Understanding how morphogens work

J. C. Smith1,*, A. Hagemann1, Y. Saka2 and P. H. Williams3

1Wellcome Trust/CR-UK Gurdon Institute, Department of Zoology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QH, UK
2Interdisciplinary Research Institute, Institut de Biologie de Lille, 1 rue du Professeur Calmette, BP447, 59021 Lille Cedex, France
3Vertebrate Development and Genetics, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1HH, UK

In this article, we describe the mechanisms by which morphogens in the Xenopus embryo exert their long-range effects. Our results are consistent with the idea that signalling molecules such as activin and the nodal-related proteins traverse responding tissue not by transcytosis or by cytonemes but by movement through the extracellular space. We suggest, however, that additional experiments, involving real-time imaging of morphogens, are required for a real understanding of what influences signalling range and the shape of a morphogen gradient.

Keywords: embryo; Xenopus; induction; morphogens; activin

1. INTRODUCTION

The specification of cell type in the early embryo frequently involves inductive interactions, during which one region of the embryo produces a signal that acts on adjacent cells and causes them to change their fates. One of the first inductive interactions during vertebrate development is mesoderm induction. In the amphibian embryo, this occurs during late blastula stages when a signal from the vegetal hemisphere of the embryo acts on overlying equatorial cells and causes them to become mesoderm rather than ectoderm (Heasman 2006). Several mesoderm-inducing factors have been identified in the embryo of the frog Xenopus laevis, and most have proved to be members of the transforming growth factor type β (TGF-β) family. These include activin (Asashima et al. 1990; Smith et al. 1990; Piepenburg et al. 2004), Vg1 (Weeks & Melton 1987; Dale et al. 1989; Thomsen & Melton 1993; Birsoy et al. 2006), the nodal-related proteins Xnr1, Xnr2 and Xnr4–6 (Jones et al. 1995; Joseph & Melton 1997; Takahashi et al. 2000) and derriere (Sun et al. 1999).

One of the most remarkable properties of these mesoderm-inducing factors is that they exert their effects in a strict concentration-dependent manner (Green & Smith 1990; Green et al. 1992; Gurdon et al. 1994; Papin & Smith 2000). In particular, low concentrations of activin, for example, induce the expression of genes such as Xbra (Smith et al. 1991), whereas higher concentrations activate genes such as goosecoid (Cho et al. 1991). This phenomenon is consistent with the suggestion that inducing factors such as activin and the nodal-related proteins establish a gradient, with a high concentration of the ‘morphogen’ near its source and lower concentrations further away (Gurdon et al. 1994; McDowell et al. 1997; Dyson & Gurdon 1998; McDowell & Gurdon 1999; Williams et al. 2004). The morphogen gradient would then activate gene expression in the appropriate spatial pattern, with genes being activated according to the local concentration of inducing factor.

There is now persuasive evidence that graded concentrations of inducing factors indeed establish specific patterns of gene expression, not only in Xenopus and other vertebrate species such as the zebrafish, but also in invertebrates such as Drosophila. For example, in Xenopus one can inhibit signalling by the nodal-related proteins by means of a truncated version of Cerberus called Cerberus-short (Agius et al. 2000). Increasing concentrations of Cerberus-short cause the progressive loss of Xbra expression in the embryo, beginning at the ventral marginal zone and extending dorsally (Agius et al. 2000). This observation is consistent with the idea that there is a graded distribution of nodal-related proteins in the embryo, with high levels on the dorsal side and lower levels ventrally. Additional experiments have involved injecting embryos with increasing concentrations of an antisense morpholino oligonucleotide directed against activin B; most sensitive to this treatment are genes such as goosecoid, which is expressed at the dorsal side of the embryo and which responds only to high concentrations of activin (Piepenburg et al. 2004).

In the zebrafish, the nodal-related protein Squint acts as a long-range morphogen (although the related protein Cyclops cannot), with different levels of nodal activity required for the formation of different mesodermal and endodermal tissues along the animal–vegetal axis of the embryo (Thisse et al. 2000; Chen & Schier 2001, 2002; Dougan et al. 2003). And in the Drosophila embryo, decapentaplegic (Dpp), a TGF-β family member most closely related to vertebrate bone morphogenetic protein 4 (BMP-4), can act as a morphogen to induce the expression of genes such as optomotor-blind and spalt (Lecuit et al. 1996; Nellen et al. 1996).
This brief introduction makes the point that spatial patterns of gene expression in the developing embryo are established, at least in part, by morphogens: molecules that are produced in a restricted region of the embryo and which then spread across a field of responding cells to activate gene expression in a concentration-dependent fashion. The important question, which this article addresses, is how morphogens exert their long-range effects, and we also touch upon the question of how cells interpret different concentrations of morphogens to activate the expression of different genes.

2. LONG-RANGE SIGNALLING: POTENTIAL MECHANISMS

The routes through which molecules exert long-range effects in the early embryo are likely to depend on the molecule itself, on the species of embryo, and on the nature of the responding tissue. For example, a small lipid-soluble molecule such as retinoic acid may traverse cells by a route that is different from that taken by a large protein such as activin; in a fast-developing species there may be little time to establish a morphogen gradient, so transmission of a signal may occur through a faster route than that employed in a slow-developing species; and transmission through an epithelium may involve different cellular mechanisms to transmission through mesenchymal tissues.

Several mechanisms have been proposed by which morphogens might traverse responding tissue. One of these involves simple diffusion, whereby molecules are presumed to move more-or-less freely in the extracellular space (Crick 1970). In practice, however, the extracellular space is a complex environment through which molecules might move with some difficulty, and mechanisms exist that might exacerbate or facilitate such movement. For example, the shape of a gradient created by such ‘restricted diffusion’ (Tabata & Takei 2004) might become steeper if cells expressed more cell surface receptors, because the receptors might trap and internalize the morphogen before it had a chance to spread far into the responding tissue. On the other hand, heparan sulphate proteoglycans (HSPGs) seem to play an important role in allowing the spread of such molecules across a field of responding cells (Tabata & Takei 2004; Zhu & Scott 2004). For example, the Drosophila gene Dally (division abnormally delayed) encodes a glycosylphosphatidylinositol- (GPI-) linked glypican that is required for the establishment of gradients of Wingless, Dpp and Hedgehog in the wing imaginal disc (Belenkaya et al. 2004; Han et al. 2004, 2005). The normal function of Dally appears to be to bind and stabilize morphogens such as Wingless in the vicinity of the cell from which they have been secreted, from which position it can move to more distal cells by disassociation and re-association with HSGPs (Han et al. 2005). In effect, therefore, Dally creates a passage along which Wingless can move. In the absence of Dally, Wingless remains associated with the cell membrane, diffuses into the wider extracellular milieu, or is degraded.

Another mechanism by which morphogens might exert long-range effects is via transcytosis, during which molecules secreted by one group of cells are endocytosed by responding tissue and then secreted again. Evidence in favour of this mechanism comes from experiments in Drosophila imaginal discs, where clones of cells deficient in dynamin prove to be unable to allow the passage of a green fluorescent protein-(GFP)-tagged form of Dpp (Entchev et al. 2000). It has been suggested, however, that there are increased numbers of cell surface receptors on such cells, and (as discussed above) that these receptors bind ligand and prevent its further progress through the responding tissue (Lander et al. 2002). Transcytosis as a mechanism for long-range signalling therefore remains an intriguing possibility, but one which is not yet proven.

Other mechanisms of long-range signalling have also been suggested. These include the possibility that lipoprotein particles act as extracellular vehicles for lipid-linked morphogens and GPI-linked proteins (Panakova et al. 2005), and ‘cytonemes’ (Ramirez-Weber & Kornberg 1999; Hsiung et al. 2005). Cytonemes are thin, actin-based extensions that project from receiving to signalling cell, and which allow direct, albeit long-range, contact between signalling and receiving cell. There is evidence that both of these mechanisms take place during Drosophila development, but the evidence that they occur during vertebrate development is less persuasive.

3. LONG-RANGE SIGNALLING IN XENOPUS

As described above, much of our understanding of the mechanism of long-range signalling comes from work on Drosophila; although Xenopus provides a powerful system for the analysis of morphogen function during vertebrate development, and especially for attempts to understand how different concentrations of morphogen activate the expression of different genes (Saka & Smith 2007), little is known about how inducing factors in this species exert long-range effects. Our attempts to understand this process have made use of GFP-tagged forms of the relevant inducing factors, together with real-time imaging of the signalling and responding tissues.

Members of the TGFβ family are synthesized as intracellular proproteins that dimerize and are then cleaved to produce the active secreted ligand. Work by the groups of Gonzalez-Gaitan (Entchev et al. 2000) and Cohen (Teleman & Cohen 2000) has shown that it is possible in the Drosophila TGFβ-family member Dpp to insert enhanced green fluorescent protein (EGFP) between the basic cleavage signal sequence and the mature region and for the resulting tagged protein to retain its biological activity. We therefore applied the same approach to Xenopus mesoderm-inducing factors, and found that although some, such as activin and Derrière, lost their activities, others, like Xnr2, retained their inducing activity (Williams et al. 2004).

Our EGFP-tagged form of Xnr2 proved to be secreted from Xenopus blastomeres and to be able to exert long-range effects. For example, when a small clone of Xnr2-expressing cells was created in the animal hemisphere of the embryo, expression of the target gene Xbra was detected in the surrounding

Phil. Trans. R. Soc. B (2008)
tissue, and when a lineage-labelled animal pole region expressing EGFP-Xnr2 was juxtaposed with a non-expressing region, target gene activation was induced in the latter tissue (Williams et al. 2004). It is unlikely that the long-range effects of EGFP-Xnr2 occur by a ‘relay’, in which Xnr2 activates the expression of another inducing signal, because expression of a constitutively active form of an Xnr2 receptor, ALK4, causes only cell-autonomous activation of target genes such as Xbra.

How does EGFP-Xnr2 traverse responding tissue? To address this question, we devised a technique in which animal pole regions were placed side by side on a fibronectin-coated cover-slip such that the two pieces of tissue could adhere to each other at one edge only. After healing, the tissue could be viewed through the cover-slip using the confocal microscope. In these experiments, one piece of tissue was derived from an embryo injected with RNA encoding EGFP-Xnr2 together with RNA encoding the fluorescent membrane marker cyan fluorescent protein- (CFP-)GPI; the other was derived from a normal embryo.

Observation of these tissue combinations revealed that tagged ligand could be detected in the vicinity of non-expressing cells within 30 min, and further experiments suggest that of the mechanisms described above, the most likely is that EGFP-Xnr2 travels through the extracellular space. Thus, the bulk of the EGFP-Xnr2 in the receiving animal pole region was extracellular and was detected in the interstices between cells, consistent with the idea that the ligand travels between the responding cells (figure 1).

In this respect, the distribution of EGFP-Xnr2 differs markedly from that of tagged Dpp in the Drosophila imaginal disc, where although some GFP-Dpp can be detected in the extracellular space, substantial amounts are also present in intracellular punctate structures (Entchev et al. 2000). Very few structures of this sort were observed in our experiments, indicating (but not, of course, proving) that transcytosis might not play a significant role in the transmission of Xnr2 in the early Xenopus embryo. If this is the case, it may be significant that the Drosophila imaginal disc is an epithelium with cells that are closely apposed at the lateral surfaces, perhaps favouring transcytosis, whereas Xenopus animal pole tissue is a more loosely packed mesenchymal tissue, perhaps favouring diffusive movement of morphogens.

Although one cannot rule out the possibility that cytonemes play a role in long-range signalling in the Xenopus embryo, this idea, like transcytosis, also appears unlikely. Receiving cells labelled with CFP-GPI do extend membrane-bound projections into their immediate environment, but none of these projections were observed to extend more than a single cell diameter.

Together, these results and others suggest that long-range signalling in the Xenopus embryo occurs by diffusion. This conclusion is consistent with the work by Gurdon and colleagues, who have interposed cells expressing a dominant-negative form of dynamin between activin-secreting cells and responding tissue; the inhibition of endocytosis did not prevent the activation of the activin signal transduction pathway in the responding cells (Kinoshita et al. 2006).

4. THE FUTURE

Our results so far suggest that morphogens in the Xenopus embryos traverse responding tissue through the extracellular space, and not through transcytosis or via cytonemes. But what controls the range of signalling or the shape of the morphogen gradient? As discussed above and elsewhere (Smith et al. 1982), one possibility is that components of the extracellular matrix can interact with signalling molecules and influence their distribution, half-life and biological function. Such interactions might occur in the case of BMP-4, a vertebrate member of the TGF-β family that is closely related to Dpp. The N-terminal region of BMP-4 contains a basic region of three amino acids that are responsible for the restricted range of action of this inducing factor; this amino acid core is believed to interact with HSPGs, and if the three basic amino acids (RRK) are substituted by alanine residues the range of the resulting molecule is substantially increased (Ohkawara et al. 2002). It is interesting that under

Figure 1. Two images of an animal cap conjugate in which cells on the left (with orange membranes) are expressing EGFP-Xnr2 (green) and cells on the right express no inducing factor. The dotted line represents the same position in each image: (a) time zero and (b) 20 min later. Note that EGFP-Xnr2 in the receiving tissue is almost exclusively extracellular, and has traversed 2–3 cell diameters over the course of the experiment.
certain circumstances, such as these, the range of a morphogenetic signal can be restricted by interaction with HSPGs, whereas interaction of morphogens with Dally can increase the range (see above).

The effective range of a morphogen can also be affected by the expression of inhibitors of their activities. The activity of BMP-4, for example, is inhibited by follistatin, chordin and noggin (Khhoka et al. 2005), all of which bind the ligand and prevent it from interacting with its receptor. At least for some of these inhibitors it is possible that they establish a ‘reverse gradient’, thereby establishing an effective range from interacting with its receptor. At least for some of inhibitors it is possible that they establish a reverse gradient, thereby establishing an effective range of activities. The activity of BMP-4, for example, is affected by the expression of inhibitors of their morphogenetic signal can be restricted by interaction with the extracellular matrix, or particular inhibitors, by RNA oligonucleotides (Heasman et al. 2000). It is clear that inhibitors of morphogen action are at least as important as the morphogens themselves in establishing spatial patterns of gene expression.

How to study the establishment and regulation of morphogen gradients, and indeed how to study the way in which they activate the expression of different genes at different concentrations? We believe it is essential to observe the setting up of morphogen gradients, and to observe their effects, in real time. We have taken the first steps to this end with our analysis of EGFP-Xnr2 movement and distribution in the early embryo, and we now plan to continue this approach by improving our imaging technologies and by manipulating the environment through which the morphogen moves. For example, one can overexpress components of the extracellular matrix, or particular inhibitors, by RNA overexpression in the embryo, or such proteins can be ‘knocked down’ by use of antisense morpholino oligonucleotides (Heasman et al. 2000). Quantitative analysis of the effects of these manipulations will reveal a great deal about how morphogens spread through responding tissue.

In addition to studies of this sort, it will be necessary to understand how different morphogen concentrations activate the expression of different genes. One approach to this problem is to follow the events that occur after the interaction of the morphogen with the responding cell. At present, our efforts focus in two areas. The first involves following the signal transduction pathway that is used by activin and members of the TGF-β family and the second involves following gene activation in situ and in real time. TGF-β family members signal by binding to serine-threonine kinase type-II receptors that then associate with, and phosphorylate, type-I receptors. The activated type-I receptors go on to phosphorylate receptor-regulated Smad proteins (the so-called R-Smads) and these associate with the common mediator Smad3 and Smad4. The Smad complexes accumulate in the nucleus and regulate the expression of downstream target genes, often by interacting with a specific cofactor.

Previous work has used GFP-tagged forms of Smad2 to monitor TGF-β signalling: the accumulation of GFP-Smad2 in the nucleus of responding cells indicates that cells have experienced signalling by a member of the TGF-β family (Grimm & Gurdon 2002; Williams et al. 2004). Although this approach has proved very useful, we are now asking whether the technique of bimolecular fluorescence complementation (Hu et al. 2002; Kerpole 2006) might provide a better signal-to-noise ratio. To achieve this, we plan to fuse the N-terminal half of EGFP to the N-terminus of Smad2, and the C-terminal half of EGFP to the N-terminus of Smad4. In the absence of signalling by members of the TGF-β family, there should only be background fluorescence. If cells receive such signals, Smad2 should associate with Smad4, thereby causing the reconstitution of functional EGFP and the appearance of nuclear fluorescence.

There are no reliable techniques for visualizing the real-time activation of endogenous gene expression, although it is possible to drive the expression of reporter genes such as GFP by the promoter of a gene of interest, such as Xbra. One potential approach, involving ‘molecular beacons’ (Bratu et al. 2003) has not proved successful in our hands, and we are now investigating a strategy based on EGFP complementation in which the two halves of EGFP are fused to two Pumilio homology domains that are targeted to adjacent sequences on the RNA of interest (Ozawa et al. 2007). If successful, this approach will allow us to test the predictions of models that attempt to explain the mechanism by which different concentrations of morphogens activate the expression of different genes. For example, our recent work predicts that the low concentrations of activin that induce stable expression of Xbra should nevertheless cause the transient induction of goosecoid shortly after exposure to the inducing factor; and high doses of activin should cause a similar transient induction of Xbra (Saka & Smith 2007). A previous analysis suggests that this may be the case (Papin & Smith 2000), but confirmation requires a more careful time course and the simultaneous observation of the two gene products in the same cell.

REFERENCES
Understanding how morphogens work J. C. Smith et al. 1391


Grimm, O. H. & Gurdon, J. B. 2002 Nuclear exclusion of Smad2 is a mechanism leading to loss of competence. *Nat. Cell Biol.* 4, 519–522. (doi:10.1038/ncb8812)


Lecuit, T., Brook, W. J., Ng, M., Calleja, M., Sun, H. & Cohen, S. M. 1996 Two distinct mechanisms for long-range patterning by decapentaplegic in the *Drosophila* wing. *Nature* 381, 387–393. (doi:10.1038/381387a0)


Phl. Tran. R. Soc. B (2008)


