The pro-domain of the zebrafish Nodal-related protein Cyclops regulates its signaling activities

Jing Tian¹, Birgit Andrée², C. Michael Jones² and Karuna Sampath¹,3,4,*

Nodal proteins are secreted signaling factors of the transforming growth factor β (TGFβ) family with essential roles in embryonic development in vertebrates. Mutations affecting the Nodal factors have severe consequences in mammals and fish. Furthermore, increased Nodal levels have been associated with melanoma tumor progression. Like other TGFβ-related proteins, Nodal factors consist of a pro-domain and a mature domain. The pro-domain of mouse Nodal protein stabilizes its precursor. However, the mechanisms by which the pro-domains exert their activities are unknown. Here, we characterize the zebrafish Nodal-related factor Cyclops (Cyc) and find unexpected functions for the pro-domain in regulating Cyc activity. We identified a lysosome-targeting region in the Cyc pro-domain that destabilizes the precursor and restricts Cyc activity, revealing the molecular basis for the short-range signaling activities of Cyc. We show that both the pro- and mature-domains of Cyc regulate its stability. We also characterize a mutation in the pro-domain of human Nodal (hNODAL) that underlies congenital heterotaxia. Heterologous expression of mutant hNODAL increases expression of Nodal-response genes. Our studies reveal unexpected roles for the pro-domain of the Nodal factors and provide a possible mechanism for familial heterotaxia.

KEY WORDS: Left-right asymmetry, Nodal signaling, Squint, Mature domain, Pro-domain, Zebrafish, Cyclops

INTRODUCTION

Intercellular signaling plays crucial roles not only in establishing the body plan and forming the germ layers during embryonic development, but also in cancer progression. For example, formation and patterning of the mesoderm, endoderm and ventral neural tube in vertebrates requires the function of the Nodal-related secreted signaling factors (Shen, 2007). Nodal and Nodal-related factors belong to the transforming growth factor β (TGFβ) family. Mutations affecting the Nodal factors have severe developmental consequences in mammals and fish (Shen, 2007). In addition to being a potent embryonic morphogen, Nodal has recently been shown to be a key regulator of melanoma tumorigenicity and progression (Topczewska et al., 2006). Nodal expression was found to be positively associated with melanoma tumor progression, and aggressive melanoma cell lines express more Nodal protein than lines that are not aggressive (Topczewska et al., 2006). Nodal expression has also been found in testicular cancer xenografts and in aggressive breast cancer cell lines (Adkins et al., 2003). Thus, Nodal protein levels can be an indicator of cancer progression in these cell types.

Nodal factors bind and signal through heteromeric complexes of type I and type II transmembrane serine-threonine kinase receptors (Reissmann et al., 2001; Yan et al., 2002; Yeo and Whitman, 2001). Studies in frogs, fish and mice have suggested that Nodal signaling via the type I receptor Alk4 and the type II receptors ActRIIA and ActRIIB leads to phosphorylation and nuclear accumulation of the downstream effectors Smad2 and/or Smad3 in a complex with Smad4 (Massague et al., 2005). The activated Smad complex interacts with DNA-binding proteins of the winged-helix and homeodomain families to regulate downstream target gene expression (Germain et al., 2000; Pogoda et al., 2000; Randall et al., 2004). Downstream targets of the Nodal pathway include Lefty, an inhibitor of Nodal signaling (Branford and Yost, 2002; Chen and Shen, 2004; Sopory et al., 1998).

In zebrafish, the Nodal-related factors Cyclops, Squint (Sqt) and Southpaw (Spw) can induce mesoderm gene expression in ectopic overexpression assays. Analysis of mutations that affect cyc and sqt has shown that these genes have overlapping functions in the embryo during mesoderm induction (Erter et al., 1998; Feldman et al., 1998; Long et al., 2003; Rebagliati et al., 1998a; Rebagliati et al., 1998b; Sampath et al., 1998). Loss of cyc function results in severe deficiencies in the ventral neural tube and in cyclopecia, whereas sqt function is required for formation of the mesendoderm and dorsal structures (Erter et al., 1998; Feldman et al., 2000; Feldman et al., 1998; Gore et al., 2005; Hagos et al., 2007; Hatta et al., 1991; Rebagliati et al., 1998b; Sampath et al., 1998; Sirotkin et al., 2000; Tian et al., 2003). Spw regulates left-right asymmetry of the visceral organs during somitogenesis (Long et al., 2003). Subsequent work also revealed differences in the signaling activities of these proteins. Whereas Cyc acts at a short-range, Sqt behaves as a long-range morphogen (Chen and Schier, 2001).

In Xenopus, the Nodal-related factor Xnr2 and the bone morphogenetic protein Bmp4 have short-range effects, whereas Activin functions as a long-range morphogen (Jones et al., 1996a; McDowell et al., 1997; Williams et al., 2004). The biochemical basis for the differences in activities of the various TGFβ proteins has remained unclear. Although studies on the mouse and Xenopus BMPs have suggested that the pro-domain may regulate mature ligand stability and processing (Constand and Robertson, 1999; Cui et al., 2001; Deggins et al., 2004; Jones et al., 1996b; Sopory et al., 2006), a recent study in zebrafish has implicated structural differences in the mature domain as being responsible (Jing et al., 2006). Furthermore, processing and glycosylation of mouse Nodal also affect the stability and signaling range of this factor (Le Good et al., 2005).
Here, we characterize the zebrafish Nodal-related factor Cyclops (Cyc) in detail and compare it to human NODAL (hNODAL) and zebrafish Squint (Sqt). Surprisingly, unlike the human and mouse Nodal proteins or zebrafish Sqt, processed Cyc has no discernible activity by itself, and the pro-domain of Cyc is essential for its functions. By deletion and domain-swap analysis, we identified regions in the Cyc pro-domain that regulate its expression and stability. Interestingly, the Cyc pro-domain also plays an important role in regulating the signaling range of this factor. We identified a potential lysosome-targeting region in the Cyc pro-domain, and show that it targets the precursor for degradation. This region, when deleted, allows Cyc to function at a distance, similar to Sqt, thus revealing the molecular basis for the short-range signaling activities of Cyc. Finally, we characterize a mutation in a conserved residue in the pro-domain of hNODAL associated with congenital heterotaxia. Our results show that the pro-domain plays important roles in regulating the interactions and functions of the Nodal-related factors.

**MATERIALS AND METHODS**

**Generation of constructs**

All deletion, epitope-tagged and chimeric constructs were generated by PCR-based methods, followed by sub-cloning into the vector pCS2+, and the nucleotide sequence was confirmed (a list of primers and constructs can be provided on request). The hNODAL cDNA was amplified by RT-PCR from undifferentiated HES3 cells and cloned into the vector pCS2+. The Activin\(^{\text{pro}}\)-Cyc\(^{\text{mat}}\) chimera was generated by fusing the pro-domain of zebrafish Activin\(^B\) with its cleavage site, to the mature domain of Cyc. The fusion constructs pCS2 Cyc\(^{\text{pro}}\)-Sqt\(^{\text{mat}}\) and pCS2 Sqt\(^{\text{pro}}\)-Cyc\(^{\text{mat}}\) were generated by PCR, with the Cyc cleavage site RRRGR in Cyc\(^{\text{pro}}\)-Sqt\(^{\text{mat}}\), and the Sqt cleavage site RRHRR in Sqt\(^{\text{pro}}\)-Cyc\(^{\text{mat}}\). Amino acid substitutions in the Cyc pro-domain were generated by PCR-based methods.

**Zebrafish embryo injections**

All experimental procedures were carried out in accordance with the guidelines of the Institutional Animal Care Use Committee at Temasek Life Sciences Laboratory. Wild-type zebrafish were maintained at 28.5°C, and embryos were obtained by standard methods for use in injection experiments. Capped mRNA was synthesized from linearized plasmids using the mMessage mMACHINE kit (Ambion), and 0.5 pg, 1.25 pg, 5 pg or 25 pg aliquots of capped mRNA were injected into zebrafish embryos at the one-cell stage. The embryos were fixed in 4% paraformaldehyde (PFA) at 50% epiboly and processed for in situ hybridization to detect gsc or ntl expression (Tian et al., 2003). For detecting the range of signaling, 5 or 10 pg aliquots of in vitro synthesized capped mRNA encoding Sqt, Cyc, Cyc\(^{\text{Δ2}}\), Cyc\(^{\text{pro}}\)-Sqt\(^{\text{mat}}\), Sqt\(^{\text{pro}}\)-Cyc\(^{\text{mat}}\), Sqt\(^{\text{pro}}\)-Cyc\(^{\text{mat}}\) were injected together with the lineage tracers fluorescent-dextran and biotinylated dextran (Molecular Probes) into a single cell of 64- to 128-cell zebrafish wild-type embryos. Injected embryos were fixed at 50% epiboly and analyzed by in situ hybridization to detect gsc and ntl. Biotin Dextran was detected by standard protocols (Vectastain). Stained embryos were mounted in glycerol, and images captured using a Zeiss Axioplan2 microscope equipped with a Nikon DXM 1200 color camera.

**Protein expression, detection and quantification**

HEK293T cells were transiently co-transfected with plasmid DNA encoding FLAG-epitope tagged Cyc or Cyc deletions, together with pEGFP as the transfection control. The cells were harvested 48 hours after transfection, and cell pellets lysed in RIPA buffer. For analysis of protein degradation, cells transfected with Myc-tagged Cyc, Cyc\(^{\text{Δ2}}\), Sqt, Sqt\(^{\text{pro}}\)-Cyc\(^{\text{mat}}\), hNODAL or Bmp2b were treated with or without either the lysosomal inhibitor chloroquine (100 μM; Sigma) or the proteosome inhibitor MG132 (20 μM; Calbiochem) for 24 hours (Zhang et al., 2004). Protein concentrations were checked using a UV spectrophotometer. Each protein sample (15 μg) was separated by SDS-PAGE and transferred to a nitrocellulose membrane. Immunoblot (IB) analysis was performed using an anti-FLAG M2 mouse monoclonal antibody (Santa Cruz). An anti-GFP rabbit polyclonal antibody (Torrey Pines) and an anti-β-Tubulin mouse monoclonal antibody (Sigma-Aldrich) were used to detect the GFP transfection control and β-Tubulin gel loading control, respectively. Anti-mouse or anti-rabbit secondary antibodies conjugated to horseradish peroxidase (DAKO) were used, and the proteins were detected by SuperSignal West Femto Maximum Sensitivity Substrate (Pierce). The NIH Image J software package was used to quantify the proteins detected by western blots. Intensity of the individual protein bands was normalized against GFP to correct for differences in transfection efficiency between experiments and are represented as histograms.

**Extraction and detection of secreted proteins**

For extraction of secreted Nodal proteins, transfected HEK293T cells were incubated for 18 hours in DMEM containing 10% fetal bovine serum. pEGFP was used as the control to score transfection efficiency. Subsequently, the medium was changed to opti-MEM I reduced serum medium (GIBCO BRL), and conditioned for 36-48 hours at 37°C. The supernatants were collected and pre-cleared by centrifugation. Cell lysates were used in western blots analysis to detect control GFP expression. Secreted proteins in the cleared supernatants were concentrated with anti-Myc Monoclonal Antibody Affinity Matrix (Covance), and detected by immunoblot analysis with anti-Myc rabbit polyclonal antibodies (Santa Cruz).

**RESULTS**

**Cyc is inactive in the absence of its pro-domain**

The mature domain of the mouse Nodal protein by itself has been shown to signal and induce downstream target gene expression in cells and embryos (Kumar et al., 2001; Le Good et al., 2005). We tested the activities of hNODAL, Sqt and Cyc by generating truncated versions (Fig. 1A) and injecting various doses into zebrafish embryos at the one-cell stage. We tested for induction of the Nodal-responsive mesendoderm gene gsc and the pan-mesodermal gene ntl at 50% epiboly. RNA encoding full-length hNODAL, Sqt, Cyc, truncated versions of Cyc, Sqt and hNODAL, and an Activin-Cyc fusion were injected. All the truncations are less active than full-length hNODAL, Cyc or Sqt. As shown previously for mouse Nodal, injection of RNA encoding only the mature domain (hNODAL\(^{\text{mat}}\)) or Sqt (Sqt\(^{\text{mat}}\)) can induce expansion of the gsc and ntl domains at 5 pg and higher doses (Fig. 1B,C; see Tables S1 and S2 in the supplementary material). By contrast, neither the mature (Cyc\(^{\text{mat}}\)) nor the pro-domain of Cyc (Cyc\(^{\text{pro}}\)) alone induce gsc or ntl expression even at 25 pg doses. Swapping the pro-domain of Cyc with the Activin pro-domain restores partial activity at intermediate and high doses. Co-injection of Cyc\(^{\text{mat}}\) with Cyc\(^{\text{pro}}\) can induce mild expansion of the gsc and ntl expression domains, but only at very high doses (Fig. 1B,C). Therefore, unlike zebrafish Sqt and the mammalian Nodal proteins, the Cyc\(^{\text{mat}}\) domain alone is not functional by itself, and the pro-domain of Cyc is required for its activity.
The Cyc pro-domain harbors regions that regulate its expression levels and activity

As the pro-domain of Cyc is required for activity, we generated serial deletions of 30 amino acids within this domain (Fig. 2A) in order to identify specific regions that regulate Cyc function. RNA encoding various deletion versions of Cyc was injected into one-cell zebrafish embryos and tested for induction of gsc (Fig. 2B; see Table S3 in the supplementary material). Deletion of regions 3, 5, 6, 10, 11 and 12 (Fig. 2A,B) from the pro-domain of Cyc causes reduced gsc inducing activity in comparison with full-length Cyc, and deletion of region 4 (CycΔ4) abolishes almost all activity. At higher doses, CycΔ4 can induce mild expansion of the gsc expression domain in some embryos (data not shown). By contrast, deletion of regions 1, 2, 7, 8, and 9 results in enhanced gsc inducing activity when compared with full-length Cyc. Strikingly, deletion of region 2 in the Cyc pro-domain results in a dramatic increase in activity with nearly 80% of injected embryos manifesting massive expansion of the gsc expression domain, 40% of which show gsc expression in the entire blastoderm (Fig. 2B; see Table S3 in the supplementary material).

We also examined the expression levels of FLAG epitope-tagged versions of the truncated Cyc proteins in HEK293T cells, normalized against control GFP protein levels (Fig. 2C). Significantly more protein is observed for CycΔ1 and CycΔ2 (P<0.05 as determined by the paired Student’s t-test), consistent with their increased gsc-inducing ability in comparison with full-length Cyc. Although deletion of regions 7, 8 and 9 does not significantly alter the levels of these truncated Cyc proteins, their ability to induce
gsc is higher than full-length Cyc. Deletion of region 4 does not alter the levels of this protein even though gsc is not induced upon its deletion. Therefore, expression levels of the various Cyc proteins do not correlate with their ability to induce gsc. However, some regions (such as region 2) clearly regulate both Cyc protein level and activity.

**Regulation of activity and stability by the Cyc pro-domain**

Deleting the pro-domain of mouse Nodal affects its stability and range of signaling (Le Good et al., 2005). As more CycΔ2 protein is detected in comparison with Cyc, and as it is more efficient in inducing gsc, we examined the range of signaling of CycΔ2 (Fig. 3A,B). Embryos with cells expressing CycΔ2 show a marked increase in the range of signaling and ntl is induced several cell diameters from the CycΔ2 source (Fig. 3A,B). Thus, CycΔ2 behaves as a long-range morphogen, similar to Sqt. In addition, unlike Cyc, which typically does not induce gsc at the doses we used in the range of signaling assays, CycΔ2 induces gsc in 80% of injected embryos (n=41) (Fig. 3A,B). CycΔ1 also shows increased range in these assays (see Table S4 in the supplementary material), consistent with its increased expression levels (Fig. 2C), and higher activity in one-cell overexpression assays (Fig. 2B). CycΔ7, CycΔ8 and CycΔ9 levels, although not increased (Fig. 2C), show increased range of signaling, consistent with their increased activity in one-cell injection assays (see Table S4 in the supplementary material). CycΔ6, CycΔ11 and CycΔ12 also exhibit a mild increase in range of signaling assays, although overexpression of these Cyc variants at the one-cell stage does not show an increase in activity (Fig. 2B). Thus, these regions in the Cyc pro-domain have distinct effects on its activity and range of signaling.

We also swapped the pro-domains of Cyc and Sqt and tested the activity of the fusion constructs. As reported previously (Chen and Schier, 2001; Jing et al., 2006), the Sqtpro:Cycmat fusion is similar to Cyc and signals at short range. However, the pro-domain of Cyc fused to Sqtmat reduces the signaling range of Sqt (Fig. 3A,B). Expression of gsc is detected in only 27% (n=22) of embryos injected with Cycpro:Sqtmat, whereas 97% of embryos that received Sqt show gsc expression in the receiving cells. Similarly, the signaling range, as measured by ntl expression, is reduced by several cell diameters when the pro-domain of Cyc is fused to Sqtmat (Fig. 3A,B). Thus, the Cyc pro-domain harbors residues that restrict the activity and signaling range of the associated mature domain.

We also examined the processing of the various Cyc and Sqt proteins in HEK293T cells. Although both the processed and unprocessed forms are detected in supernatants from cells expressing Sqt, we typically do not detect unprocessed Cyc in the supernatants, and processed Cyc is detected at very low levels. Thus, unprocessed Cyc is very unstable (Fig. 3C) (Jing et al., 2006). We also examined the secreted proteins expressed from cells transfected with the Cycpro: sqtmat or sqtpro:cycmat fusion constructs. The unprocessed and processed forms of Sqtpro:Cycmat are increased relative to Cyc in supernatants of expressing cells, but the amount of processed protein detected is low in comparison to Sqt. We do not detect the Cycpro: Sqtmat precursor in supernatants, and only the processed form of this protein is detected, suggesting that the Cycpro region targets the precursor for processing (Fig. 3C). Thus, the pro-domain of Cyc renders the precursor unstable and regulates its processing. In addition, Cycmat is not very stable even when fused with Sqtpro. Taken together, these experiments show that Cyc is not stable in comparison with Sqt or mouse Nodal (data not shown), and that regions in the Cyc pro-domain affects its stability and signaling range.

**Region 2 of the Cyc pro-domain restricts Sqt activity and targets Cyc for lysosomal degradation**

Region 2 of Cyc restricts its activity. As Sqtpro does not harbor these residues (see Fig. S1 in the supplementary material), we tested whether swapping this region of Cyc into Sqt can affect its activity. **Fig. 2. Specific regions within the Cyc pro-domain regulate its expression levels and activity. (A)** Schematic representation of the Cyc precursor. The Cyc leader sequence is shown by the hatched box, the pro-domain in pink and the mature domain in blue. Regions 1-12 are 30 amino acid stretches in the pro-domain. **(B)** Analysis of induction of gsc in wild-type embryos injected with 1.25 pg aliquots of RNA encoding Cyc with deletions in regions 1-12 of the pro-domain. Some deletions, such as CycΔ2 and CycΔ9, have stronger gsc-inducing activity than wild-type Cyc, whereas some have either no detectable activity (CycΔ4) or reduced activity. Scale bar: 100 μm. **(C)** Expression of FLAG epitope-tagged versions of the various Cyc proteins. Immunoblots on extracts of HEK293T cells expressing FLAG-tagged Cyc proteins show more CycΔ1 and CycΔ2 protein than Cyc or the other deletions. The protein samples were loaded on two gels (lanes 1-7 and 8-14, respectively), and electrophoresed in parallel. Western blots of the input proteins are shown, detected with antibodies towards the FLAG epitope for the Cyc proteins, α-GFP for transfection control and α-Tubulin for loading controls. The graph shows normalized protein expression levels relative to the GFP control, from three independent experiments (mean±e.m.). Asterisks indicate significant differences; P<0.05 by paired Student’s t test. A representative blot is shown.
Injection of one-cell embryos with chimeric Sqt (Sqt<sup>pro:Cyc2:Sqt</sup>) containing the Cyc region 2 residues results in reduced gsc and ntl inducing ability in comparison with Sqt (Fig. 4A-C). In contrast to embryos injected with sqt RNA that show substantial expansion of the gsc and ntl expression domains (Fig. 4B,C; see Tables S5 and S6 in the supplementary material), embryos injected with sqt<sup>pro:Cyc2:Sqt</sup> RNA manifest only modest expansion of the gsc and ntl expression domains (see Tables S5 and S6 in the supplementary material). We also examined the range of signaling of Sqt<sup>pro:Cyc2:Sqt</sup> in comparison with Cyc and Sqt. Although ntl expression is induced by cells expressing Sqt<sup>pro:Cyc2:Sqt</sup>, the range of ntl expression is reduced in comparison with embryos expressing Sqt (Fig. 4D; see Table S7 in the supplementary material). Furthermore, gsc expression is not detected in most embryos expressing Sqt<sup>pro:Cyc2:Sqt</sup> (88%; n=41), similar to embryos injected with RNA encoding Cyc (Fig. 4D; see Table S7 in the supplementary material). Supernatants from HEK293T cells expressing Myc-tagged Sqt, Cyc<sup>pro:Sqtmat</sup>, Sqt<sup>pro:Cycmat</sup>, Cyc and CycΔ2 proteins. Both unprocessed (~46 kDa) and processed (~17 kDa) forms of Sqt are detected. The Cyc precursor is not detected, whereas substantial amounts of the Sqt<sup>pro:Cycmat</sup> precursor (~45 kDa) are detected. For Cyc<sup>pro:Sqtmat</sup>, no precursor is detected. Very low levels of processed Cyc<sup>mat</sup> (~16 kDa) is detected from supernatants of cell expressing Cyc, and marginally higher amounts of processed Cyc is detected in the CycΔ2 lane (~16 kDa). The samples were electrophorased on a SDS-PAGE gel; lane 6 (CycΔ2) is from a different part of the same gel. Expression of the transfection control GFP in cell lysates is shown in the lower panel.
Fig. 4. Region 2 of the Cyc pro-domain renders chimeric Sqt unstable, and targets the precursor for degradation. (A) Schematic representation of Cyc, Sqt and Sqt<sup>proCyc2</sup>:Sqt. Hatched boxes indicate the leader sequences, the Cyc pro-domain is shown in pink, the Sqt pro-domain in dark blue, the Sqt mature domain in pale blue and the Cyc mature domain in mid-blue. (B, C) Analysis of induction of <i>gsc</i> (B) or <i>ntl</i> (C) in wild-type embryos injected at the one-cell stage with 0.5, 5 or 25 pg of RNA encoding Cyc, Sqt or Sqt<sup>proCyc2</sup>:Sqt. Animal pole views of embryos at 50% epiboly showing endogenous <i>gsc</i> or <i>ntl</i> expression (I), and mild (II) or massive (III and IV) expansion of the <i>gsc</i> or <i>ntl</i> expression domains. Embryos were assessed and counted, and percentages for each class are shown in histograms. Scale bars: 100 μm. (D) 5 pg of RNA encoding Cyc, Sqt or Sqt<sup>proCyc2</sup>:Sqt was injected into one cell at the 64- to 128-cell stage together with a lineage tracer (brown). Range of signaling was examined by in situ hybridization to detect <i>ntl</i> (blue, top panel) and <i>gsc</i> (blue-brown, bottom panel) expression. The Sqt<sup>proCyc2</sup>:Sqt fusion behaves like Cyc, with its reduced <i>ntl</i> expression domain and the lack of <i>gsc</i> expression in most embryos (see Table S7 in the supplementary material). Scale bars: 100 μm. (E) Supernatants from HEK293T cells expressing Myc-tagged Sqt, Cyc and Sqt<sup>proCyc2</sup>:Sqt proteins. Both unprocessed (~46 kDa) and processed (~17 kDa) forms of Sqt are detected. For Sqt<sup>proCyc2</sup>:Sqt, the precursor is typically not detected and the processed Sqt (~17 kDa) protein levels are reduced. The processed form of Cyc (~16 kDa) is detected at very low levels. Supernatants from cells expressing pCS2Myc were used as negative control. GFP expression in the cell lysates was used to measure transfection efficiency. (F) ImmunobLOTS to detect Myc-tagged Cyc (~58 kDa), Cyc<sup>Δ2</sup> (~55 kDa), Sqt (~46 kDa), Sqt<sup>proCyc2</sup>:Sqt (~46 kDa), hNODAL (~41 kDa) or Bmp2b (~48 kDa) in extracts of cells grown in the presence or absence of the lysosomal inhibitor chloroquine. Cyc<sup>Δ2</sup> expression is not significantly altered by chloroquine treatment. By comparison, Cyc and the other proteins are stabilized by chloroquine. GFP was used as the transfection control and tubulin was used as the loading control. The histogram shows normalized protein expression levels relative to the GFP control, from three independent experiments (mean±s.e.m.). The asterisks (*) indicate significant differences (P<0.05 as determined by paired Student’s t-test) between cells not treated versus those treated with the lysosomal inhibitor for 24 hours. A representative gel from three independent experiments is shown.
presence of chloroquine (Fig. 4F). By contrast, CycΔ2 is stabilized by treatments for 24 hours with the proteosome inhibitor MG132, similar to the other proteins (see Fig. S2 in the supplementary material). We detected more Sqt, SqtproCyc2:Sqt, hNODAL and Bmp2b protein in the presence of chloroquine, indicating that these TGFβ-related proteins may also be targeted to the lysosome. Thus, region 2 of the Cyc pro-domain can render the precursors less stable, and restricts their signaling activity by targeting to the lysosome for degradation.

A mutation in the pro-domain of hNODAL renders it more active and is responsible for familial heterotaxia

A mutation in the human NODAL gene has been reported in members of a family with left-right asymmetry defects (Casey, 1998; Gebbia et al., 1997). This mutation affects an arginine residue in the pro-domain of hNODAL (see Fig. S1 in the supplementary material). We tested the activity of mutant hNODALR183Q in comparison with that of wild-type hNODAL. Injection of the mutant

---

**Fig. 5.** Mutation of a conserved arginine residue in hNODAL renders the protein more active and can induce left-right asymmetry defects. (A,B) CycR314Q and hNODALR183Q are more active than Cyc and hNODAL. Zebrafish embryos at the one-cell stage were injected with 5 or 25 pg of RNA encoding either Cyc, CycR314Q, hNODAL or hNODALR183Q. Expression of gsc (A) or ntl (B) was examined at 50% epiboly. The Arg-Gln mutations render CycR314Q and hNODALR183Q more active than the wild-type proteins. Animal pole views shown. Scale bar: 100 μm.

(C) Overexpression of CycR314Q and hNODALR183Q in Xenopus embryos causes perturbations in left-sided expression of Xpitx2. Xenopus embryos at the four-cell stage were injected with 25 pg of DNA expression constructs encoding either Cyc, CycR314Q, hNODAL or hNODALR183Q in the dorsal-left or dorsal-right blastomeres and cultured until stage 28 and processed for in situ hybridization to detect Xpitx2 expression. Embryos were scored as left, right, bilateral or no Xpitx2 expression. Ventral views of stage 28 embryos with Xpitx2 expression in the cardiac primordia are shown. Injections of hNODALR183Q or CycR314Q can induce right-sided, bilateral or no Xpitx2 expression at a higher frequency than observed after injections of hNODAL or Cyc. (D) Gut looping is also perturbed by injection of hNODALR183Q. Xenopus embryos at the four-cell stage were injected with 25 pg of DNA encoding either hNODAL or hNODALR183Q in the dorsal-left or dorsal-right blastomeres and cultured until stage 45 and scored for the direction of intestinal coiling. Injected embryos manifested either normal, reversed or abnormal looping. In some embryos, heterotaxia of intestinal coiling and cardiac laterality was observed. Right-sided injections of hNODALR183Q increase the frequency of reversed intestinal coiling in comparison with hNODAL.
hNODALR183Q RNA into zebrafish embryos at the one-cell stage causes increased expansion of the gsc andntl expression domains, in comparison with wild-type hNODAL RNA at the same doses (Fig. 5A,B). The pro-domain of the zebrafish Nodal factor Cyc also harbors this Arg residue. We also tested mutant Cyc harboring a mutation in the conserved residue (see Fig. S1 in the supplementary material). Upon injection of low doses of CycR314Q RNA into zebrafish embryos, we observed a similar increase in activity compared with Cyc (Fig. 5A,B). Thus, this Arg residue may have conserved functions in the zebrafish Nodal and hNODAL proteins.

We then tested the ability of the mutant hNODALR183Q and CycR314Q proteins to perturb left-right asymmetry in Xenopus embryos and compared it with that of wild-type hNODAL or Cyc DNA. Left- as well as right-sided microinjections of mutant hNODALR183Q or CycR314Q DNA expression constructs into Xenopus embryos can induce perturbations in expression of the Nodal target gene, Xpitz2 (Fig. 5C). By contrast, left-sided injection of the wild-type hNODAL or Cyc DNA does not cause right-sided Xpitz2 expression at the same doses. In addition, right-sided expression of mutant hNODALR183Q or CycR314Q results in a higher proportion of either bilateral or right-sided Xpitz2 expression in comparison with hNODAL or Cyc. We also examined the direction of intestinal coiling in Xenopus embryos injected with DNA expression constructs encoding either hNODAL or hNODALR183Q. The proportion of embryos manifesting normal or reversed intestinal coiling is comparable in embryos injected on the left side with either wild type hNODAL or mutant hNODALR183Q. However, right-sided expression of hNODALR183Q leads to a higher frequency of reversal in intestinal coiling in comparison with hNODAL (Fig. 5D). Thus, the proportion of embryos manifesting perturbations in Xpitz2 expression and reversal of gut looping is increased in embryos injected with mutant hNODALR183Q DNA. Taken together, the results from overexpression in zebrafish and Xenopus embryos suggest that the mutation in the hNODAL pro-domain can lead to gain-of-function activity and left-right asymmetry defects.

**DISCUSSION**

Although the Nodal-related proteins are similar in that they can act as mesoderm inducers, there appear to be some differences in the activities of these factors in embryos. The Xenopus Xnr-1 and Xnr-2 proteins differ in their ability to induce mesoderm and to affect left-right asymmetry (Jones et al., 1995; Sampath et al., 1997). Genetic analysis in zebrafish has shown that Cyc has cell autonomous as well as non-autonomous functions, and the cell autonomous activity of Cyc is particularly important in development of the ventral neural tube (Hatta et al., 1991; Sampath et al., 1998). Other experiments have suggested that Cyc acts over a short range, whereas Sqt behaves as a long-range morphogen. The mouse Nodal protein is known to have long-range signaling activities (Chen and Schier, 2001; Le Good et al., 2005). The biochemical basis for these differences is just beginning to be elucidated and a recent study showed that acidic residues in the mature domain of Cyc contribute towards its short-range activity (Jing et al., 2006). Although Cyc and Sqt share 68% identity and 78% sequence similarity in their mature domains, their pro-domains are more divergent (Erter et al., 1998; Rebagliati et al., 1999a). In particular, the Cyc pro-domain has several proline-rich regions that are not shared by any of the other vertebrate Nodal-related factors (Hudson and Yasuo, 2005; Jones et al., 1995; Morokuma et al., 2002; Rebagliati et al., 1999a; Sampath et al., 1998; Tian et al., 2003; Toyama et al., 1995; Yu et al., 2002; Zhou et al., 1993). Although it is not clear how this region evolved, the presence of the proline-rich repeats is suggestive of insertion by retroposon-like elements. Our experiments where we disrupt the Cyc pro-domain via deletions and point mutations indicate that this region of Cyc contributes to the instability of this Nodal-related factor. Similar findings have been reported for mouse Nodal. Although expression of wild-type Bmp4 could be detected in COS cells, no mature BMP4 was detected when expressed as a chimeric Nodalpro:Bmp4mat protein, and, conversely, fusion of Dorsalinpro to Nodalmat had a stabilizing effect on Nodalmat expression (Constam and Robertson, 1999). Interestingly, in our biochemical studies, we typically do not detect the Cyc precursor, suggesting that the Cyc precursor is rapidly turned over. Similar findings have been reported for a Cyc-GFP fusion protein expressed in COS cells (Jing et al., 2006). Thus, regions in the Cyc pro- as well as Cyc mature-domain are responsible for the overall instability of this Nodal-related factor, and this is probably the reason that Cyc has only short-range signaling activities in embryos.

Remarkably, swapping the pro-domain of Sqt with that of Cyc restricts the long-range signaling ability of Sqtmat. A cleavage-resistant form of mouse Nodal can induce expression of Bmp4 in cells at a distance from its source (Ben-Haim et al., 2006). This raises the possibility that the zebrafish Nodal precursors may act at a distance from their source, and the Cyc pro-domain reduces the signaling range of the Cyc:Sqt precursor. In addition, we observed that no secreted precursor of the Cycpro:Sqtmat fusion is detected, further verifying that the pro-domain of Cyc contributes to the turnover of the precursor proteins. This contrasts with previous reports using a Cycpro:SqtmatGFP fusion that found that activity of a Cyc:Sqt chimera is indistinguishable from that of Sqt (Jing et al., 2006). We cannot explain the reason for the difference in activity of the Cyc:Sqt fusion observed by the two groups, but one possibility is that the larger GFP epitope tag used by Jing et al. may allow their fusion protein to be more stable than our Cyc:Sqt fusion. As observed by Jing et al., we also find that the Sqtpro:Cycmat fusion protein is, by and large, similar in activity to Cyc. In some injected embryos (43%, n=44), we observed a mild expansion (two or three cells) of the ntl expression domain. Interestingly, we detect substantial amounts of the Sqtpro:Cycmat precursor in supernatants of expressing cells, whereas secreted Cyc precursor is typically not detected in our assays. However, the amount of processed Cycmat from the Sqt:Cyc fusion is low, similar to Cyc. Therefore, although the Sqt pro-domain can stabilize the Sqtpro:Cycmat precursor, mature Cyc is inherently unstable. These findings also point towards fundamental differences between the pro-domains of Cyc and Sqt.

Surprisingly, unlike zebrafish Sqt or the mammalian Nodal factors, the Cyc mature domain by itself has no activity. Swapping the Cyc pro-domain with that of Activin restored some activity, but not to the levels observed with Cyc, indicating that the Activin:Cyc chimera is less active than wild-type Cyc. Thus, the Cyc pro-domain is vital for its functions. We typically detect less processed Cyc in comparison to processed Sqt and mouse Nodal in culture supernatants. Therefore, the lack of activity of Cycmat may be because it is much more unstable than processed Sqt and mouse Nodal. We also find it difficult to detect hNODAL protein in our assays (data not shown) (Topczewska et al., 2006), indicating similarities between Cyc and hNODAL proteins. As increased hNODAL levels have been associated with tumor progression, the turnover of hNODAL may be an important step in regulation of its activities.

We identified a region in the N terminus of the Cyc pro-domain (region 2), which when deleted increases Cyc activity and allows it to function as a long-range mesoderm inducer, similar to Sqt and...
Functions of pro-domain of Nodal factors

mouse Nodal. How does region 2 regulate Cyc protein activity? One possibility is that the residues in region 2 target the precursor for degradation, leading to its rapid turnover in embryos. In support of this possibility, region 2 in Cyc harbors a YRHY motif, which has some similarity to consensus endosome/lysosome-targeting sequences (Bonifacio and Dell’Angelica, 1999). Removal of this region makes Cyc more stable and, conversely, a chimeric Sqt protein that contains region 2 of the Cyc pro-domain is less stable and less active than Sqt, and sensitive to the lysosomal inhibitor chloroquine. Studies with Xenopus Bmp4 have also shown that the processing of BMP4 is regulated by the pro-domain of this factor, and cleavage within the pro-domain directs degradation of mature Bmp4 (Sopory et al., 2006). Therefore, it is conceivable that these residues in the Cyc pro-domain regulate its degradation. Nonetheless, we cannot rule out other mechanisms that may regulate Cyc protein synthesis, secretion or processing. For example, glycosylation regulates the stability of mouse Nodal protein (Le Good et al., 2005). There are four predicted N-glycosylation sites in Cyc, all of which reside in the Cyc pro-domain, in regions 1, 2, 4 and 11, respectively. Of the Cyc deletions in regions with the predicted glycosylation sites, CycA1 and CycA2 exhibit enhanced activity of Cyc, whereas CycA11 and CycA4 display reduced and no activity, respectively. Therefore, we do not observe a direct correlation between glycosylation sites and Cyc stability or activity. However, as our experiments make use of deletions, it remains possible that glycosylation may regulate Cyc activity.

A conserved Arg residue in the hNODAL and Cyc pro-domains, when mutated, causes increased activity of these Nodal proteins. It is not clear whether the R>Q mutation increases hNODAL activity by stabilizing the protein or by regulating its interactions with other components of the NODAL signaling pathway. We find that this increase in activity is modest, so it may allow germ layer formation and patterning in early embryos. Nonetheless, this modest increase in activity is sufficient to cause laterality defects. Human laterality defects have been found to be associated with loss-of-function mutations in the Nodal co-receptor CFC1 (Cryptic), the zinc-finger transcription factor Zic3, the Nodal antagonist Lefty and the type II activin receptor Act2b (Bamford et al., 2000; Gebbia et al., 1997; Kosaki, K. et al., 1999; Kosaki, R. et al., 1999). As expression of the mutant hNODAL protein can enhance laterality defects in frogs, it is likely that it is the gain-of-function activity of the mutant hNODAL that is responsible for the heterotaxia defects found in individuals harboring this mutation. Thus, the pro-domain of the Nodal factors not only regulates their stability, but is also important for other aspects of signaling by these factors. Understanding the precise mechanisms by which the pro-domain influences Nodal activity will provide further insights into the regulation of these essential signaling factors.

We thank Ray Dunn, Aniket Gore, Sudipto Ray and members of the Sampath laboratory for discussions and suggestions; Ray Dunn for human ES cell RNA; Chan Aye Thu, Wang Xin, Siew Peng Tan, Hong Xin and Helen Ngoc Bao Quach for technical assistance; Shi Min Lim for assistance with the manuscript; Srinivas Ramasamy and the TLL Bioinformatics core; and the TLL laboratory of K.S. is supported by Temasek sequencing and fish facilities for support. B.A. and C.M.J. are supported by A*STAR, Singapore. Work in the laboratory of K.S. is supported by the recognition of tyrosine-based sorting signals. J. Cell Biol. 145, 923-926.

Supplementary material
Supplementary material is available at http://dev.biologists.org/cgi/content/full/135/15/2649/DC1

References


