, P. (1991) *Cell* **66,** 1105–1119.
10. He, X., Treacy, M. N., Simmons, D. M., Ingraham, H. A., Swanson, L. W. & Rosenfeld, M. (1989) *Nature (London)* **340,** 35–42.
11. Castrillo, J.-L., Theill, L. E. & Karin, M. (1991) *Science* **253,** 197–199.
12. Bodmer, R., Jan, L. Y. & Jan, Y. N. (1990) *Development* **110,** 661–669.
13. Fahner, K., Hogan, B. L. M. & Flavell, R. A. (1987) *EMBO J.* **6,** 1265–1271.
14. Komuro, I., Schalling, M., Jahn, L., Bodmer, R., Jenkins, N. A., Copeland, N. G. & Izumo, S. (1993) *EMBO J.* **12,** 1387–1401.
15. Sanchez, A., Jones, W. K., Gulick, J., Doetschman, T. & Robbins, J. (1991) *J. Biol. Chem.* **266,** 22419–22426.
16. Kim, Y. & Nirenberg, M. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7716–7720.
17. Price, M., Lazzaro, D., Pohl, T., Mattei, M.-G., Ruther, U., Olivo, J.-C., Duboule, D. & Di Lauro, R. (1992) *Neuron* **8,** 241–255.
18. Dohrmann, C., Azpiazu, N. & Frasch, M. (1990) *Genes Dev.* **4,** 2098–2111.
19. Lazzaro, D., Price, M., De Felice, M. & Di Lauro, R. (1991) *Development* **113,** 1093–1104.
20. Bodmer, R. (1993) *Development* **118,** 719–729.
21. Kissinger, C. R., Liu, B., Martin-Blanco, E., Kornberg, T. B. & Pabo, C. O. (1990) *Cell* **63,** 579–590.
22. Otting, G., Qian, Y. Q., Billeter, M., Muller, M., Affolter, M., Gehring, W. J. & Wuthrich, K. (1990) *EMBO J.* **9,** 3085–3092.
23. Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D. & Pabo, C. O. (1991) *Cell* **67,** 517–528.
24. Jan, Y. N. & Jan, L. Y. (1990) *Trends NeuroSci.* **13,** 493–498.
25. Holland, P. W. H. & Hogan, B. L. M. (1988) *Genes Dev.* **2,** 773–782.
26. DeRuiter, M. C., Poelmann, R. E., Van der Plas-de Vries, I., Mentink, M. M. T. & Gittenberger-de Groot, A. C. (1992) *Anat. Embryol.* **185,** 461–473.
27. Rugh, R. (1990) *The Mouse: Its Reproduction and Development* (Oxford Univ. Press, Oxford).
28. Robert, B., Sasoon, D., Jacq, B., Gehring, W. & Buckingham, M. (1989) *EMBO J.* **8,** 91–100.
29. Miller-Hance, W. C., Evans, S. M., Shubeita, H. E., Robbins, J. & Chien, K. R. (1993) *J. Cell. Biochem. Suppl.* **17D,** 200 (abstr.).
30. Lyons, G. E., Schiaffino, S., Sassoon, D., Barton, P. & Buckingham, M. (1990) *J. Cell Biol.* **111,** 2427–2436.