

Darmin is a novel secreted protein expressed during endoderm development in *Xenopus*

Edgar M. Pera, Sarah L. Martinez, John J. Flanagan, Mariel Brechner,
Oliver Wessely, E.M. De Robertis*

Howard Hughes Medical Institute and Department of Biological Chemistry, University of California, Los Angeles, CA 90095-1662, USA

Received 4 November 2002; received in revised form 19 December 2002; accepted 7 January 2003

Abstract

Endoderm development is an area of intense interest in developmental biology, but progress has been hampered by the lack of specific markers for differentiated endodermal cells. In an unbiased secretion cloning screen of *Xenopus* gastrula embryos we isolated a novel gene, designated *Darmin*. *Darmin* encodes a secreted protein of 56 kDa containing a peptidase M20 domain characteristic of the glutamate carboxypeptidase group of zinc metalloproteases. We also identified homologous *Darmin* genes in other eukaryotes and in prokaryotes suggesting that *Darmin* is the founding member of a family of evolutionarily conserved proteins. *Xenopus Darmin* showed zygotic expression in the early endoderm and later became restricted to the midgut. By secretion cloning of *Xenopus* cleavage-stage embryos we isolated another novel protein, designated *Darmin*-related (*Darmin-r*) due to its sequence similarity with *Darmin*. *Darmin-r* was maternally expressed and showed at later stages expression in the lens and pronephric glomus. The endoderm-specific expression of *Darmin* makes this gene a useful marker for the study of endoderm development.

© 2003 Elsevier Science B.V. All rights reserved.

Keywords: *Darmin*; *Darmin*-related; M20 peptidase; Endoderm; Gut; Lens; Kidney; Glutamate carboxypeptidase; Zinc metalloprotease; Endodermin; Secretion cloning; *Xenopus*

1. Results and discussion

1.1. *Darmin* is a novel secreted protein of the glutamate carboxypeptidase family

In metazoans, the endoderm constitutes the innermost germ layer and differentiates into the epithelium of the digestive tract and associated organs (Wells and Melton, 1999; Stainier, 2002; Shivdasani, 2002). Molecular studies in *Xenopus* have identified multiple factors implicated in endoderm formation. These include transcription factors, such as VegT, Mix1, Mixer, Xsox17, and XGATA5 (Henry and Melton, 1998; Lemaire et al., 1998; Hudson et al., 1997; Zhang and King, 1998; Weber et al., 2000), the RNA-binding molecule Bicardal-C (Wessely and De Robertis, 2000), and secreted proteins like Nodal-related TGF- β ligands, Chordin, Noggin, and Cerberus (Henry et al., 1996; Sasai et al., 1996; Bouwmeester et al., 1996). Despite intense interest, our knowledge of how the endoderm develops is still limited.

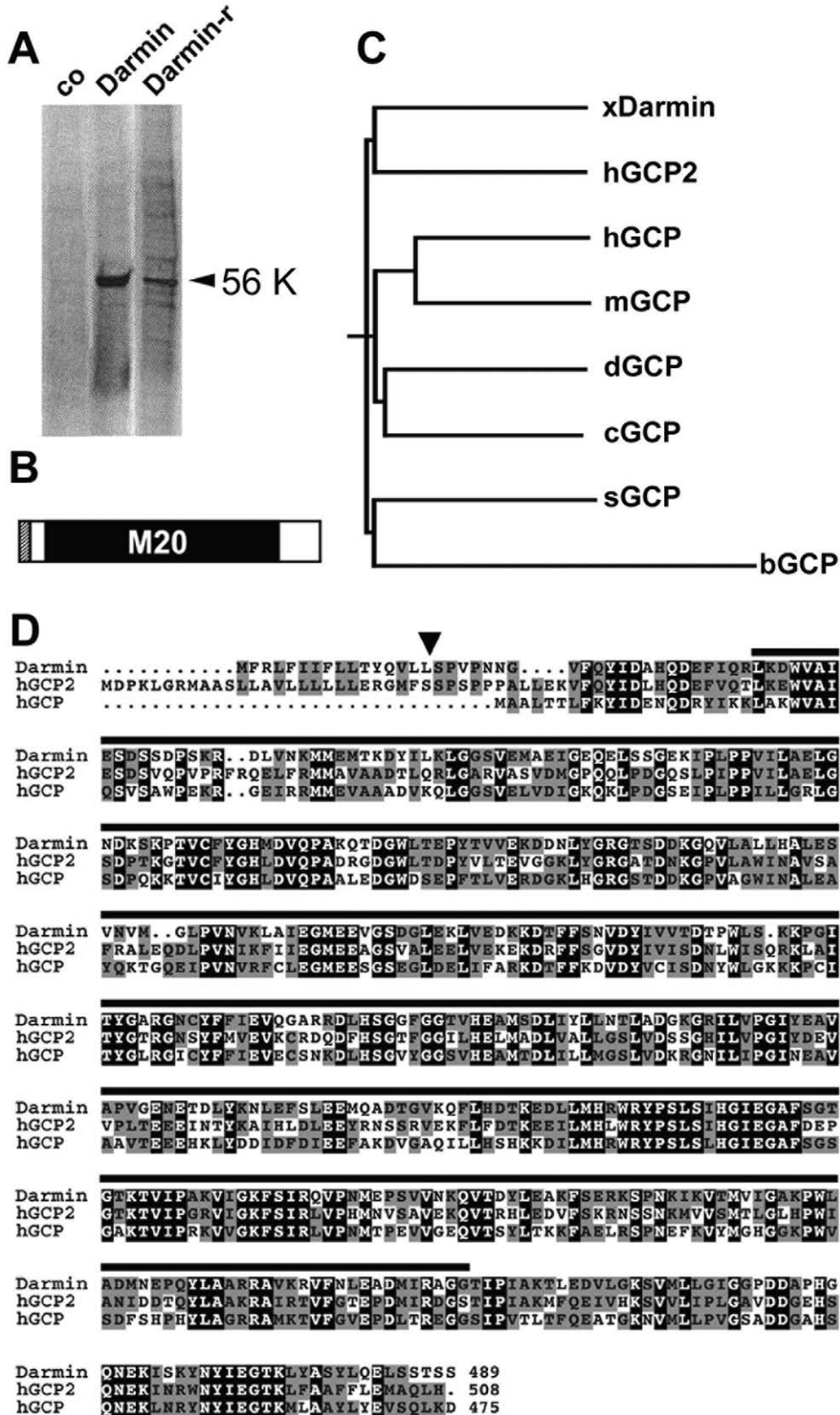
This is partly due to a lack of genes uniquely expressed in this germ layer. Therefore, early markers specific for differentiated endoderm are needed. We have recently reported a method, designated secretion cloning, that allows the direct identification of secreted proteins (Pera and De Robertis, 2000). Using an expression cDNA library from *Xenopus* midgastrula embryos, we identified a novel gene, which we describe here and call *Darmin* ('Darm' means gut in German; GenBank accession number AY166869). It encodes a secreted metalloproteinase that is exclusively expressed in the endoderm and early midgut.

Xenopus Darmin was first identified as a 56 kDa protein in the supernatant of transfected human 293T cells after metabolic labeling with ^{35}S -methionine and ^{35}S -cysteine (Fig. 1A). The full-length *Darmin* cDNA clone encoded a 489 amino acid protein containing a putative hydrophobic signal peptide at the N-terminus and a predicted peptidase M20 domain (pfam 01546; NCBI Conserved Domain Search; Figs. 1B,D). M20 domains are found in glutamate carboxypeptidases (Rawlings and Barrett, 1995; Sherwood and Melton, 1998). These proteins belong to the group of zinc metallopeptidases that hydrolyze peptide bonds and

* Corresponding author. Tel.: +1-310-206-1401; fax: +1-310-206-2008.
E-mail address: derobert@hhmi.ucla.edu (E.M. De Robertis).

depend on zinc for their activity. Glutamate carboxypeptidases release C-terminal glutamate residues from many N-acyl moieties, including peptidyl, aminoacyl, benzoyl, benzoyloxycarbonyl, foyl, and pteroyl groups.

Xenopus Darmin is most closely related to a human hypothetical protein called glutamate carboxypeptidase-like protein 2 (hGCP2, Figs. 1C,D). The two proteins share 56% amino acid identity. hGCP2 has an N-terminal signal



peptide, which is predicted to be cleaved at the same position as Darmin (indicated by the arrowhead in Fig. 1D). Other hypothetical proteins related to Darmin include human glutamate carboxypeptidase (hGCP, 55% amino acid identity) and homologous hypothetical proteins in mouse, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Schizosaccharomyces pombe*, and *Bacillus halodurans* (55, 55, 48, 42 and 32% amino acid identity, respectively). All these sequences contain a predicted peptidase M20 domain. Darmin is weakly homologous to human N-acetylated- α -linked acidic dipeptidase (NAALADase; 19% amino acid identity; Rawlings and Barrett, 1997). NAALADase is a well-studied glutamate carboxypeptidase that hydrolyzes the neurotransmitter N-acetyl-aspartyl-glutamate into N-acetyl-aspartate and glutamate. However, in contrast to Darmin, NAALADase is a transmembrane protein and belongs to the M28 peptidase family of catalytic zinc metallopeptidases. We conclude that Darmin is a member of a novel family of evolutionarily conserved glutamate carboxypeptidases.

1.2. Expression of *Xenopus* Darmin during early development

The expression pattern of *Darmin* was determined by whole-mount in situ hybridization in *Xenopus* embryos (Fig. 2). We did not detect maternal transcripts at the four-cell stage (Fig. 2A). At the onset of gastrulation, *Darmin* transcripts appeared on the dorsal side of the vegetal cell mass (Fig. 2B). Nuclear staining was evident in yolky endoderm precursors that lined the invaginating dorsal mesoderm in Spemann's organizer (Fig. 2C). As gastrulation proceeded, *Darmin* expression soon expanded throughout the endoderm (Fig. 2D). In late gastrula embryos, *Darmin* strongly demarcated the endoderm, but excluded the anterior region fated to become the foregut (Fig. 2E, see arrow indicating the liver diverticulum). In late neurula embryos, *Darmin* was abundantly expressed in the ventral and lateral walls of the gut (Figs. 2F–H). By tail bud stage, the gut was densely stained (Fig. 2I). Expression in the intestinal tract persisted until the swimming tadpole stage (Figs. 2J,K). We performed in situ hybridization on guts

dissected from 4 to 7 day-old embryos as described in Chalmers and Slack (2000), but did not detect any *Darmin* transcripts at this stage (data not shown). We conclude that *Darmin* is a zygotic gene exclusively expressed in the endoderm during the early phases of gut development.

1.3. Darmin is a midgut-specific marker

We compared the expression of *Darmin* to that of *Endodermin* (*Edd*; Figs. 3A,B). *Edd* is a pan-endodermal gene marker that encodes a secreted proteinase inhibitor of the α 2-macroglobulin family (Sasai et al., 1996). In tail bud embryos, *Edd* demarcated the entire gut (Fig. 3A), whereas *Darmin* transcripts were excluded from the fore- and hindgut (Fig. 3B). This confirms that *Darmin* marks the midgut.

1.4. Expression of *Xenopus* Darmin-r

By secretion cloning using a *Xenopus* 64-cell stage cDNA library, we identified another novel protein, which we refer to as Darmin-related (*Darmin-r*). Like *Darmin*, this protein had an apparent molecular weight of 56 kDa (Fig. 1A). Partial sequencing of the *Darmin-r* cDNA (GenBank accession number BU993920) revealed considerable homology to *Darmin* (data not shown). Whole-mount in situ hybridization revealed a very different expression pattern for *Darmin-r* in *Xenopus* embryos (Figs. 3C,D). *Darmin-r* was not expressed in the endoderm. High levels of maternal *Darmin-r* transcripts were abundant at the four-cell stage (Fig. 3C). Zygotic *Darmin-r* expression was not detectable until the tail bud stage (data not shown). By stage 32, *Darmin-r* was weakly expressed in the lens and the pronephric glomus (Fig. 3D). To investigate the activity of *Darmin* and *Darmin-r*, we microinjected mRNA into early *Xenopus* embryos, but did not observe any obvious phenotypic effects (data not shown).

1.5. Conclusions

We have isolated two novel members of the glutamate carboxypeptidase family of zinc metalloproteases. *Darmin*

Fig. 1. Darmin is a novel secreted glutamate carboxypeptidase. (A) Supernatant culture medium of human 293T cells transfected with *Xenopus* *Darmin* cDNA (*Darmin*) or *Xenopus* *Darmin-related* cDNA (*Darmin-r*). Transfected cells were labeled with ^{35}S -methionine and ^{35}S -cysteine, and supernatants analyzed by SDS-PAGE and autoradiography. Note that *Darmin* and *Darmin-r* are secreted as 56 kDa proteins. (B) Schematic representation of *Darmin*. The N-terminal signal peptide (gray box) and the conserved M20 peptidase domain (black box) are indicated. (C) Evolutionary relationships of *Darmin* and other members of the glutamate carboxypeptidase family. Proteins displaying the greatest sequence similarity group together, and branch lengths are proportional to distance (alignments and phylogenetic tree generated by T-COFFEE, <http://www.ch.embnet.org/software/TCoffee.html>, and PhyloDendron, <http://iubio.bio.indiana.edu/treeapp/treeprint-form.html>). GenBank accession numbers are as follows: hGCP-2, human glutamate carboxypeptidase-like protein 2 (CAD10388); hGCP, human glutamate carboxypeptidase (CAC69883); mGCP, mouse hypothetical protein (BAB22991); dGCP, *D. melanogaster* hypothetical protein (AAM48397); cGCP, *C. elegans* hypothetical protein (NP506610); sGCP, *S. pombe* hypothetical protein (NP595077); bGCP, *B. halodurans* hypothetical protein (NP244742). (D) Amino acid sequence alignments of *Xenopus* *Darmin* (*Darmin*), human glutamate carboxypeptidase-like protein 2 (hGCP2) and human glutamate carboxypeptidase (hGCP). Identical amino acid residues are shaded in black, and similar residues in gray. The signal peptide cleavage site (triangle) and the peptidase M20 domain (bar) are indicated.

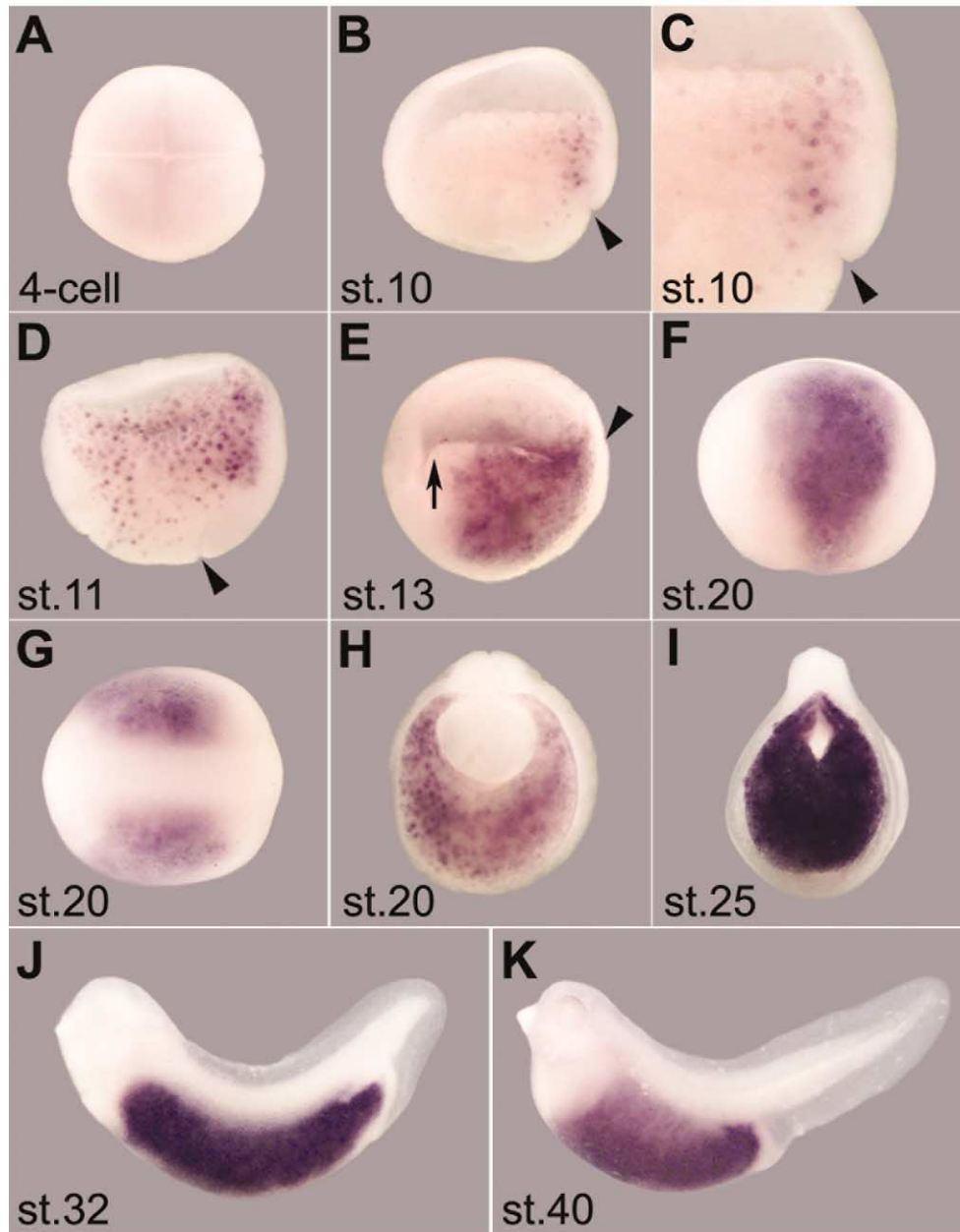


Fig. 2. *Darmin* is an endoderm-specific marker. Whole-mount in situ hybridization was performed on intact (A, F, G, J, K) or hemisectioned *Xenopus* embryos (B–E, H, I). (A) Animal view of 4-cell stage embryo; note the lack of maternal transcripts. (B) Internal view of a hemisectioned early gastrula embryo (stage 10); *Darmin* expression starts in the dorsal vegetal cell mass; the arrowhead indicates the dorsal blastopore lip in this and subsequent panels. (C) Close-up of embryo in (B); note nuclear staining in the yolk endoderm adjacent to invaginating mesoderm of Spemann's organizer. (D) At mid-gastrula (stage 11), *Darmin*-labeled cells are found throughout the endoderm. (E) Internal view of late gastrula embryo (stage 13) showing increased transcript levels in the endoderm excluding the anterior region; the arrow marks the liver diverticulum, a derivative of the foregut. (F, G) Lateral and dorsal views of late neurula (stage 20) displaying *Darmin* expression in the midgut. (H) Internal view showing signals in the ventral and lateral walls of the midgut. (I) At early tail bud stage (stage 25), staining in the midgut becomes very abundant. (J, K) Lateral views of advanced tail bud embryos (stages 32 and 40); *Darmin* expression remains strong in the midgut.

was transiently expressed in the endoderm and developing midgut at the time of yolk resorption. Future experiments will address whether proteinase inhibitors such as Edd regulate the catalytic activity of *Darmin*. Due to its specific and strong expression in the early midgut, *Darmin* provides a useful marker for the study of endoderm development in vertebrate embryos.

2. Experimental procedures

Embryos were staged according to [Nieuwkoop and Faber \(1994\)](#). An amplified cDNA library from *Xenopus* UV-ventralized stage 11 embryos in pCDNA3 (kind gift of Dr Leonard Zon, Harvard Medical School, Boston) and an unamplified library from *Xenopus* stage 6 embryos in pCS2

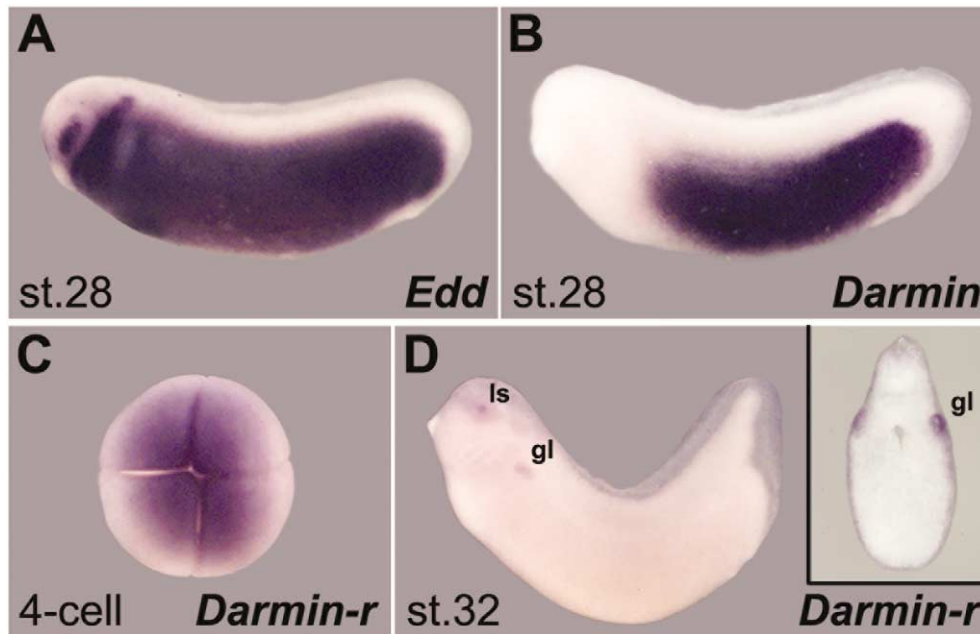


Fig. 3. Expression of *Endodermin*, *Darmin* and *Darmin-related*. Embryos are shown in lateral (A, B, D) and animal views (C). (A) Tail bud embryo (stage 28); note expression of *Endodermin* (*Edd*) in the fore-, mid-, and hindgut. (B) Sibling embryo showing *Darmin* expression restricted to the midgut. (C) Abundant maternal transcripts of *Darmin-related* (*Darmin-r*) in 4-cell stage embryo. (D) Advanced tail bud embryo (stage 32) displaying weak expression of *Darmin-r* in the lens (ls) and pronephric glomus (gl). Transversal section in inset shows transcripts in the glomera.

were used. Secretion cloning was performed as described by Pera and De Robertis (2000). Briefly, 293T cells were transfected with pools of cDNA clones and 1.5 days later labeled with ^{35}S -methionine and ^{35}S -cysteine under serum-free conditions. Secreted proteins were identified by SDS-PAGE and autoradiography, and positive clones individualized in a second transfection step by sib-selection. For antisense RNA, pCDNA3-*Darmin* and pCS2-*Darmin-r* were linearized with *Bam*HI and transcribed with SP6 and T7 RNA polymerase, respectively. The probe for *Edd* was synthesized as described (Sasai et al., 1996). Whole-mount in situ hybridizations were performed as described (<http://www.hhmi.ucla.edu/derobertis/>).

Acknowledgements

We thank Drs C. Coffinier, J. Kim, and Z. Pettway Unno for critical reading of the manuscript. We are grateful to A. Cuellar and U. Tran for technical assistance. E.M.P and O.W. were HFSPO fellows. This work was supported by the National Institutes of Health (R37 HD21502-16). E. M. D. R. is an Investigator of the Howard Hughes Medical Institute.

References

Bouwmeester, T., Kim, S.-H., Sasai, Y., Lu, B., De Robertis, E.M., 1996. Cerberus is a head-inducing secreted factor expressed in the anterior endoderm of Spemann's organizer. *Nature* 382, 595–601.

Chalmers, A.D., Slack, J.M.W., 2000. The *Xenopus* tadpole gut: fate maps and morphogenetic movements. *Development* 127, 381–392.

Henry, G.L., Brivanlou, I.H., Kessler, D.S., Hemmati-Brivanlou, A., Melton, D.A., 1996. TGF-beta signals and a pattern in *Xenopus laevis* endodermal development. *Development* 122, 1007–1015.

Henry, G.L., Melton, D.A., 1998. *Mixer*, a homeobox gene required for endoderm development. *Science* 281, 91–96.

Hudson, C., Clements, D., Friday, R.V., Stott, D., Woodland, H.R., 1997. *Xsox17- α* and $-\beta$ mediate endoderm formation in *Xenopus*. *Cell* 91, 397–405.

Lemaire, P., Darras, S., Caillol, D., Kodjabachian, L., 1998. A role for the vegetally expressed *Xenopus* gene *Mix.1* in endoderm formation and in the restriction of mesoderm to the marginal zone. *Development* 125, 2371–2380.

Nieuwkoop, P.D., Faber, J., 1994. Normal Table of *Xenopus laevis*, Garland Publishing, New York, NY.

Pera, E.M., De Robertis, E.M., 2000. A direct screen for secreted proteins in *Xenopus* embryos identifies distinct activities for the Wnt antagonists Crescent and Frzb-1. *Mech. Dev.* 96, 183–195.

Rawlings, N.D., Barrett, A.J., 1995. Evolutionary families of metalloproteases. *Methods Enzymol.* 248, 183–228.

Rawlings, N.D., Barrett, A.J., 1997. Structure of membrane glutamate carboxypeptidase. *Biochim. Biophys. Acta* 1339, 247–252.

Sasai, Y., Lu, B., Piccolo, S., De Robertis, E.M., 1996. Endoderm induction by the organizer-secreted factors chordin and noggin in *Xenopus* animal caps. *EMBO J.* 15 (17), 4547–4555.

Sherwood, R.F., Melton, R.G., 1998. Glutamate carboxypeptidases. In: Barrett, A.J., Rawlings, N.D., Woessner, J.F. (Eds.), *Handbook of Proteolytic Enzymes*, Academic Press, London, pp. 1416–1420.

Shivdasani, R.A., 2002. Molecular regulation of vertebrate early endoderm development. *Rev. Dev. Biol.* 249, 191–203.

Stainier, D.Y.R., 2002. A glimpse into the molecular entrails of endoderm formation. *Rev. Genes Dev.* 16, 893–907.

Weber, H., Symes, C.E., Walmsley, M.E., Rodaway, A.R., Patient, R.K., 2000. A role for GATA5 in *Xenopus* endoderm specification. *Development* 127, 4345–4360.

Wells, J.M., Melton, D.A., 1999. Vertebrate endoderm development. *Annu. Rev. Cell Dev. Biol.* 15, 393–410.

Wessely, O., De Robertis, E.M., 2000. The *Xenopus* homologue of

Bicaudal-C is a localized maternal mRNA that can induce endoderm formation. *Development* 127, 2053–2062.

Zhang, J., King, M.L., 1998. The role of maternal VegT in establishing the primary germ layers in *Xenopus* embryos. *Cell* 94, 515–524.