Primary Structure Requirements for *Xenopus* Nodal-related 3 and a Comparison with Regions Required by *Xenopus* Nodal-related 2*

Received for publication, December 22, 1999, and in revised form, February 18, 2000

Carin Hansen Ezal, Christopher D. Marion, and William C. Smith‡

*From the Department of Molecular, Cellular, and Developmental Biology, and the Neuroscience Research Institute, University of California, Santa Barbara, California 93106*

Transforming growth factor-β superfamily members play important roles in the early development of animals. Activin and the *Xenopus* nodal-related proteins 1, 2, and 4 induce muscle actin from *Xenopus* ectodermal explants, whereas the bone morphogenetic proteins 4 and 7 induce ectoderm to differentiate as epidermis. Bone morphogenetic proteins are antagonized by soluble binding proteins such as noggin and chordin, which leads to expression of neural cell adhesion molecule in animal caps. The transforming growth factor-β superfamily member *Xenopus* nodal-related 3 also induces the neural cell adhesion molecule through inhibition of bone morphogenetic proteins. Therefore, whereas *Xenopus* nodal-related 2 and 3 share a high amount of sequence homology, they lead to very different cell fates. This study investigates the functional domains that distinguish the activities of these two factors. It was found that mutually exclusive regions of nodal-related 2 and 3 were required for activity. The central region of the mature domain is required for nodal-related 2 to induce muscle actin, whereas the N- and C-terminal ends of the mature domain are required for nodal-related 3 to induce neural cell adhesion molecule. These results help to define the minimal domains required for the unique activities of these factors.

The transforming growth factor-β (TGF-β) superfamily is composed of a myriad of related secreted proteins that are important regulators of development and physiology in both vertebrates and invertebrates (1). Most TGF-βs are synthesized as preproteins that are biologically inactive until proteolytically processed at R-X-(K/R)-R and R-X-X-R consensus sequences by subtilisin-like proprotein convertases (2–4). Active TGF-β proteins consist of two 12–16-kDa peptides that show varying ability to function as homo- and heterodimers. Superfamily members have seven highly conserved cysteines in the C-terminal mature domain that form intra- and interchain disulfide bonds. Within the monomer, disulfide bond pairs are formed between the first and fifth, second and sixth, and third and seventh cysteines. The fourth conserved cysteine makes an interchain bond between dimer subunits. Another conserved feature of the superfamily is a glycine residue between the second and third cysteines, within the consensus sequence CXGXC. It is thought that steric hindrance would require a glycine (Fig. 1, indicated by an asterisk) in this position for proper folding, because two disulfide bonds form a closed ring on either side of this residue, which prevents its substitution (5, 6).

Three-dimensional structural studies of TGF-βs 1, 2, and 3, as well as bone morphogenetic proteins (BMPs) 2 and 7 predict a conserved “cysteine knot” structure that has been described as a “left hand” (Fig. 1) (5–10). The “heel” of the hand structure is an α-helix formed by the amino acids between the third and fourth cysteines. Extending out from this are two long loops consisting of β-sheets that form “fingers.” The N-terminal amino acid region of the mature domain extending from the cleavage site to the first conserved cysteine represents the “thumb” (also referred to as the proknot sequence, (10)). Members of the TGF-β subfamily (TGF-βs 1–5) and activin have two additional conserved cysteines in this region, which form a disulfide bond anchoring a short α-helix to the first β-sheet of finger 1. Because most other superfamilies share only the seven conserved cysteines, they lack this additional disulfide bond. In BMP2 and BMP7, the thumb region is disordered and cannot be resolved in electron density maps, so the structure of this region in these superfamilies is unknown. Members of the nodal subfamily of the TGF-βs play important roles in vertebrate mesoderm induction and patterning (11–13). Whereas the mouse and chick appear to have single members of the nodal gene family, duplications have led to at least four family members in *Xenopus* (Xnr1–4) and two in zebrafish (*squint* and *cyclops*) (14–16). Comparison of predicted amino acid sequences indicates that Xnr1, Xnr2, and Xnr3 are more closely related to each other than to Xnr4. Most notably, Xnr1, Xnr2, and Xnr3 share the unique feature of having the sequence CXXC between the fourth and fifth conserved cysteines. This sequence is also found in chick nodal and zebrafish squint. Xnr4, mouse nodal, and zebrafish cyclops, as well as the majority of other TGF-β superfamilies, have the sequence CC for these two cysteines. *Xenopus* nodal-related 3 (Xnr3) is unique among the nodal subfamily in having several primary structure features that diverge from the TGF-β superfamily consensus (17). First, Xnr3 is missing the last of the seven conserved cysteines. Second, whereas all other superfamilies have a glycine located between the second and third cysteines, Xnr3 has a serine in this position. Perhaps this substitution is allowed because Xnr3 lacks one of the two disulfide bonds that constrain this residue to a glycine in other superfamilies. Together these observations suggest that Xnr3 does not form the characteristic knot structure (Fig. 1).

Xnr3 also has biological activity in developing *Xenopus* embryos that differ markedly from the other nodal-related factors.

---

* This work was supported by Grant GM52835 from the National Institutes of Health and the Beckman Young Investigators Program (to W. C. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: (805) 893-7698; Fax: (805) 893-2005; E-mail: w_smith@lifesci.ucsb.edu.

*The abbreviations used are: TGF-β, transforming growth factor-β; BMP, bone morphogenetic protein; Xnr, *Xenopus* nodal-related; MA, muscle actin; NCAM, neural cell adhesion molecule; TBS, Tris-buffered saline; TBSM-T, TBS with 3% dried milk and 0.1% Tween.
In *Xenopus* animal cap induction assays, Xnr1, 2, and 4 induce the mesodermal markers brachyury (when assayed at early gastrula stage) and muscle actin (MA) at tailbud stage (11, 16, 20). However, Xnr3 blocks the activities of the TGF-β family members BMP4 and activin, does not induce brachyury or MA, and instead induces the mesodermal markers brachyury (when assayed at early gastrula stage) and muscle actin (MA) at tailbud stage (11, 16, 20). Twenty animal caps were cultured together in ~100 µl of 1/3X MMR (1X MMR: 0.1 M NaCl, 2 mM KCl, 1 mM MgSO4, 2 mM CaCl2, 5 mM HEPES, pH 7.8, 0.1 mM EDTA, 1.5 mM sodium phosphate pH 7.5, 0.75 mM NaHCO3) with 25 mg/ml gentamicin and bovine serum albumin added to 1 mg/ml.

Site-directed Mutagenesis—Oligonucleotide-mediated site-directed mutagenesis was performed using a *dUTP*-containing strand of *E. coli* DNA polymerase (28). Single-strand uracil-containing templates were prepared from Xnr3 and Xnr2 and annealed with the following oligonucleotides to make the indicated mutants: cmXnr3, 5'-GGTGAATTGTTGTCGAGAAGGAGGACG-3'; substitution of amino acids of Xnr3, 5'-AAATGTGTGGAATTTCCCTCTTCAATGCAGGA-3'; chimeras 1–12 (Xnr2 + Xnr1), 5'-CAATGCATATAGTGTTGAGGACG-3'; chimeras 13–14 (Xnr3 + Xnr2), 5'-GAATATATAGTGTTGAGGACG-3'; chimeras 3–12 (Xnr3 + Xnr1), 5'-GGAATATATAGTGTTGAGGACG-3'; substitution of amino acids 386–390 of Xnr2 with the Xnr3 sequence NEDFI, 5'-GCCCATGACGAAAGGACG-3'; substitution of the Xnr3 thumb domain is only two residues long because the structure for the N-terminal residues was not determined (14).

For making epitope-tagged Xnr3, polymerase chain reaction was used to introduce a *SalI* site into Xnr3. The 5'-end (nucleotides 1–940 of Xnr3) was amplified using the M13 reverse primer and the oligonucleotide 5'-CATCACGAGG-3'. Single-strand uracil-containing templates were prepared from Xnr3 and Xnr2 and annealed with the following oligonucleotides to make the indicated mutants: cmXnr3, 5'-GGTGAATTGTTGTCGAGAAGGAGGACG-3'; substitution of amino acids of Xnr3, 5'-AAATGTGTGGAATTTCCCTCTTCAATGCAGGA-3'; chimeras 1–12 (Xnr2 + Xnr1), 5'-CAATGCATATAGTGTTGAGGACG-3'; chimeras 13–14 (Xnr3 + Xnr2), 5'-GAATATATAGTGTTGAGGACG-3'; chimeras 3–12 (Xnr3 + Xnr1), 5'-GGAATATATAGTGTTGAGGACG-3'; substitution of amino acids 386–390 of Xnr2 with the Xnr3 sequence NEDFI, 5'-GCCCATGACGAAAGGACG-3'; substitution of the thumb domain is only two residues long because the structure for the N-terminal residues was not determined (14).

In *Xenopus* animal cap induction assays, Xnr1, 2, and 4 induce the mesodermal markers brachyury (when assayed at early gastrula stage) and muscle actin (MA) at tailbud stage (11, 18). All three also have the ability to rescue mesoderm in *VegT*-null gastrula stage and muscle actin at tailbud stage (11, 16, 20). Twenty animal caps were cultured together in ~100 µl of 1/3X MMR (1X MMR: 0.1 M NaCl, 2 mM KCl, 1 mM MgSO4, 2 mM CaCl2, 5 mM HEPES, pH 7.8, 0.1 mM EDTA, 1.5 mM sodium phosphate pH 7.5, 0.75 mM NaHCO3) with 25 mg/ml gentamicin and bovine serum albumin added to 1 mg/ml.
linearized by XhoI digestion and RNA- transcribed with T3 RNA polymerase. Injection of 1 ng of RNA was at one cell stage 1, and animal cap explants were isolated at stages 8–9.

RNA Extraction and Analysis—Total RNA was isolated from animal cap explants and embryos using Trizol Reagent (Life Technologies, Inc.). For Northern analysis, 20 caps were used/sample and electrophoresed on formaldehyde-containing agarose gels (29). Gels were transferred to Hybond-N nylon membrane (Amersham Pharmacia Biotech) by capillary action overnight and hybridized with QuikHyb (Stratagene). Random-primed 32P-labeled probes were prepared using the Prime-a-gene system (Promega) with isolated fragments from NCAM (data not shown). In contrast to other TGF-

Western Blotting—Xenopus embryos were injected with 2.5 ng of RNA, and 20 animal caps were isolated and grown to stages 20–25. Animal caps were homogenized in reducing Laemmli sample buffer (33) and run on a 12% polyacrylamide gel at 175 V for an hour. Gels were transferred to Hybond-ECL nitrocellulose (Amersham Pharmacia Biotech) for 1 h at 40 V and blocked for 90 min in TBS with 3% dried milk (filtered through Whatman 4 paper) and 0.1% Tween (TBSM-T). Anti-Xnr3 antibody was diluted 1:2500 and incubated overnight. The blots were then washed 4 times for 10 min in TBSM-T, incubated with goat anti-rabbit horseradish peroxidase secondary diluted 1:5000 in TBSM-T for 40 min, and washed 6 times for 10 min in TBSM-T with a final 10 min wash in TBS without detergent. Detection was done with SuperSignal (Pierce).

Antibodies—The Xnr3 antibody was made in rabbit against the synthetic peptides HVSTVPPKPIEEIKPEC and CHVSTVPPKPIEEIKPE, which correspond to the amino acids at the N terminus of the putative mature domain (the thumb domain). These peptides were linked to keyhole limpet hemocyanin and were used to immunize two rabbits. Serum from both rabbits was tested for binding of antibody to antigen coupled to keyhole limpet hemocyanin and Myc, a cleavage site mutation (cmXnr3), or a Xnr2–Xnr3 fusion protein. For cmXnr3, ribosomal bands were used as a loading control in place of elongation factor 1 (32).

RESULTS

Xnr3 Requires N-terminal Domain of Predicted Mature Protein—Epitope tagging is a commonly used technique in the study of the structure and function of TGF-β superfamily members. Epitope tags have been successfully inserted into the thumb regions of many TGF-β superfamily members, including dorsalin, Xnr1, Vg1, and activin (34, 35). However, construction of similarly modified versions of Xnr3 yielded surprising results that provide clues to the structural requirements for Xnr3 activity. The addition of tags (Myc or hemagglutinin) to the Xnr3 mature domain. Because the results with epitope tagging suggested that some feature of the thumb region was important for activity, the fusion of the pro- and mature domains was made amine-termed N-terminal of the Xnr3 putative cleavage site. As with the activin fusion protein, no NCAM activity of the fusion protein could be detected (Fig. 2). Although negative data are by their nature not conclusive, our observations suggest that the ability to swap domains results in one of the Xnr3 mature domain.

TABLE I

| Chimera number | Amino acid sequence composition |
|----------------|--------------------------------|
| 1              | Xnr2 1–325, Xnr3 326–401       |
| 2              | Xnr2 1–325, Xnr2 332–405       |
| 3              | Xnr2 1–378, Xnr3 373–401       |
| 4              | Xnr2 1–370, Xnr2 377–405       |
| 5              | Xnr2 1–331, Xnr2 326–370, Xnr2 377–405 |
| 6              | Xnr2 1–325, Xnr2 332–378, Xnr3 373–401 |
| 7              | Xnr2 1–298, Xnr2 304 replaced with G when AvoI is added, Xnr2 305–378, Xnr3 373–401 |
| 8              | Xnr2 1–298, Xnr2 304 replaced with G when Aval is added, Xnr2 305–403, Xnr3 398–401 |
| 9              | Xnr2 1–329, Xnr2 336 (the G), Xnr3 331–370, Xnr2 377–405 |
| 10             | Xnr2 1–325, Xnr2 332–378, Xnr3 373–397, last two amino acids CA |
| 11             | Xnr2 1–325, Xnr2 323–345, Xnr3 340–342, Xnr2 349–378, Xnr3 373–401 |
| 12             | Same as 10 with Myc tag added between 291 and 292 |

2 W. Smith, unpublished observations.

FIG. 2. A, diagram of Xnr3 showing the sequence around the putative cleavage site, location of epitope tag insertion, and fusion protein junction. B, Northern blot for NCAM and elongation factor 1 α (EF1α) (loading control) from Xenopus animal caps injected with RNAs encoding wild-type Xnr3 or Xnr2, epitope-tagged versions of Xnr3 (hemagglutinin and Myc), a cleavage site mutation (cmXnr3), or a Xnr2/Xnr3 fusion protein. For cmXnr3, ribosomal bands were used as a loading control in place of elongation factor 1 α.
The most likely processing site for Xnr3 is the sequence RRLRR (amino acids 270–274), although another less likely site is found in the flanking region (residues 179–182). The processing site at residues 270–274 was mutated to RKLSS (cmXnr3) to assess the importance of the site for activity. We hoped to use this strategy to make a dominant negative cleavage mutant Xnr3, as had been done successfully with TGF-β1, activin, BMPs, and Xnr2 (41–45). Surprisingly, cmXnr3 is still functional as a neural inducer (Fig. 2). Western blotting of extracts from mRNA-injected animal caps using a polyclonal antibody directed at a mature domain peptide showed that the cleavage site mutation dramatically reduced the amount of processed protein, with the subsequent appearance of the unprocessed proprotein (Fig. 3). Even though it is possible that some processing could occur at amino acids 179–183, it would be unlikely. Therefore, the unprocessed protein or a protein containing a significant amount of prodomain has biological activity.

We further examined the importance of the Xnr3 thumb region, including the putative processing site and surrounding residues. A chimera was made in which a 22-amino acid sequence was deleted (residues 270–291), extending from the 5 residues in the putative cleavage site toward the C terminus to include an additional 17 amino acids. This deletion mutant was therefore lacking the region immediately N-terminal of the location where epitope tags were inserted and has seven amino acids before the first conserved cysteine. This chimera had no apparent activity when injected into whole Xenopus embryos (data not shown) or when assayed in explanted animal caps for NCAM induction (Fig. 2). Together these results point to structural requirements around the cleavage site for Xnr3 activity that are not found in Xnr2 or nodal.

**Chimeras of Xnr3 and Xnr2 Reveal Different Domains for Activity**—To further characterize regions of Xnr3 required for its unique NCAM-inducing activity, we made chimeric proteins fusing regions of Xnr3 to the closely related factor Xnr2, which does not have direct NCAM-inducing activity (46). First, we divided the mature region of Xnr2 and Xnr3 approximately into thirds by making a set of six chimeras (Fig. 4A). Because prior experiments indicated that Xnr3 likely required different regions for activity than TGF-β1, this approach was used to narrow down the possible sequences that might be involved in the activity of Xnr3. The N-terminal one-third includes the thumb and “finger 1,” the center region the α-helix heel, and the C-terminal portion contains “finger 2.” To make the fusion chimeras, we added a number of restriction sites, most of which resulted in no change in amino acid sequence. However, a BanII site was added to Xnr3 at nucleotide 1127 to correspond with a BanII site found in Xnr2, which resulted in changing methionine 371 to a serine. This substitution alone resulted in no apparent change in Xnr3 activity (data not shown). The activity of the chimeras was examined by injecting in vitro transcribed mRNA into the animal pole of Xenopus embryos at the one-cell stage. For each chimera, 20 animal caps were dissected at late blastula stage and grown until stage 25. RNA isolated from the animal caps was analyzed by Northern blotting for NCAM, MA, and elongation factor 1 α transcripts. The symbol +/– is used in cases where activity was slightly above background, but much weaker than the activity of the controls. B, representative Northern blots. Each Northern blot is shown with four controls: whole embryo, un.injected animal caps, Xnr2 RNA-injected animal caps, and Xnr3 RNA-injected animal caps. Experimental lanes that are grouped together are from the same experiment and were assayed together. In some cases it was necessary to cut out intervening lanes that were not relevant to the figure. The MA probe cross-reacts to a lesser extent with cytoskeletal actin; the MA-specific signal is the lower band. EF1α, elongation factor 1 α.

**Activity—**To further characterize regions of Xnr3 required for its unique NCAM-inducing activity, we made chimeric proteins fusing regions of Xnr3 to the closely related factor Xnr2, which does not have direct NCAM-inducing activity (46). First, we divided the mature region of Xnr2 and Xnr3 approximately into thirds by making a set of six chimeras (Fig. 4A). Because prior experiments indicated that Xnr3 likely required different regions for activity than TGF-β1, this approach was used to narrow down the possible sequences that might be involved in the activity of Xnr3. The N-terminal one-third includes the thumb and “finger 1,” the center region the α-helix heel, and the C-terminal portion contains “finger 2.” To make the fusion chimeras, we added a number of restriction sites, most of which resulted in no change in amino acid sequence. However, a BanII site was added to Xnr3 at nucleotide 1127 to correspond with a BanII site found in Xnr2, which resulted in changing methionine 371 to a serine. This substitution alone resulted in no apparent change in Xnr3 activity (data not shown). The activity of the chimeras was examined by injecting in vitro transcribed mRNA into the animal pole of Xenopus embryos at the one-cell stage. For each chimera, 20 animal caps were dissected at late blastula stage and grown until stage 25. RNA isolated from the animal caps was analyzed by Northern blotting for NCAM, MA, and elongation factor 1 α transcripts. The symbol +/– is used in cases where activity was slightly above background, but much weaker than the activity of the controls. B, representative Northern blots. Each Northern blot is shown with four controls: whole embryo, un.injected animal caps, Xnr2 RNA-injected animal caps, and Xnr3 RNA-injected animal caps. Experimental lanes that are grouped together are from the same experiment and were assayed together. In some cases it was necessary to cut out intervening lanes that were not relevant to the figure. The MA probe cross-reacts to a lesser extent with cytoskeletal actin; the MA-specific signal is the lower band. EF1α, elongation factor 1 α.
and chimera 4 had Xnr3 sequence except for 28 amino acids at the C terminus. Neither of these chimeras had NCAM- or MA-inducing activity. The low level of MA seen in one case with chimera 4 was perhaps because of contamination. Thus, whereas the C-terminal 74 residues of Xnr2 were able to confer MA-inducing activity to the fusion protein, a smaller segment consisting of only the C-terminal 28 residues was inactive. When the middle third of the mature domain of Xnr2 was replaced with the corresponding domain of Xnr3 (chimera 5; Fig. 4A), a NCAM-inducing activity comparable to that of wild-type Xnr3 (Fig. 5, A and B).

### Residues Important for Proper Folding of Other TGF-β Family Members Appear Unimportant for Xnr3—As detailed in the Introduction, Xnr3 has several features that differ from the consensus for TGF-β superfamily members, including the lack of the seventh conserved cysteine and the substitution of a serine for a glycine residue between the second and third conserved cysteines. Mutation of this serine back to glycine did not alter the biological activity of Xnr3, even in chimeras in which a seventh cysteine was added (data not shown). We used chimera 9 to test whether the substitution of the conserved glycine residue by serine in Xnr3 was responsible for the lack of activity of chimera 4, which contains the center region of Xnr3. Chimera 9 was identical to chimera 4 except that the serine to glycine substitution had been made (Fig. 5A). This chimera yielded an unexpected result. The fusion protein consistently induced high levels of NCAM transcript (Fig. 5B). The result, and the fact that Xnr3 can tolerate changes to positions known to be crucial in other TGF-β superfamily members (47), suggest fundamental structural differences.

### The Center Third of the Xnr2 Mature Domain Is Sufficient for Muscle Actin Induction—Whereas the presence or absence of the seventh cysteine did not alter the activity of wild-type Xnr3, the importance of this alteration to the activity became evident in the context of the Xnr2/Xnr3 chimeras. In chimera 6, which had strong NCAM- and no MA-inducing activity, only the middle third of the mature domain of Xnr3 was substituted with the Xnr2 sequence (Fig. 4). However, when the seventh cysteine was added back (chimera 10; Fig. 6A), the chimera acquired strong MA-inducing activity in three and weak activity in one of six independent trials (Fig. 6B). Despite the variability in the induction of MA transcript, chimera 10 induced NCAM above background levels in all six trials (Fig. 6B).

Additional mutations were made to determine if the sequences responsible for conferring MA-inducing activity could be narrowed further. Within this 46-amino acid segment, 28 amino acids are identical between Xnr2 and Xnr3. To find out which of the remaining 18 remaining amino acids were critical, we hypothesized that necessary residues would be ones conserved among Xnr1, 2, and 4, which all induce MA, but not with Xnr3. The amino acid sequence FKP found in this region of Xnr2 (residues 346–348) would appear to be a likely candidate (Fig. 6C). The identical sequence is found in Xnr1, whereas the corresponding sequence in Xnr4 is VKP, representing one conservative change. However, in Xnr3, the analogous amino acids chimera was still active for NCAM induction at wild-type levels (Fig. 5B).

The results of experiments with chimeras 6 and 7 point to regions at the C- and N-terminal regions of the mature domain of Xnr3 that are important for activity. We made an additional chimera (Fig. 5A; chimera 8) to examine the C-terminal region of Xnr3 more closely. The C-terminal four amino acids of Xnr3 were hypothesized to be important for activity because this region was most divergent between Xnr3 and the consensus sequence for TGF-β superfamily members. Chimera 8 is identical to chimera 7, except that chimera 8 has only the four C-terminal amino acids of Xnr3. This fusion protein combining the Xnr3 proregion, N-terminal 22 amino acids of the Xnr3 mature domain, and four C-terminal residues had NCAM-inducing activity comparable to that of wild-type Xnr3 (Fig. 5, A and B).

**Fig. 5.** A, Xnr2/Xnr3 chimeras indicate that 22 amino acids at the N terminus and four residues at the C terminus of the mature domain are sufficient for NCAM-inducing activity (chimera 8). B, representative Northern blots. Samples were processed and scored as in Fig. 3. EF1α, elongation factor 1α.
DISCUSSION

Requirements for Xnr3 NCAM-inducing Activity—Experimental manipulation of Xnr3 primary structure helped to reveal regions important for activity. Specifically, proregion fusions, epitope tags, and site-directed mutations indicated that the thumb domain of Xnr3 is similarly unstructured. Therefore, Xnr3 would appear to be unique among related family members in having such strict structural requirements for this domain. This domain may be involved in a receptor interaction or it could play an indirect role in the overall folding of the protein, in which case disruption of this domain would affect the structure of another region.

Although many other TGF-βs, including nodal and Xnr2, are able to function when produced as fusion proteins with heterologous proregions, this does not appear to be the case for Xnr3. Xnr3 fusions failed to produce active protein when the junction was in the thumb domain, which is done commonly for other TGF-βs, or even N-terminal of the putative cleavage site. Surprisingly, even fusion of Xnr3 to the prodomain of the closely related Xnr2 failed to make an active protein. In contrast, the Xnr3 prodomain can be substituted into Xnr2 (e.g. chimera 2, Fig. 4). Previous studies on nodal have shown that prodomains play a crucial role in determining stability of the mature domain (39). Therefore, one possibility is that the strict requirements for Xnr3 prodomain relate to protein stability and that the presumed structural anomalies of Xnr3 can not be stabilized by other prodomains. Mutations to the putative cleavage site appear to indicate that processing is not required for Xnr3 activity, even though residues surrounding this region are necessary for activity. Because there are other distal potential cleavage sites, it is possible that Xnr3 may still be processed to a lesser extent. Mutations of the putative cleavage site in Xnr2 result in a protein with dominant negative activity, whereas similar mutations to Xnr1 and Xnr4 do not diminish activity (44). It was speculated that the residual activities of the mutant Xnr1 and Xnr4 were because of processing at alternative sites.

The results of the initial domain swaps between Xnr2 and Xnr3 (Fig. 4) showed that the prodomain, first, and last one-third of the Xnr3 mature domain were required for activity. Additional chimeras revealed that NCAM-inducing activity re-
quired the thumb domain but not the region between the first and second cysteines. Residues at the C terminus are required as well, as indicated by the first series of chimeras. However, just the most C-terminal four amino acids appear to be sufficient for NCAM induction in combination with the Xnr3 pro- and thumb domains (chimera 8). Structural predictions of Xnr3 based on BMP7 predict that the C-terminal region is in close proximity to the thumb domain, and that it is possible these two areas interact (49–51).

The activity of chimera 9 is particularly hard to explain. Chimera 9 differs from chimera 4 only by the substitution of a glycine for a nonconsensus serine residue found between the second and third cysteine residues of Xnr3 (Fig 5A). Whereas chimera 4 had virtually no MA- or NCAM-inducing activity, chimera 9 had strong NCAM-inducing activity. Thus with chimera 9, NCAM-inducing activity was present even though the chimera lacked an Xnr3 sequence at the C terminus. It is possible that changing serine to glycine allows disulfide bonds to form between the second and sixth and the third and seventh cysteines, as in most TGF-βs, which could mimic the folding of Xnr3 in the absence of the N-terminal 4 amino acids.

Requirements for Xnr2 MA-inducing Activity—Unlike Xnr3, Xnr2 has the conserved features of a typical TGF-β superfamily member, including all seven conserved cysteines and the stereotypically conserved glycine between cysteines two and three. Tertiary modeling based on the BMP7 and TGF-β2 structures predict that Xnr2 has a very similar structure (49–51). In chimera 2, the entire prodomain and first third of the mature domain of Xnr2 was substituted with the corresponding region of Xnr3. This chimera retained MA-inducing activity similar to that of wild-type Xnr2. On the other hand, chimera 1, which contained the reciprocal swaps between Xnr2 and Xnr3, had very little or no activity. The results from chimera 10 showed that the requirements for MA-inducing activity could be reduced even further to the center-one-third of the mature domain provided that a C-terminal seventh cysteine was added. Without the seventh cysteine (chimera 6), the fusion protein had strong NCAM-inducing activity but no MA-inducing activity. Having identified the central third of the Xnr2 mature domain as containing essential features for MA-inducing activity, we compared the sequences of the four Xenopus nodal-like genes within this region to find differences that might account for the divergent activities of Xnr3 versus Xnr1, -2, and -4. Near the center of this region we identified the 3-amino acid sequence FKP in the central domain of Xnr3. Site-directed mutations to change these three amino acids in Xnr2 to the corresponding sequence in Xnr3 resulted in a fusion protein with greatly reduced MA-inducing activity but that retained strong NCAM-inducing activity. If the alignment of Xnr2 with the three-dimensional structures of BMP7 and TGF-β is used to model Xnr2, these three amino acids would be just N-terminal of the α-helix forming the heel of the hand-shaped fold (49–51).

The domain that we have identified as being essential for MA-inducing activity by Xnr2 differs from regions known to be required for the activity of TGF-β1 and TGF-β2. Similar domain swapping studies have identified residues that are responsible for the divergent activities in TGF-β1 and TGF-β2 (52, 53). It was shown that exchanging residues 92–98 of the mature domain was sufficient to change the activity of TGF-β1 to resemble that of TGF-β2 in a LS513 cell growth assay. In addition, this protein no longer bound to the TGF-β1 receptor, TβRII, which TGF-β1 recognizes, but TGF-β2 does not. Significant ly, this part of the protein forms an extended surface loop forming the end of finger 2 and therefore is likely to be involved in receptor interactions (52). To test if similarly positioned residues played a role in distinguishing Xnr2 and Xnr3 activities, the sequences of Xnr2, Xnr3, TGF-β1, and TGF-β2 were first compared to determine which residues in the Xnrs were analogous to positions 92–98 of TGF-β. In Xnr2, the corresponding amino acid sequence is ENEDFIL, whereas in Xnr3 it is ENEDFIL. We were confident in the alignment of these sequences because all four proteins share a tyrosine immediately preceding this sequence, and it is followed in 12 residues by the sixth conserved cysteine. Because the first and last amino acids are already shared between the two proteins, mutagenesis was done to change residues DEEVV to Xnr2 to NEDFI. Surprisingly, this mutant Xnr2 protein retained wild-type MA-inducing activity (data not shown). Whereas it is possible that the residues in Xnr2 analogous to those identified in TGF-β1 and TGF-β2 lie slightly more N- or C-terminal to those mutated, our Xnr2/Xnr3 chimera results suggest that very different regions of the noda and TGF-βs may be required for activity. Recent results consistent with this possibility have suggested that signaling by Xnr2 might be atypical for a TGF-β superfamily member. It has been found that a mutant Xnr2 with the fourth cysteine changed to a serine retains MA-inducing activity (44). This cysteine forms the interchain bond in the ligand dimer in other TGF-β superfamily members (5–10). In activin, mutation of this residue results in protein with only 2% of wild-type biological activity (47).

Both Direct Neural and Mesoderm Inducing Activities may Co-exist in One Molecule—Chimera 10 had properties consistent with both Xnr2 and Xnr3 activities. Although it induced MA in three of six independent assays, it had strong NCAM-inducing activity every time it was tested. Alterations to the structure of chimera 10 appear to independently disrupt one type of activity or the other. When the seventh cysteine was absent (chimera 6), making the protein more Xnr3-like, MA-inducing activity was lost, but NCAM-inducing activity remained strong. Likewise, as discussed above, if the amino acid sequence FKP in the central domain of Xnr2 was mutated to ENA as in Xnr3 (chimera 11), MA inducing but not NCAM-inducing activity was reduced. Finally, if an epitope tag was added in a position known to disrupt Xnr3 activity, MA-inducing activity was retained, but NCAM-inducing activity was greatly reduced or absent (chimera 12).

Fig. 8 summarizes this work and shows the regions necessary for NCAM- or MA-inducing activity. The shaded areas required for activity are mutually exclusive. We show that a protein containing both domains has NCAM- and MA-inducing functions (chimera 10). We speculate that chimera 10 may be able to bind both as an agonist to the putative nodal receptor and as an antagonist to the BMP4 receptor. As suggested by chimera 11, mutation of critical residues in this middle portion of the protein may interfere with binding to the hypothetical nodal receptor and therefore prevent MA induction but without reducing NCAM induction. Conversely, the addition of an epitope tag in the thumb domain would block the inhibitory binding to the BMP receptor but not the activation of a nodal receptor.

The nodal-related proteins are a distinct subfamily of the TGF-βs. There is a single nodal protein in mouse, whereas duplications have led to multiple nodal-like genes in zebrafish and Xenopus, which have two and four nodal-related genes,
respectively. *Xenopus* has the most diverged family member, Xnr3. The structure and function of Xnr3 is unique among the nodal relatives, and a similar protein has yet to be found in any other animal. Perhaps Xnr3 was able to evolve its unusual characteristics because *Xenopus* has several redundant copies of the nodal gene. Xnr2 has an activity similar to the other nodals. However, this work has shown that the regions necessary for Xnr2 activity are different than those required by TGF-β itself and that the nodals may have evolved a different signaling strategy from other members of the superfamily.

Acknowledgments—We thank Chris Wright for the gift of the Xnr2 plasmid, Shannon Davis and Lisa Belluzzi for their comments on the manuscript, and Rolf Christoffersen for expert computer assistance.

REFERENCES

1. Heldin, C.-H., Miyazono, K., and ten Dijke, P. (1997) *Nature*** 390, 465–471
2. Creemers, J. W., Kormelink, P. J., Roebroek, A. J., Nakayama, K., and Van de W, J. (1995) *FEBS Lett.* 336, 65–69
3. Hosaka, M., Murakami, K., and Nakayama, K. (1994) *Biomed. Res.* 16, 383–390
4. Molloy, S. S., Bresnahan, P. A., Leppla, S. H., Klimpel, K. R., and Thomas, G. (1992) *J. Biol. Chem.* 267, 16396–16402
5. Griffith, D. L., Rebagliati, M. R., Toyama, R., and Dawid, I. B. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 515–524
6. O'Connell, M. L., Wright, C. V. E., Kuehn, M. R., and Dawid, I. B. (1999) *J. Mol. Biol.* 287, 687–702
7. Dale, C., Matthews, G., and Colman, A. (1993) *EMBO J.* 11, 395–401
8. Osada, S. I., and Wright, C. V. (1999) *Curr. Opin. Genet. Dev.* 9, 139–149
9. Gray, A. M., and Mason, A. (1999) *Science* 284, 1261–1271
10. Scheufler, C., Sebald, W., and Hulsmeyer, M. (1999) *J. Mol. Biol.* 287, 103–115
11. Jones, M., Hogan, B. L. M., Kuehn, M. R., Smith, J. C., and Wright, C. V. E. (1999) *Development* 126, 383–391
12. Zhou, X., Sasaki, H., Lowe, L., Hogan, B. L., and Kuehn, M. R. (1993) *Nature* 361, 543–547
13. Feldman, B., Gates, M. A., Egan, E. S., Dougan, S. T., Rennebeck, G., Sirotnik, H. I., Schier, A. F., and Talbot, W. S. (1998) *Nature* 395, 181–185
14. Rebagliati, M. R., Toyama, R., Haffter, P., and Dawid, I. B. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 9932–9937
15. Sambath, K., Ruheinstein, A. L., Cheng, A. M. S., Liang, J. O., Fekany, K., Solnica-Krezel, L., Korch, V., Halpern, M. E., and Wright, C. V. (1998) *Nature* 395, 185–189
16. Smith, W. C., Mc Kendrick, R. M., Ribisi, S., and Harland, R. M. (1995) *Cell* 82, 37–46
17. Joseph, E. M., and Melton, D. A. (1997) *Dev. Biol.* 184, 367–372
Primary Structure Requirements for *Xenopus* Nodal-related 3 and a Comparison with Regions Required by *Xenopus* Nodal-related 2
Carin Hansen Ezal, Christopher D. Marion and William C. Smith

*J. Biol. Chem.* 2000, 275:14124-14131.
doi: 10.1074/jbc.275.19.14124

Access the most updated version of this article at [http://www.jbc.org/content/275/19/14124](http://www.jbc.org/content/275/19/14124)

Alerts:
- When this article is cited
- When a correction for this article is posted

[Click here](http://www.jbc.org/content/275/19/14124.full.html#ref-list-1) to choose from all of JBC’s e-mail alerts

This article cites 51 references, 22 of which can be accessed free at [http://www.jbc.org/content/275/19/14124.full.html#ref-list-1](http://www.jbc.org/content/275/19/14124.full.html#ref-list-1)