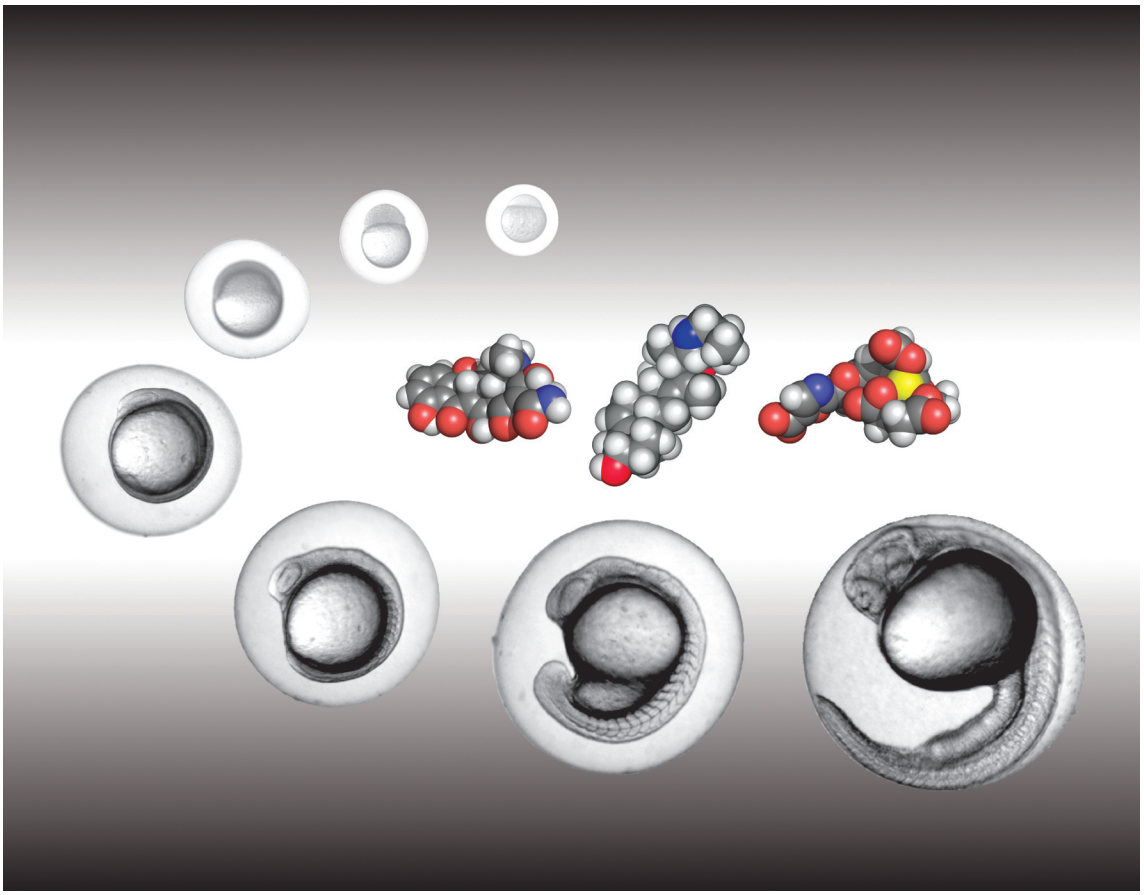
Chem Soc Rev

This article was published as part of the

2008 Chemistry–Biology Interface Issue

Reviewing research at the interface where chemistry meets biology

Please take a look at the full table of contents to access the other papers in this issue

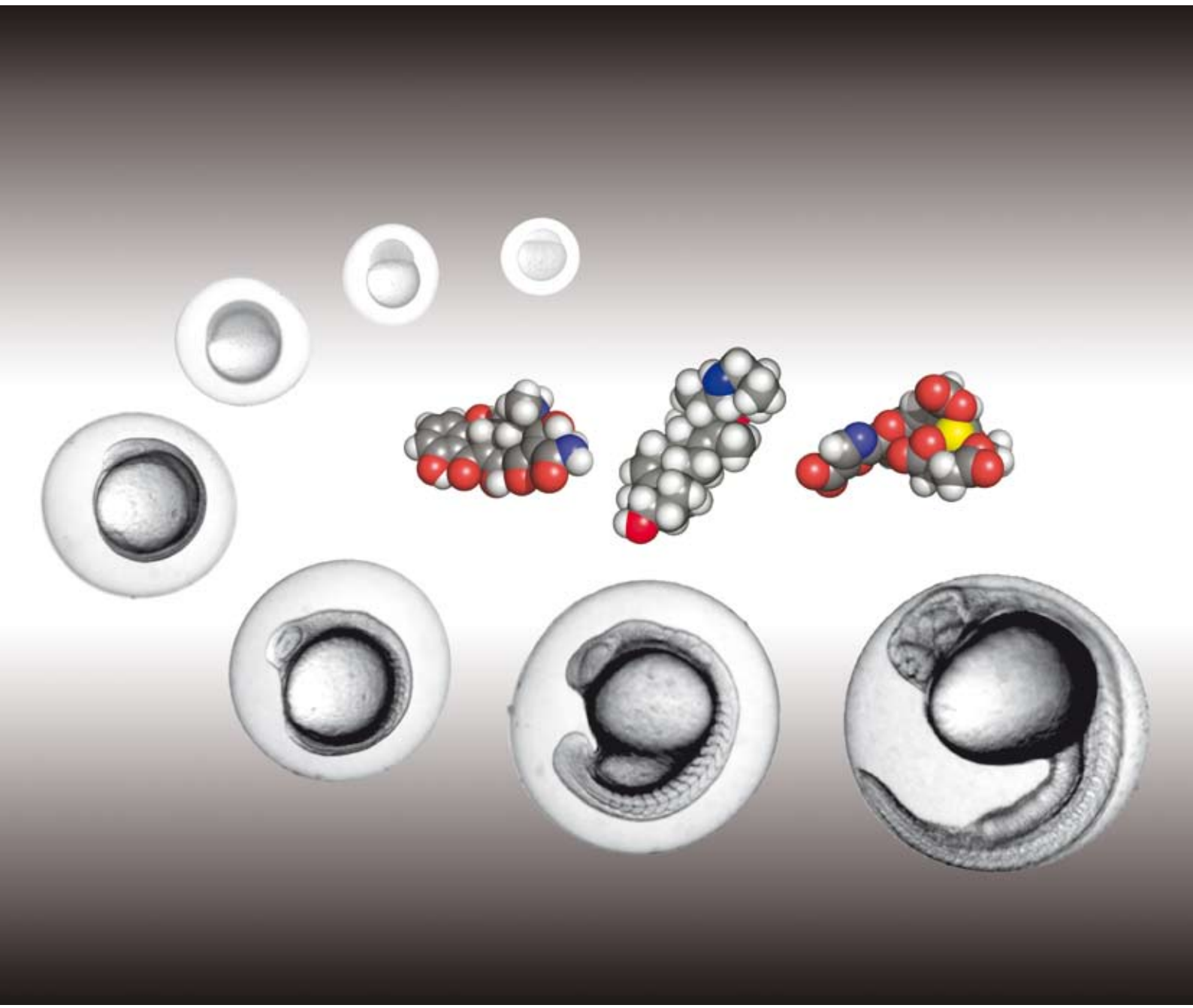


Chem Soc Rev

Chemical Society Reviews

www.rsc.org/chemsocrev

Volume 37 | Number 7 | July 2008 | Pages 1281–1452



ISSN 0306-0012

RSC Publishing

THEMATIC ISSUE: CHEMISTRY–BIOLOGY INTERFACE
Guest editor: David Spring

0306-0012(2008)37:1;1-C

Chemical technologies for probing embryonic development†

Ilya A. Shestopalov and James K. Chen*

Received 3rd February 2008
First published as an Advance Article on the web 7th May 2008
DOI: 10.1039/b703023c

Embryogenesis is a remarkable program of cell proliferation, migration, and differentiation that transforms a single fertilized egg into a complex multicellular organism. Understanding this process at the molecular and systems levels will require an interdisciplinary approach, including the concepts and technologies of chemical biology. This *tutorial review* provides an overview of chemical tools that have been used in developmental biology research, focusing on methods that enable spatiotemporal control of gene function and the visualization of embryonic patterning. Limitations of current approaches and future challenges are also discussed.

1 Introduction

How a single fertilized egg gives rise to the human form is one of the great mysteries of biological science. Each zygote must generate distinct cell types in a spatially and temporally controlled manner, ultimately assembling complex organs with specific structures and functions. For centuries, our understanding of this remarkable process was largely descriptive, starting with Aristotle's account of chicken (*Gallus gallus*) embryogenesis in *Generation of Animals*. The scientific tradition of embryological observation has since expanded to include several invertebrate and vertebrate model organisms, such as the sea urchin (*Strongylocentrotus purpuratus*, *Lytechinus variegates*, and other species), the worm (*Caenorhabditis elegans*), the fruitfly (*Drosophila melanogaster*), the fish (*Danio rerio* and *Oryzias latipes*), the frog (*Xenopus laevis* and *Xenopus tropicalis*), and the mouse (*Mus musculus*). While these metazoans appear dissimilar in adult structures and forms, their embryos share some common features with the developing human fetus. For example, embryonic cells in these organisms segregate during a process called gastrulation to form three "germ" layers (the ectoderm, the mesoderm, and the endoderm) and become asymmetrically patterned with respect to each body axis (*e.g.* anterior–posterior, dorsal–ventral, and left–right). The differentiation of each germ layer into specialized tissues is also grossly conserved, with epidermis and nerves arising from the ectoderm, muscle from the mesoderm, and the digestive system from the endoderm.

Using these model systems, embryologists have interrogated the origins of multicellular pattern and function through a "perturb and observe" paradigm. Perturbation strategies have varied with each organism, according to their amenability to available technologies. Prior to the advent of modern genetics and molecular biology, embryological manipulations were primarily limited to surgical procedures and the labeling of certain cell populations with visible dyes, favoring the study of embryos that develop *ex utero* and are readily dissected. At the turn of the 20th century, cell dissociation and transplantation experiments with sea urchin embryos helped establish fundamental concepts such as regulative development, inductive interactions between cells, and the existence of morphogen

Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA 94305, USA.
E-mail: jameschen@stanford.edu; Tel: +1 (650) 725-3582
† Part of a thematic issue examining the interface of chemistry with biology.

Ilya A. Shestopalov

Ilya Shestopalov received his BS degree in Biological Chemistry from the University of Chicago in 2004. As an undergraduate he worked with Rustem Ismagilov to engineer a microfluidic system for nanoparticle synthesis, and as a graduate student at Stanford University he has worked with Dr Chen to develop photoactivatable morpholinos that enable spatiotemporal control of embryonic gene expression. He is currently utilizing the caged morpholino system to investigate roles of T-box transcription factors in vertebrate embryogenesis.

James K. Chen

James Chen received his AB and PhD degrees in Chemistry and Chemical Biology from Harvard University, working with George Whitesides and Stuart Schreiber, respectively. After postdoctoral studies with Philip Beachy at Johns Hopkins School of Medicine he joined the Department of Chemical and Systems Biology at Stanford University School of Medicine as an Assistant Professor. His research interests include developmental signaling pathways and new technologies for probing the molecular mechanisms of embryogenesis.

gradients. Subsequent transplantation studies using frog embryos led to the discovery of cellular "organizing" domains, named the Nieuwkoop center and Spemann organizer, that establish the dorsal–ventral axis during gastrulation, and surgical procedures on chick embryos revealed specific tissue structures that regulate limb patterning. Other insights were obtained from spontaneous mutants with embryonic defects, such as the *antennapedia* fruitfly which has ectopic legs on its head rather than antennae and the *talpid* chicken which has extra toes and craniofacial abnormalities.

Elucidation of the genetic code and development of molecular biology techniques transformed embryology from a descriptive science to one deeply immersed in molecular mechanism. Large-scale mutagenesis screens pioneered by Nüsslein-Volhard and Wieschaus in the 1980s yielded several hundred fruitfly mutants with distinct developmental abnormalities,[1] and the positional cloning of these mutated genes has yielded many of the key molecules that regulate embryonic patterning. What has emerged from these and subsequent studies is that a relatively limited number of signaling mechanisms—such as the Hedgehog (Hh), Wnt, transforming growth factor-β (TGF-β), bone morphogenetic protein (BMP), fibroblast growth factor (FGF), and Notch pathways—are used iteratively throughout development to coordinate cell proliferation, migration, and differentiation. We now know that Wnt pathway activation in a subset of cells within the frog blastula (the embryo prior to gastrulation) establishes the Nieuwkoop center, which in turn induces the Spemann organizer. The organizer then secretes inhibitors of BMP signaling to promote dorsal cell fates. Ectopic expression of Wnt proteins or BMP antagonists in the frog embryo can consequently induce a secondary body axis, recapitulating the phenotypes observed upon organizer transplantation. Similarly, digit identity in the developing limb bud is controlled by a gradient of Hh pathway activation established by a posterior domain of Hh ligand-secreting cells, and ectopic Hh protein expression in the nascent chick wing can cause dramatic mirror-image duplications of this structure.

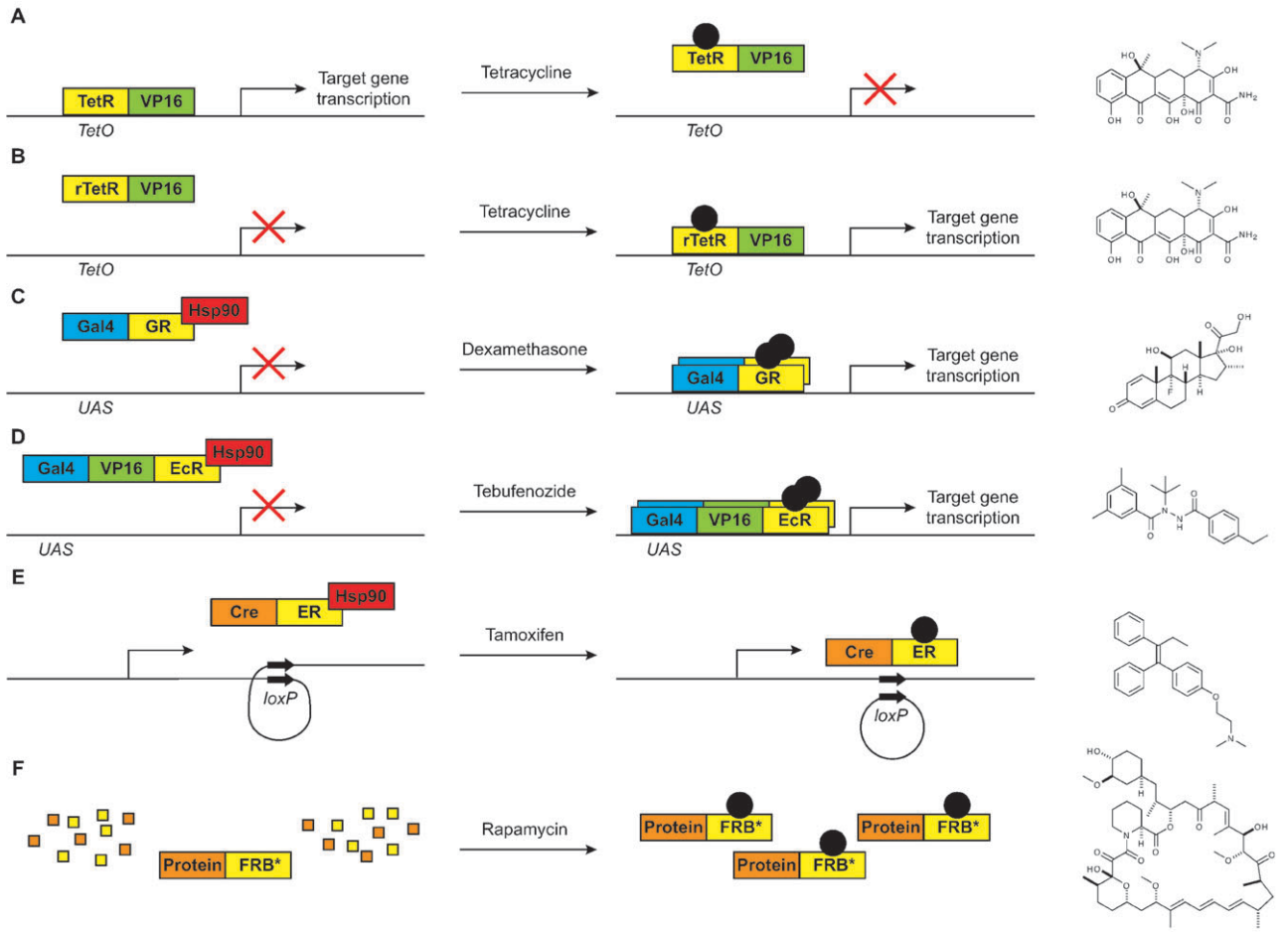
The recent completion of multiple genome sequencing projects promises to further revolutionize our understanding of embryonic development, since we now have a comprehensive parts list for Nature's genetic circuitry. Determining how these genes orchestrate embryogenesis at the molecular and systems levels is the challenge that lies ahead, and realizing this goal will require the expertise of multiple scientific disciplines. Genetic approaches will certainly continue to make significant contributions to this effort, building upon current technologies for controlling embryonic gene function, generating mutant and transgenic organisms, and analyzing spatiotemporal changes in gene expression levels. Computational approaches will be necessary to understand how diverse developmental signaling mechanisms integrate to create specific morphological outcomes. While currently under-represented in this initiative, chemists have an important role to play in embryological research as well. Our ability to synthesize novel compounds can create new ways to "perturb and observe" embryological processes, circumventing the limitations of Nature's building blocks. In addition, our intuition about chemical structure and reactivity can bring a unique perspective to the molecular mechanisms of embryogenesis. This tutorial review will summarize areas in which chemical concepts and technologies have advanced our knowledge of embryonic patterning, in the hope that these examples will guide and inspire other chemists to explore developmental biology. We will first focus on chemical methods that can alter embryonic gene expression or function with spatiotemporal control, enabling the interrogation of complex patterning mechanisms. Complementary tools for visualizing the molecular and cellular processes that constitute embryogenesis will then be discussed, as well as future research opportunities in chemical embryology.

2 Ligand-dependent strategies for conditional gene regulation

As exemplified by the Nieuwkoop center and Spemann organizer studies described above, the misexpression of developmental genes is a common strategy for determining *in vivo* function. In this particular case, the Wnt proteins and BMP antagonists were introduced into individual cells (also called blastomeres) of the early frog embryo by mRNA microinjection, taking advantage of the physical accessibility and relatively large size of these cells. Since this approach is not universally applicable to other embryonic tissues and lacks conditionality, alternative strategies for manipulating exogenous or endogenous gene expression have been developed for embryological studies. Several of these technologies are derived from naturally occurring small molecule-dependent transactivators or repressors, enabling chemically controlled gene expression (Scheme 1).

Establishing these systems for embryological studies requires methods for introducing the necessary genetic and chemical components into the developing organism. Embryos that are amenable to microinjection techniques, such as those from sea urchins, fruitflies, worms, frogs and fish, can be injected with mRNAs or cDNAs encoding the necessary genetic factors. Oligonucleotides can also be electroporated, as is commonly performed in chicken embryos. The transient nature of these approaches, however, restricts their utility. Gene expression conveyed by exogenous mRNAs typically persists for only one day, and although cDNAs exhibit greater perdurance, their mosaic distribution during embryogenesis complicates the interpretation of resulting phenotypes. As a result, certain applications of these chemical technologies may require the generation of transgenic organisms that stably express the ligand-responsive machinery.

Introduction of the chemical ligand is more straightforward. Small molecule perturbations of embryos that develop *ex utero* can usually be achieved by compound addition to the culture medium. For embryos that develop *in utero*, pregnant adults can be treated with the chemical ligand by intraperitoneal injection, oral gavage, or solubilization in drinking water. Through these approaches, developmental biologists have been able to spatially and temporally regulate gene expression during embryogenesis, in some cases even with dosage control. These experimental capabilities have provided insights into *in vivo* gene function that cannot be discerned from conventional knockout or mutant organisms.



Scheme 1 Ligand-regulated gene expression systems. (A–F) Schematic representations of Tet-off, Tet-on, Gal4-GR, GV-EcR, Cre-ER and FRB* systems. Small molecules used in each of these technologies are shown to the right of each diagram.

2.1 Tetracycline-regulated gene expression

The most commonly used ligand-inducible gene expression system exploits an antibiotic resistance mechanism found in *Escherichia coli*. When this bacterium is challenged with tetracycline, a transmembrane efflux pump called TetA is expressed to clear the antibiotic from the bacterial cytosol. This small molecule-induced transcriptional activation is mediated through a repressor protein (TetR), which binds to tetracycline with nanomolar affinity. In the absence of antibiotic, TetR recognizes a palindromic tetracycline operator (*tetO*) sequence in *tetA* regulatory elements with a dissociation constant of approximately 10^{-11} M. Upon complexation by tetracycline, however, TetR undergoes a conformational change that reduces its affinity for *tetO* elements by several orders of magnitude. The resulting dissociation of TetR from *tetO* sites promotes *tetA* transcription and tetracycline clearance from the bacterium.

The TetR/*tetO* system therefore can be viewed as a chemically gated transcriptional switch, composed of a small molecule-dependent transcription factor, its chemical ligand, and a specific regulatory element. Since the TetR protein and *tetO* operator are absent from eukaryotic genomes, tetracycline-regulated transcription can in principle be used to interrogate embryonic patterning mechanisms with minimal off-target perturbations. Toward that goal, two tetracycline-dependent expression systems have been developed for *in vivo* studies. In the "Tet-off" strategy, the DNA- and ligand-binding domains from TetR are fused with an activation domain from the herpes simplex virus protein VP16, creating a tetracycline-controlled transcriptional activator (tTA), which is a functional inverse of TetR (Scheme 1(A)).[2] Cells that express tTA and contain a *tetO*-dependent transgene therefore constitutively express the targeted gene, and the addition of tetracycline or its derivative doxycycline abrogates this transcription. Tetracycline-activated gene expression can also be achieved using the "Tet-on" system, which utilizes a tTA variant (rtTA) that has been mutagenized to reverse the effects of ligand binding (Scheme 1(B)).[3] Rather than dissociating from regulatory elements in the presence of tetracycline, rtTA only binds *tetO* sequences when it is complexed with the antibiotic.

The power of tTA/rtTA systems for *in vivo* studies is perhaps best exemplified by their use to investigate endothelin receptor B (Ednrb)-dependent patterning in the mouse embryo.[4] In this study, Tilghman and co-workers integrated *tetO* elements into the endogenous *Ednrb* locus and coding sequences for either tTA and rtTA into the other allele. This approach restricts transactivator function in compound heterozygous embryos (genotypes $Ednrb^{tTA}/Ednrb^{tetO}$ and $Ednrb^{rtTA}/Ednrb^{tetO}$) to cells that would normally express this G protein-coupled

receptor. Inducible *Ednrb* expression is therefore spatially identical to that of the endogenous gene, with doxycycline providing temporal control of transcription. By treating these transgenic mice with doxycycline at different developmental stages, the Tilghman laboratory revealed that *Ednrb* is required between embryonic days 10 and 12.5 (abbreviated E10 and E12.5) for neural crest cell development. Originating from the dorsal region of the developing spinal cord at E9 and then migrating extensively by E10.5, these pluripotent, migratory cells give rise to the peripheral nervous system, craniofacial bones and cartilage, smooth muscle cells, and melanocytes. *Ednrb* function therefore appears to be required for development of migratory neural crest cells and their derivatives, but not for their initial emergence or post-migratory survival.

Although tetracycline-regulated gene expression has been widely used in mice, few applications in other metazoans have been reported. This may reflect limitations in the generality of this approach, perhaps arising from differences in compound bioavailability between species. For example, although tTA can induce *tetO*-regulated transgene expression in frog embryos, this activity cannot be blocked by adding tetracycline to the culture medium.[5] Rather, the antibiotic must be injected into the embryo, indicating that the compound cannot efficiently traverse the embryonic membrane or vitelline envelope. Since ligand introduction by microinjection is practical for only early-stage embryos, the utility of tetracycline-regulated transactivators in frogs is quite limited. Efforts to implement the rtTA system in zebrafish have been more successful, as Tsai and co-workers recently reported doxycycline-activated green fluorescent protein (GFP) expression in cardiac myocytes.[6] However, GFP expression in these transgenic fish fails to plateau even after four days of doxycycline treatment, suggesting that antibiotic penetrance into the zebrafish embryo is also rate-limiting. Since zebrafish organogenesis is largely complete within two days after fertilization, the kinetics of tetracycline action in this organism may preclude its use to interrogate embryological questions.

2.2 Dexamethasone-regulated gene expression

Nuclear hormone receptors represent another class of natural small molecule-modulated transactivators, mediating transcriptional responses to steroids and other endogenous lipids. For example, the glucocorticoid receptor (GR) is a modular transcription factor that is normally sequestered by a cytosolic complex containing heat shock protein 90 (Hsp90) and therefore is transcriptionally inactive. The binding of cortisol or other glucocortioids to GR induces a conformational change that dissociates the transactivator from this complex, allowing it to translocate to the nucleus, bind glucocorticoid response elements in genomic DNA, and activate the transcription of targeted genes. As in other nuclear hormone receptors, each of these GR activities is mediated by a discrete polypeptide module, including a ligand-binding domain, a DNA-binding domain, and an activation domain. This modularity allows the GR ligand-binding domain to regulate protein function in *cis*, and heterologous transcription factors fused to this polypeptide can be used to achieve ligand-dependent transgene expression in embryos.

Kolm and Sive demonstrated this principle by fusing the GR ligand-binding domain to MyoD, a basic helix-loop-helix transcription factor that is a potent inducer of muscle cell differentiation.[7] Frog embryos injected with mRNA encoding this transactivator (MyoD-GR) developed normally; however, the addition of the synthetic GR agonist dexamethasone to the culture medium induced ectopic muscle cells as determined by genetic markers. This strategy can be modified to enable dexamethasone-dependent expression of any gene by coupling the GR ligand-binding domain with another transcriptional activator, preferably one that is not endogenous to metazoans. The *Saccharomyces cerevisiae* transcription factor Gal4 is an ideal candidate for this approach, as it recognizes an upstream activating sequence (UAS) that is unique to yeast. Orthogonality of the Gal4/UAS system to metazoan biology has been further demonstrated by its extensive use to interrogate gene function in fruitfly embryos.[8] Zivkovic and colleagues therefore generated a chimeric transactivator by replacing the GR DNA-binding domain with that of Gal4 (Gal4-GR; Scheme 1(C)).[9] After injecting zebrafish zygotes with plasmids encoding the Gal4-GR transactivator and constructs for UAS-driven *XactivinβB* or *Xwnt* transcription, they were able to activate the expression of these morphogens with dexamethasone. Patterning defects associated with the ectopic expression of these genes were observed, illustrating the potential of this technology.

Several other ligand-inducible transcription factors have been created using the GR/dexamethasone system, demonstrating the generality of this method. One limitation of GR-based technologies, however, is the potential for interactions between endogenous GR signaling and the exogenous transactivator system. For example, exogenous glucocorticoid agonists can alter pituitary gland development, tissue regeneration, and other physiological processes in zebrafish.[10,11]

2.3 Ecdysone- and dibenzoylhydrazine-regulated gene expression

Crosstalk between small molecules and endogenous steroid hormone receptors could be minimized by using transactivators from evolutionarily distant species. For studies of vertebrate embryology, technologies based upon the insect-specific ecdysone receptor (EcR) therefore might enable transgene regulation without off-target effects. EcRs are activated by steroid hormones such as 20-hydroxyecydsone to regulate molting and metamorphosis, and as with the GR, the apoform of EcR is localized to the cytosol by macromolecular chaperones and is transcriptionally inactive. Upon ligand binding, the EcR heterodimerizes with another hormone receptor called Ultraspiracle, which is a member of the retinoid X receptor (RXR) family. This complex then translocates into the nucleus and initiates the transcription of target genes by binding to ecdysone response elements in promoter regions. Building upon their discovery of this natural transcriptional switch, Evans and co-workers developed an inducible gene expression system for mammalian cells by replacing the EcR activation domain with that from VP16 and co-expressing this transactivator (V-EcR) with mammalian RXR.[12] Using this chimeric protein, they successfully regulated gene expression

in adult transgenic mice by injecting them with muristerone A, a plant-derived ecdysone derivative.

The limitations of the tetracycline and Gal4-GR systems for frog or zebrafish studies, prompted our laboratory to modify this EcR-based technology further to enable zebrafish models with conditional transgene expression.[13] One possible drawback of the V-EcR system is its requirement for RXR co-activators; the co-expressed exogenous RXR could perturb embryonic signaling since this hormone receptor has pleiotropic roles during development. In addition, tissue-specific differences in endogenous RXR expression might influence V-EcR function. To minimize these issues, we investigated an EcR-based transactivator initially used to control transgene expression in plants,[14] which lack any endogenous RXR homologs. This chimeric protein (GV-EcR) is composed of the Gal4 DNA-binding and homodimerization domains, the VP16 activation domain, and the EcR ligand-binding domain (Scheme 1(D)). After evaluating several GV-EcR isoforms that contain different regions of the VP16 domain, we identified an optimal transactivator that induces the expression of UAS-dependent transgenes in zebrafish embryos in response to EcR agonists. In this case, we found that non-steroidal EcR ligands such as the dibenzoylhydrazine derivative tebufenozide were much more effective than ecdysone derivatives, possibly reflecting structural differences between the ligand-transactivator complexes. The kinetics of tebufenozide-dependent transgene expression appears to be significantly faster than that of the tetracycline-based system, and since our original report of this technology, we have demonstrated its efficacy in transgenic fish by inducing heart-specific GFP expression in response to tebufenozide-containing medium (J. K. Chen, unpublished observations).

2.4 Tamoxifen-regulated genetic recombination

In principle, the ligand-binding domains of nuclear hormone receptors could be used to convey small molecule-dependence to a variety of protein functions. Due to the cytosolic localization of the unliganded domains, factors that require nuclear localization for their activity are especially amenable to this strategy. McMahon and co-workers have extended this approach to DNA-modifying enzymes, achieving small molecule-induced genomic recombination.[15] These studies utilized the bacteriophage P1 recombinase Cre, which catalyzes site-specific recombination between tandem 34-base pair *loxP* sequences. DNA sequences that are flanked by *loxP* sites are efficiently excised by Cre, and integration of a *loxP*-flanked cassette into regulatory elements can therefore effect Cre-dependent gene expression. The McMahon laboratory generated an inducible Cre recombinase by fusing this enzyme to the estrogen receptor (ER) ligand-binding domain, taking advantage of an ER mutant that responds to the synthetic agonist 4-hydroxytamoxifen but not the endogenous ligand 17β-estradiol (Scheme 1(E)). Transgenic mice that specifically express this chimeric protein (Cre-ER) in the developing central nervous system were crossed with mice that contained a β-galactosidase reporter disrupted by a *loxP*-flanked insert, and the resulting pregnant mice were injected with either 4-hydroxytamoxifen or vehicle alone. While embryos containing both transgenes exhibited only background β-galactosidase activity in the absence of 4-hydroxytamoxifen, strong β-galactosidase activity was observed in the developing brain and spinal cord upon ligand treatment.

The Cre-ER system can also be used to conditionally knockout gene function, as illustrated by a functional analysis of the *Otx2* homeobox gene by Lamonerie and colleagues.[16] In this study, one *Otx2* allele was replaced with the *Cre-ER* gene, while *loxP* sites were integrated into the other locus such that they flanked the second exon. $Otx2^{loxP}/Otx2^{loxP}$ and $Otx2^{loxP}/Otx2^{Cre-ER}$ mice were crossed, and the pregnant females were injected with 4-hydroxytamoxifen at different developmental stages to induce cell-specific loss of *Otx2* expression. Since Cre recombinase expression is controlled by the *Otx2* promoter, excision of the *Otx2* exon occurs only in cells expressing this transcription factor at the time of 4-hydroxytamoxifen treatment. Through this systematic analysis, the temporal windows during which Otx2 patterns craniofacial and midbrain structures were determined to be E8.5 to E10.5 and E10.5 to E16.5, respectively.

2.5 Ligand-controlled protein degradation

An alternative approach to controlling gene expression levels is to chemically regulate protein degradation rather than transcriptional mechanisms. The Wandless and Crabtree laboratories pioneered this strategy after fortuitously observing that cellular proteins fused to a mutant form of the FK506-binding protein/rapamycin-binding domain (FRB*) rapidly degrade in a manner than can be rescued upon addition of FRB* ligands such as rapamycin or its analog 20-methyallyl-rapamycin (Scheme 1(F)).[17] Crabtree and Longaker then collaboratively applied this technology to study glycogen synthase kinase-3β (GSK3β) function *in utero*, using mice that have an FRB* tag inserted into the 3′ end of the GSK3β locus.[18] Mouse embryos homozygous for this mutant allele had greatly diminished levels of GSK3β expression and displayed cleft palate and midline skeletal defects, consistent with the role of GSK3β in multiple developmental signaling pathways (see Section 4). However, GSK3β levels in these embryos could be restored to wildtype levels upon rapamycin treatment after E13.5 (earlier treatments were teratogenic since rapamycin potently inhibits the serine/threonine kinase TOR). By varying the timing of rapamycin administration, the GSK3β mutant phenotype could be rescued to different extents, revealing distinct temporal windows for GSK3β-dependent formation of the palate and midline skeleton. The Wandless laboratory has since developed new protein-destabilizing domains by engineering mutants of the FK506-binding protein FKBP12.[19] These domains can be stabilized by the small molecule SLF* (also called Shield-1), which may exhibit lower inherent teratogenicity than rapamycin and therefore have greater versatility for *in vivo* studies.

While ligand-controlled protein degradation has demonstrated utility in mouse models, its applicability to other organisms has not yet been established. In the case of GSK3β-FRB*, protein levels were restored to that of endogenous GSK3β within 34 h of compound treatment, which may restrict its use in more rapidly developing animals such as

frogs and fish. Other genes will exhibit different kinetics, since ligand-induced stabilization is rapid and the synthesis of new protein is the rate-limiting step. Thus, experimental parameters for this technology will need to be determined in a gene- and species-specific manner.

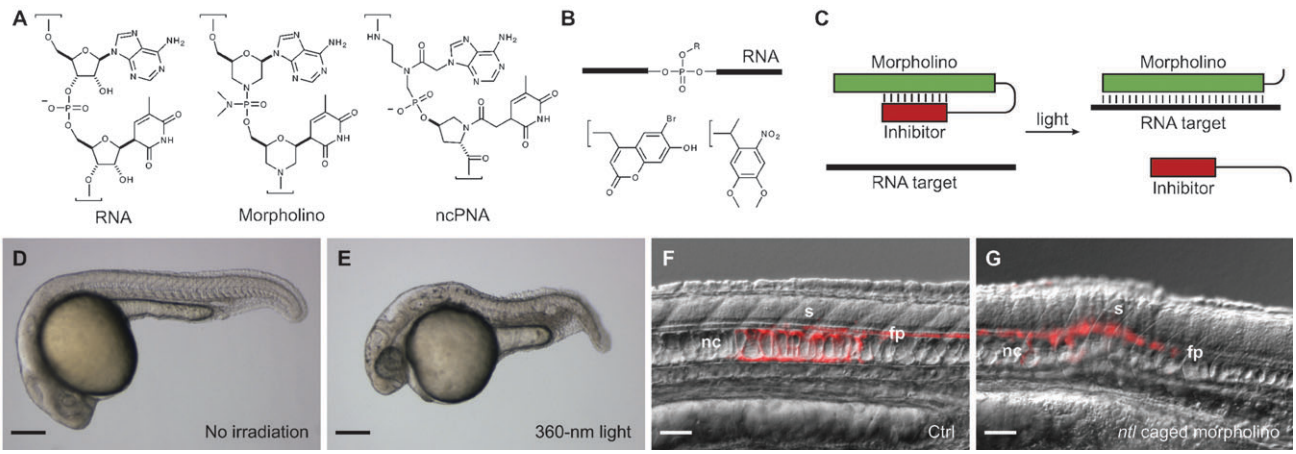
3 Caged-molecule strategies for conditional gene regulation

While ligand-dependent effectors can restrict gene expression to a defined temporal window, they rely upon tissue-specific promoters to enable spatial control of gene function. When such regulatory sequences have been identified, they can be used to dynamically recapitulate the expression patterns of endogenous genes. However, targeted cell populations within the embryo do not always coincide with known enhancer elements. Chemical approaches provide an opportunity to bypass these constraints, allowing developmental biologists to direct genetic perturbations according to morphological cues alone. The ability to focus light with single-cell precision makes caging strategies with photocleavable groups particularly attractive (Scheme 2). In this paradigm, latent molecules would be introduced into every cell of the developing embryo, and light-directed uncaging of these factors would localize their activities to selected tissues. If the activated molecules are membrane-impermeant, they would also function in a cell-autonomous manner. Embryos that have been studied using caged reagents include those of the fruitfly, albino frog, and zebrafish, with the latter emerging as the model of choice for optical perturbations and imaging. These organisms are perhaps best suited for caged-molecule applications, since they develop *ex utero* and have sufficient transparency during embryogenesis. In addition, the amenability of their embryos to microinjection techniques facilitates the delivery of membrane-impermeant caged reagents.

3.1 Caged Gal4-VP16

One of the first attempts to achieve photoactivatable gene expression targeted the hybrid transcription factor Gal4-VP16, which contains the Gal4 DNA-binding and homodimerization domains and the VP16 activation domain.[20] Cambridge and co-workers functionalized bacterially expressed Gal4-VP16 with 4,5-dimethoxy-2-nitrobenzyl chloroformate, presumably modifying lysines in the transcription factor in a stochastic manner. Injection of this caged Gal4-VP16 protein into fruitfly embryos containing a UAS-β-galactosidase transgene enabled the light-induced expression of β-galactosidase with single-cell resolution. While this study focused on cell lineage analyses, the caged Gal4-VP16 reagent could be used in conjunction with the variety of UAS-transgene lines that have been generated in the developmental biology community. Caged Gal4-VP16 was also reported to be an effective gene expression tool in frog and zebrafish embryos that contain UAS-driven transgenes.[20]

Although these initial findings are promising, there appear to be some limitations to this approach. Activation of Gal4-VP16-mediated gene expression in fruitfly embryos is restricted to the first eight hours after cellularization, and frog and zebrafish experiments were also conducted at early developmental time points. This narrow temporal window may reflect the *in vivo* instability of this reagent. Gal4-VP16 caging also requires modifications to over half of the 14 lysines in the protein, with conjugation ratios fluctuating between each preparation. Differing levels of residual protein activity and variable uncaging efficiencies may therefore complicate embryological studies that utilize this reagent.



Scheme 2 Caging technologies for spatiotemporal control of gene function. (A) Chemical structures of natural and artificial oligonucleotides used to regulate embryonic gene expression. (B) Bhc- or nitrobenzyl-based groups can be used to cage the phosphate backbone of RNAs and siRNAs. (C) Schematic depiction of caged morpholinos and their photoactivation. (D–E) A caged morpholino targeting the zebrafish *ntl* gene. Wildtype zebrafish embryos injected at the one-cell stage with the caged *ntl* morpholino and cultured in the dark develop normally, while similarly injected embryos irradiated with 360-nm light at 4 h post fertilization (hpf) lack notochord and tail structures. (F–G) Spatiotemporal control of *ntl* function illustrates its role in cell fate choice. Cell lineage studies using the photoconvertible fluorescent protein Kaede demonstrate that a population of post-gastrulation (10 hpf) cells are fated to form the notochord by 24 hpf. Similarly irradiated cells in embryos injected with the caged *ntl* morpholino do not contribute to the notochord but rather populate the floorplate domain of the developing spinal chord. 24-hpf embryos are shown, and notochord (nc), floorplate (fp), and somitic (s) tissues are labeled. Scale bars: 200 μm, (D) and (E); 100 μm, (F) and (G). Figure panels (C)–(G) are adapted from ref. 26.

3.2 Caged mRNA and siRNAs

Rather than caging gene expression at the transcriptional level, another strategy would be to modulate the activity of the RNA transcripts themselves. RNA is routinely microinjected into the embryos of worms, fruitflies, frogs and fish to evaluate gene function, and the addition of photocleavable groups to these oligonucleotides could provide some degree of conditionality. Okamoto and colleagues developed a method for randomly modifying the phosphate backbone of RNAs with 6-bromo-4-diazomethyl-7-hydroxycoumarin, thereby caging mRNA encoding the transcription factor Engrailed2a (Eng2a) with an average conjugation ratio of 30 bromohydroxycoumarin (Bhc) groups per kilobase (Scheme 2(A) and (B)).[21] Zebrafish embryos injected with the caged *eng2a* mRNA appeared to develop normally in the dark; however, exposure to 365-nm light caused an eyeless phenotype, deletion of the forebrain, and an anterior shift of the midbrain, suggesting that this factor regulates cell fate choice in the developing central nervous system. Interesting, these patterning defects were more severe than those observed with unmodified *eng2a* mRNA, and *in situ* hybridization analyses indicated that the Bhc modifications significantly increased transcript stability. The Okamoto laboratory was also able to spatially restrict mRNA uncaging by spot illumination, as demonstrated by whole-mount staining with anti-Eng2a antibodies.

More quantitative experiments with caged GFP and β-galactosidase mRNAs, however, revealed some pitfalls to this strategy. While the Bhc modifications could block mRNA translation by $>97\%$, the gene expression induced by light irradiation was less than 15% of that observed with an equivalent amount of unmodified mRNA. Basal and induced levels of gene expression consequently differed by less than five-fold. This limited dynamic range likely is due to the inefficient rate of oligonucleotide uncaging, since dozens of Bhc groups are conjugated to each transcript. Attempts to increase uncaging efficacy by reducing the number of Bhc groups would cause a commensurate increase in basal activity, and it is unclear whether the dynamic range of these caged reagents can be significantly improved.

Similar issues have been observed with caged versions of RNA duplexes such as small interfering RNAs (siRNAs). The Friedman laboratory derivatized a GFP-targeting siRNA with 1-(1-diazoethyl)-4,5-dimethoxy-2-nitrobenzene to varying extents, modifying 5 to 15% of the backbone phosphates (Scheme 2(B)).[22] Co-transfection of these reagents and a GFP expression construct into cultured cells provided a range of basal and light-induced gene silencing, with absolute activities inversely correlated with the number of caging groups per siRNA. In all cases, however, the relative changes in GFP expression upon irradiation were two-fold or less. This limitation can be overcome to some extent by placing a single caging group at the 5′ phosphate, as demonstrated by McMaster and co-workers.[23] Since stability of the RNA-induced silencing complex is promoted by the siRNA 5′ phosphate, siRNAs caged in this manner exhibit lower activities. While cell culture experiments indicated that the 5′ phosphate modification did not completely block siRNA function (5–10% residual activity), photolysis of the single nitrobenzyl-based caging group restored gene silencing to levels approaching that of unmodified siRNAs. RNA duplexes modified in this manner therefore may prove to be useful tools in embryological research, especially since RNA interference is commonly used to interrogate worm, fruitfly and frog development.

3.3 Caged morpholinos

In addition to RNA interference-mediated transcript degradation, RNA functions can be disrupted by synthetic oligonucleotides called morpholinos. These antisense reagents differ from DNA or RNA in two aspects: (1) the ribose sugar is replaced with a morpholino ring, and (2) the negatively charged phosphate backbone is replaced with neutral phosphorodiamidates (see Scheme 2(A)). These structural differences render morpholinos resistant to endogenous nucleases yet allow them to hybridize to complementary RNA sequences with high affinity. Formation of the morpholino/RNA duplex blocks either RNA splicing or translation, depending on whether intron/exon junctions or sequences proximal to the start codon are targeted, and in contrast to RNA interference technologies, transcript degradation is not induced. As first reported by Heasman and Ekker, these properties allow 25-base morpholinos to efficiently block gene expression in frog or zebrafish embryos for several days.[24,25] In frogs this is achieved by microinjecting the morpholino into the zygote or even individual blastomeres at the 32-cell stage. Since cell fate maps for this organism are well established, this latter approach can afford some degree of spatial control. For zebrafish studies, morpholinos are typically injected into the embryonic yolk prior to the 16-cell stage to ensure a uniform distribution of the antisense reagent throughout the embryo. Different morpholino doses can be used to generate phenotypes of varying severity, and two or three morpholinos can be co-injected into embryos to study genetic interactions.

Morpholinos are especially important for zebrafish studies, since targeted gene knockouts by homologous recombination have not yet been achieved in this organism and siRNAs appear to be relatively ineffective in fish embryos. Nevertheless their global and immediate activity limits their utility for studying genes with pleiotropic functions. To address this issue, our laboratory has developed caged morpholinos that can be activated with 360-nm light, taking advantage of the optical transparency of zebrafish embryos.[26] We envisioned that morpholino activity could be gated by tethering the antisense reagent and a complementary inhibitor through a photocleavable linker (Scheme 2(C)). Intramolecular hybridization of the morpholino and inhibitor would abrogate RNA targeting, whereas dissociation of the inhibitor upon linker photolysis would enable morpholino/RNA binding.

As a proof of concept, we targeted the *no tail* (*ntl*) gene, a T-box transcription factor that is required for development of the notochord and posterior mesoderm. Zebrafish embryos injected with the caged *ntl* morpholino and cultured in the dark developed normally, but those irradiated briefly with 360-nm light developed patterning defects associated with *ntl* loss of function, such as loss of notochord and tail structures (Scheme 2(D) and (E)). We subsequently confirmed our ability to spatially and temporally regulate *ntl* function by irradiating

the caged morpholino-injected embryos at different times in either a global or localized manner (Scheme 2(F) and (G)). These studies demonstrated that *ntl* is required for the specification of notochord progenitor cells and their morphogenetic movement during gastrulation, corroborating earlier studies of *ntl* mutants. Our observations further indicated that notochord fate commitment does not occur within the zebrafish organizer during gastrulation as has been suggested previously; rather, *ntl* is required at later time points to maintain the specification of notochord cells and ultimately to promote their differentiation.

3.4 Caged peptide nucleic acid analogs

Other synthetic oligonucleotides have been explored as reverse-genetic tools for embryological research. Although peptide nucleic acids (PNAs) bind to complementary DNA and RNA sequences with high affinity and are resistant to both proteases and nucleases, they do not effectively block gene function in zebrafish embryos, perhaps because of their low solubility. To circumvent this problem, Farber and co-workers designed a negatively charged PNA analog (ncPNA), composed of alternating *trans*-4-hydroxy-L-proline and phosphonate PNA monomers (see Scheme 2(A)).[27] Although these reagents are more challenging to synthesize than morpholinos, 18-base ncPNAs can block target gene expression in zebrafish embryos with efficacies rivalling those of morpholino reagents.

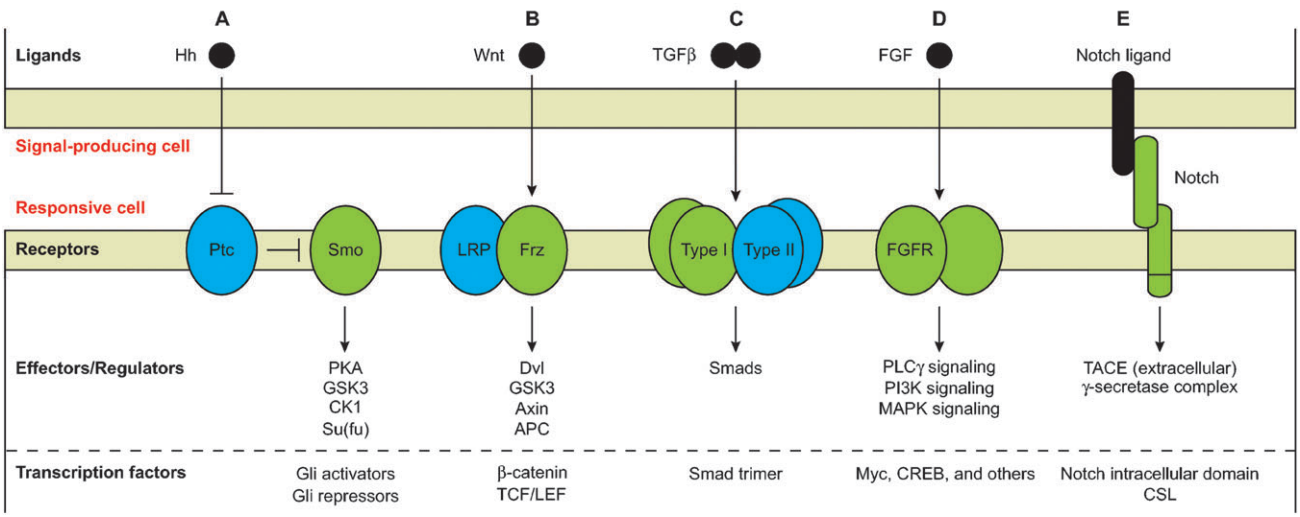
Using a strategy similar to our caged morpholino technology, Dmochowski and Weinberg collaboratively developed photoactivatable versions of ncPNAs by attaching complementary 2′-*O*-methyl RNAs through a nitrobenzyl-based linker.[28] Zebrafish embryos injected with caged ncPNAs targeting either the organizer-inducing homeobox gene *bozozok* (*boz*) or the BMP signaling antagonist *chordin* (*chd*) exhibited developmental phenotypes consistent with loss of *boz* or *chd* expression after global irradiation. Although spatiotemporal aspects of *boz* or *chd* function were not investigated, in principle these reagents could be used to convey conditional control of embryonic gene function with cellular resolution.

Caged morpholinos and ncPNAs are therefore promising tools for dissecting the molecular mechanisms of embryonic patterning. It should be noted, however, that their efficacies will vary according to protein degradation rates for each targeted gene. It is also likely that these caged reagents will be subject to the same biological constraints as their uncaged counterparts, which in the zebrafish limits their use to the first three to five days of development.

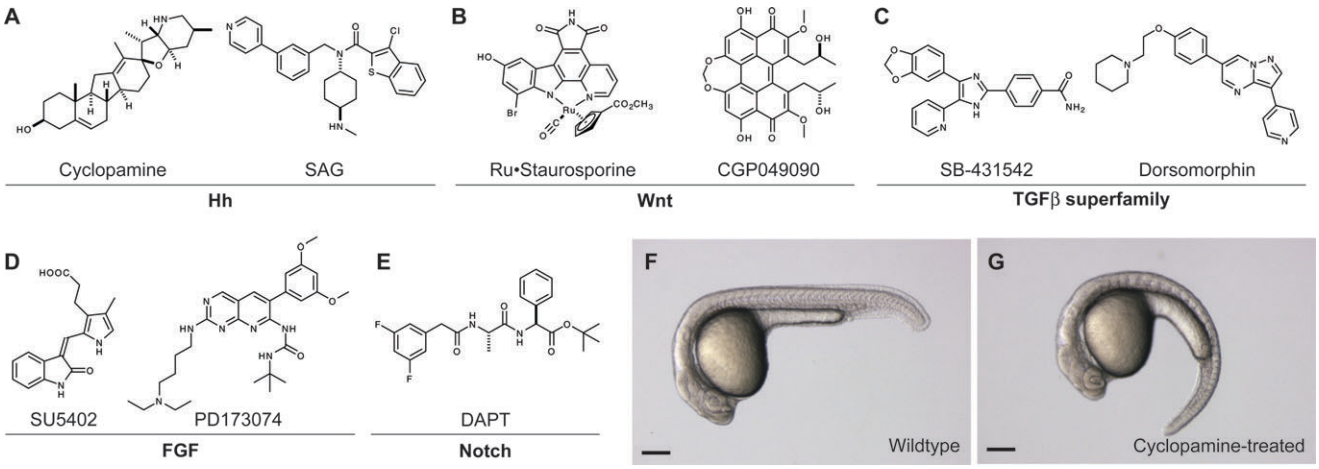
4 Small molecule modulators of developmental signaling pathways

While chemical strategies described above can provide spatiotemporal control of embryonic gene expression, phenotypes arising from these perturbations are contingent upon either the kinetics of RNA and protein synthesis or the degradation rates of these macromolecules. Small molecules that act directly on developmental signaling proteins would bypass this temporal delay, permitting embryological studies on the minute time scale (Schemes 3 and 4). In addition, chemical perturbations can induce phenotypes that are distinct from those obtained from genetic methodologies, since compound binding to a protein target may disrupt only a subset of its cellular functions. These unique attributes have prompted several laboratories to screen large chemical libraries for new teratogens.

Some of these "chemical genetic" screens have utilized biochemical assays to identify ligands for a specific target, whereas others have employed phenotypic screens that involve cultured cells or whole organisms. Each approach has unique challenges. Biologically active compounds discovered through *in vitro* screens must have sufficient bioavailability and specificity to merit their use in embryological studies. Small molecules discovered in cell-based or organismal assays require target identification for meaningful interpretations of their corresponding phenotypes. Efforts toward the latter are typically done on a case-by-case basis, since no universal target-



Scheme 3 Developmental signaling pathways. (A)–(E) Schematic depiction of five major pathways that mediate cell–cell communication during embryogenesis: Hh, Wnt, TGFβ, FGF or Notch signaling. Ligands, receptors, downstream effectors or regulators, and transcription factors for each pathway are shown.



Scheme 4 Small-molecule modulators of developmental signaling pathways. (A)–(E) Representative compounds that perturb the Hh, Wnt, TGFβ, FGF, or Notch pathway. (F–G) Small molecules can phenocopy genetic lesions in whole organisms, as illustrated by the Hh loss-of-function defects observed in zebrafish embryos treated with the Smo antagonist cyclopamine. 24-hpf embryos are shown. Scale bars = 200 μm.

identification strategy has yet been described. Another potential limitation of small molecules is that differences in embryo permeability and target protein structure can restrict their use to certain organisms. Nevertheless, small molecules of sufficient potency and specificity can be valuable reagents in the developmental biologist's toolbox.

4.1 Hh pathway modulators

The impact of small molecules on embryonic patterning was perhaps first appreciated by scientists at the Poisonous Plant Research Unit of the United States Department of Agriculture over fifty years ago. During an investigation of cyclopic lambs born in ranches abutting the Boise, Sawtooth, and Challis National Forests, Binns and co-workers discovered that pregnant sheep fed a flowering plant called *Veratrum californicum* frequently gave birth to lambs with profound craniofacial and limb defects.[29] Subsequent analyses demonstrated that the teratogenic principle in this plant was a steroidal alkaloid, which they aptly named cyclopamine (Scheme 4(A)).

How cyclopamine induces congenital deformities remained unresolved for the next forty years, until genetics provided our first insights into the molecular mechanisms of embryogenesis. In particular, Beachy and colleagues found that mice lacking the *Sonic hedgehog* (*Shh*) gene, one of the three Hh homologs in mammals, exhibited patterning defects reminiscent of the cyclopic lambs. This observation suggested that cyclopamine might target a component of the Hh signaling pathway, and subsequent studies of chick neural tube patterning confirmed that this steroidal alkaloid blocks cellular responses to the Shh ligand.[30] As initially revealed through the Nüsslein–Volhard and Wieschaus mutagenesis screens, Hh signal transduction involves a number of cellular proteins, including the Hh receptor Patched (Ptc), the seven-pass transmembrane protein Smoothened (Smo), and the Cubitus interruptus (Ci)/Gli family of transcription factors (Scheme 3(A)).[31] Ci/Gli proteins can be either activators or repressors, and Smo activity regulates the intracellular ratio of these forms to dictate the Hh pathway state. In the absence of Hh ligands, Ptc represses Smo function and Ci/Gli proteins therefore can be constitutively phosphorylated by protein kinase A (PKA), GSK3β, and casein kinase 1α (CK1α) and then proteolytically converted into their repressor forms. Ci/Gli function is also negatively regulated by the protein Suppressor of Fused (Su(fu)). Hh binding to Ptc, however, leads to Smo activation and shifts the balance of Ci/Gli factors to promote the expression of Hh target genes. The mechanisms of this process are not well understood and may differ between invertebrate and vertebrate species, with the latter having multiple Hh, Ptc and Gli isoforms.

To determine how cyclopamine inhibits Hh signaling, a combination of genetic and chemical approaches was pursued. Epistatic mapping showed that cyclopamine acts at the level of Smo within the pathway,[32] and fluorescent and ^{125}I-labeled-photoaffinity derivatives of cyclopamine were used to demonstrate that the teratogen binds directly to this transmembrane receptor.[33] Since these studies by the Beachy laboratory, other Hh pathway modulators have been discovered in high-throughput cell-based screens of chemical libraries. These include several structurally diverse Smo antagonists and two pathway agonists, a benzothiophene derivative called SAG (also known as Hh–Ag; see Scheme 4(A)) and a purine analog called purmorphamine.[34,35] Studies by the Beachy group and our laboratory have demonstrated that both of these compounds directly target Smo as well.[36,37] Indeed, Smo appears to be particularly sensitive to small-molecule modulation in comparison to other Hh pathway components, and efforts to identify compounds that target downstream effectors are now underway in our research group.

While these compounds can modulate Hh pathway activation in cultured mouse cells, their utility for embryological studies hinges upon their efficacy and pathway specificity in whole organisms. In the case of cyclopamine, the congenital defects observed in lambs half a century ago suggest that this alkaloid is highly selective for the Hh pathway. Accordingly, zebrafish embryos treated continuously with cyclopamine exhibit phenotypes that are essentially identical to those observed in mutants lacking Smo expression, and this small molecule has been widely used to interrogate the role of zebrafish Hh signaling in hematopoiesis, craniofacial

patterning, gut organogenesis, and other developmental processes. By varying the timing of cyclopamine treatment in these studies, developmental biologists have gained a temporal understanding of Hh-dependent patterning that would be impossible to achieve with *smo* mutant embryos alone. Cyclopamine has also been used to explore patterning mechanisms of the salamander, frog, chicken, mouse, and other organisms; however, fruitfly Smo is not inhibited by this compound.[33] Similarly, SAG can partially rescue patterning in *Shh* knockout embryos,[34] but it cannot activate zebrafish Smo (J. K. Chen, unpublished observations). Species-specific activity is therefore one of the limitations of chemical genetic approaches, underscoring the importance of matching screening conditions to the model organism to be used for *in vivo* studies.

4.2 Wnt pathway modulators

The fruitfly screens by Nüsslein-Volhard and Wieschaus also led to the discovery of signaling proteins in the Wnt pathway, which plays a critical role in body axis specification and the subsequent patterning of multiple tissues.[38] As in the Hh pathway, Wnt signaling is initiated by secreted ligands (Wnts) that bind to cell-surface receptors in responsive cells and activate specific gene expression programs (Scheme 3(B)). The membrane receptors in this case are composed of low-density-lipoprotein receptor-related protein (LRP) and Frizzled (Frz) family members, which communicate with a cytoplasmic complex containing Axin, adenomatosis polyposis coli (APC) protein, and GSK3β through the Dishevelled (Dvl) family of multi-module proteins. In the absence of Wnt ligands, this intracellular complex promotes the proteolytic degradation of a transcriptional activator called β-catenin, allowing members of the T-cell factor (TCF) family to repress the expression of Wnt target genes. Activation of the Frz receptors promotes Dvl-dependent inhibition of the Axin/APC/GSK3β complex, leading to β-catenin accumulation and Wnt target gene transcription by TCF/β-catenin complexes.

While cell-based screens for chemical Wnt pathway modulators have been reported, strategies targeting specific Wnt signaling proteins have been more effective thus far (Scheme 4(B)). For example, Meggers and co-workers designed a ruthenium-cyclopentadienyl derivative of staurosporine that is a highly potent and specific inhibitor of GSK3 isoforms (IC50 <1 nM).[39] This unusual organometallic compound can induce β-catenin accumulation and Wnt target expression in cultured cells, and zebrafish embryos treated with this inhibitor exhibit head and eye defects that have been previously associated with GSK3β blockade. The TCF/β-catenin complex can also be effectively targeted by small molecules, as demonstrated by Shivdasani and co-workers.[40] Using a high-throughput *in vitro* assay for TCF4/β-catenin binding, his research group was able to identify six natural products that blocked TCF4/β-catenin heterodimerization from a collection of approximately 7000 compounds, refuting the common notion that protein–protein interactions cannot be disrupted by small molecules. All six compounds were able to block Wnt target gene expression in cell culture with low micromolar potencies, and three were able to inhibit β-catenin-induced axis duplication in frog embryos (compounds PKF118-310, PKF115-584, and CGP049090). Although these Wnt pathway antagonists and agonists have not been extensively evaluated *in vivo*, they illustrate the potential utility of chemical tools for studying Wnt-dependent patterning.

4.3 TGFβ superfamily modulators

The TGFβ superfamily comprises a large class of secreted morphogens and their downstream effectors that regulate cell proliferation and differentiation in diverse embryonic contexts.[41] For example, certain family members establish the body axes during early embryogenesis and then pattern the limbs, spinal cord, and other tissues. The superfamily comprises at least thirty genes in higher organisms and can be categorized into four major subgroups, including the TGFβ, BMP, Activin/Nodal, and growth differentiation factor (GDF) families. In spite of this genetic complexity, the mechanisms by which these ligands effect transcriptional responses are highly conserved. All TGFβ superfamily ligands dimerize to bridge two pairs of serine/threonine kinase receptors, forming an active, hetero-tetrameric signaling complex (Scheme 3(C)). Upon ligand-binding, the type I receptor is phosphorylated by the type II receptor, and cytosolic effector proteins called Smads are recruited to these phosphorylation sites. The translocated Smads are then phosphorylated by the type I receptors, enabling them to interact with other Smad isoforms to form a trimeric factor that regulates target gene transcription. Multiple isoforms of each receptor and effector exist in complex organisms such as vertebrates, providing functional specificity and redundancy during embryogenesis. In addition, the various type I and type II receptors can interact combinatorially to diversify ligand responsiveness. TGFβ signaling mechanisms, however, can be grossly separated at the level of the type I receptor into two distinct cascades, with the first receptor class (Alk4, Alk5 and Alk7 in vertebrates) interacting with one subset of Smads (Smad2 and Smad3) and the second (Alk1, Alk2, Alk3 and Alk6) communicating with another (Smad1, Smad5 and Smad8). Smad4 is utilized by both signaling branches to form trimeric transcriptional regulators.

Due to the complexity and redundancy of TGFβ signaling, identifying small molecule modulators that rival the specificity of genetic perturbations is a significant challenge. Nevertheless, inhibitors that can distinguish between the two signaling branches have been identified (Scheme 4(C)). Laping and colleagues identified a dihydropyrroloimidazole called SB-431542 that selectively inhibits the kinase activity of type I receptors Alk4, Alk5 and Alk7 using an *in vitro* assay of Alk5-mediated phosphorylation of Smad3.[42] The ligand specificity of these receptors makes SB-431542 selective for signaling induced by TGFβ and Activin/Nodal ligands, and frog and zebrafish embryos treated with this inhibitor exhibit phenotypes consistent with a loss of Nodal signaling, such as severe head, midline and somitic defects.[43] A complementary inhibitor called dorsomorphin was recently discovered by Peterson and co-workers in a zebrafish embryo screen for dorsalizing compounds.[44] In contrast to SB-431542, dorsomorphin targets the type I receptors Alk2, Alk3 and Alk6 to selectively block

BMP signaling, which initially induces ventral cell fates during embryogenesis and then promotes bone development. Taking advantage of the temporal control afforded by small molecules, the Peterson laboratory was able to functionally separate these two BMP activities by exposing zebrafish to dorsomorphin after establishment of the dorsal–ventral axis. Embryos treated in this manner have diminished vertebral and craniofacial bone calcification but otherwise appear to be morphologically and functionally normal.

4.4 FGF pathway modulators

Another large group of morphogenetic proteins with diverse embryological functions is the FGF family of secreted factors.[45] While invertebrates such as the fruitfly typically have two or three members of this family, vertebrates can have over twenty FGF ligands, which regulate brain morphogenesis, limb patterning, and other developmental processes. Metazoans can also have multiple FGF receptors (FGFRs), with splice variants providing additional tissue-specific complexity. As in TGFβ signaling, the binding of FGF to its cell-surface receptors induces their aggregation, in this case the formation of homodimers that autophosphorylate tyrosine residues in their cytoplasmic tails (Scheme 3(D)). FGFR phosphorylation then recruits proteins containing Src homology 2 (SH2) or phosphotyrosine binding (PTB) domains to the plasma membrane, which in turn initiate downstream signaling processes such as the mitogen-activated protein kinase (MAPK), phospholipase C-γ (PLCγ), and phosphatidylinositol-3-kinase (PI3K) pathways. As a result of this complex interaction network, FGF signaling can lead to the activation of several transcription factors.

Since FGF signaling intersects with other pathways that can have FGF-independent functions, targeting the growth factors or receptors themselves provides the best chance for *in vivo* specificity. Efforts by the pharmaceutical industry to block FGF signaling for therapeutic purposes have yielded several FGFR inhibitors, of which the indolinone derivative SU5402 has been most widely used in embryological studies (Scheme 4(D)). For example, SU5402-induced phenotypes have demonstrated that FGF signaling is required for tooth and ear formation in zebrafish.[46,47] The relatively weak potency of SU5402 (IC50 ~10–20 μM) limits its *in vivo* efficacy, however, and more potent FGFR inhibitors would facilitate chemical genetic studies of FGF-dependent patterning. One promising lead is the nanomolar inhibitor PD173074, which was recently used to interrogate FGF-mediated retina regeneration in chick embryos.[48]

4.5 Notch pathway modulators

While secreted morphogens such as Hh, Wnt, TGFβ and FGF play critical roles during embryogenesis, intercellular communication can also be achieved with membrane-bound ligands. One prominent example of this signaling mechanism is the Notch pathway, which regulates cell fate decisions in the developing heart, vasculature, pancreas, mammary gland, and other tissues.[49] Notch proteins are expressed as single polypeptides that are cleaved intracellularly to yield heterodimeric, single-pass transmembrane receptors (Scheme 3(E)). Notch ligands are similarly localized to the cell surface through a single transmembrane domain, such that Notch receptor/ligand interactions typically involve cells that are immediately adjacent to each other. Upon ligand binding, the Notch extracellular domain is proteolytically cleaved from the cell surface by TNFα-converting enzyme (TACE) and subsequently endocytosed by the ligand-expressing cell. The remaining portion of the Notch receptor is then cleaved by a multicomponent complex called γ-secretase, which allows the cytoplasmic domain to translocate into the nucleus and form a transcriptionally active complex with the protein CSL (CBF1/Suppressor of Hairless/Lag1).

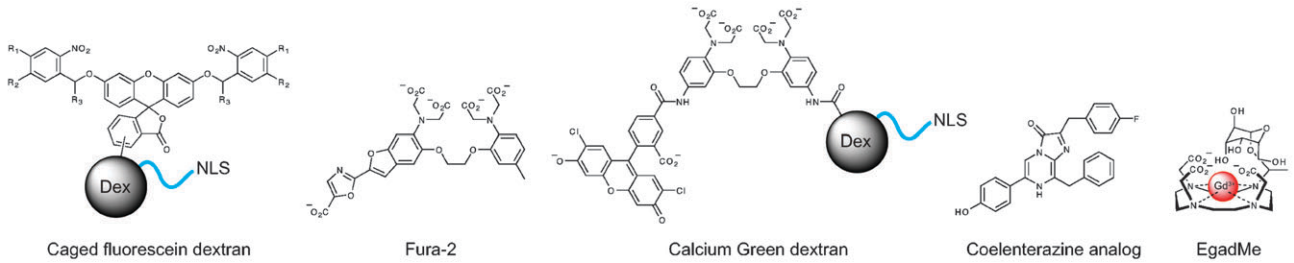
The remarkably simple architecture of the Notch signaling pathway provides few targets for small-molecule modulation. Among the pathway components, the γ-secretase complex is perhaps the most attractive, especially since its parallel role in β-amyloid formation has prompted the pharmaceutical development of γ-secretase blockers as possible Alzheimer's therapies. One γ-secretase inhibitor known as DAPT (*N*-[*N*-(3,5-difluorophenacetyl)-1-alanyl]-(*S*)-phenylglycine *tert*-butyl ester) produces phenotypes in fruitfly and zebrafish embryos that are indistinguishable from those observed in Notch pathway mutants (Scheme 4(E)).[50,51] This pan-species activity is noteworthy and undoubtedly reflects conservation of the DAPT-binding site in γ-secretase between invertebrate and vertebrate organisms. The utility of DAPT for studying embryonic patterning has been further demonstrated by Reh and co-workers, who used this inhibitor to determine the timing of Notch-dependent neuronal differentiation in the developing chick retina.[52]

5 Chemical tools for visualizing embryonic processes

Chemical approaches can also make significant contributions to the latter half of the "perturb and observe" paradigm. Visualizing embryogenesis through chemical reagents can be traced to the pioneering studies of Lindahl during the 1930s, in which Nile Blue sulfate dye was used to stain individual blastomeres of the sea urchin embryo. Advances in imaging technologies have transformed this tradition of direct embryological observation, and chemical probes have been developed in parallel to illuminate the molecular and cellular mechanisms of ontogeny (Scheme 5). Selected examples of reagents that enable the visualization of cell lineages, morphogenetic movements, cell signaling mechanisms, and embryonic gene expression are described below.

5.1 Cell lineage tracers

Although dyes such as Nile Blue sulfate provided the first developmental fate maps, the labeling of individual cells with these reagents becomes increasingly challenging as the embryo grows in complexity. As discussed above, focused light can be used to irradiate selected embryonic cells with spatial precision, and photoactivatable fluorophores are now commonly used in cell tracking studies. Vincent and O'Farrell created one of the first photoactivatable lineage tracers by conjugating caged fluorescein and a nuclear localization sequence (NLS) to dextran (molecular weight ~70 kD).[53] Fruitfly embryos



Scheme 5 Chemical sensors for embryological studies. Representative reagents that have been used to visualize molecular and cellular processes during embryogenesis are shown. R groups in the caged fluorescein structure can be varied to optimize the excitation wavelength or efficiency of photoactivation.

injected with this reagent uniformly distribute the tracer among its cells, which can then be specifically marked by microbeam illumination. The localization peptide restricts visible fluorescence to nuclei, which facilitates the tracking of individual cells and also enables paraformadehyde fixation for subsequent analyses. Using this caged probe, Vincent and O'Farrell were able to demonstrate that anterior–posterior cell fates within each fruitfly body segment are not intrinsic, cell-autonomous decisions but rather dictated by extracellular signals (now known to include fruitfly orthologs of Hh and Wnt). Caged fluorescein reagents subsequently have been used to demonstrate that the first cell cleavage in sea urchin embryos does not specify the axes of bilateral symmetry and to reveal the role of notochord cells in regulating heart specification in zebrafish.[54,55] A complementary caged rhodamine tracer has also been described by Gee and colleagues,[56] enabling cell tracking analyses in embryos co-labeled with green fluorophores (*e.g.* transgenic organisms with tissue-specific GFP expression).

The recent development of fluorescent proteins that are photoactivatable or photoconvertible has provided genetically encoded alternatives to caged fluorescein and rhodamine dextrans. However, the synthetic reagents still have some advantages. One particularly useful property is that the photoactivated fluorophore can be specifically detected in paraformaldehyde-fixed embryos using anti-fluorescein or anti-rhodamine antibodies. This versatility allows cell tracking to be combined with methods for visualizing cellular proteins or transcripts. Thus, the utility of synthetic, photoactivatable tracers in embryological research presents opportunities for chemical biologists, particularly since commercial vendors have discontinued production of caged fluorescein and rhodamine reagents.

5.2 Calcium signaling sensors

Chemical reagents can also provide a window into the cell signaling mechanisms that regulate embryonic patterning. This is perhaps best exemplified by the application of calcium-sensitive dyes pioneered by the Tsien laboratory. Chemical sensors were first used in embryos by Tsien and Poenie, who injected sea urchin eggs with the Ca^{2+}-sensitive fluorophore fura-2 and visualized the calcium wave that traverses the oocyte upon fertilization.[57] Precise measurements of intracellular calcium levels can be obtained by ratiometric fluorescence imaging of the fura-2 probe, providing a quantitative description of calcium signaling in the developing embryo. Since this seminal experiment, other calcium sensors have been used to visualize calcium dynamics during zebrafish development. Calcium Green-conjugated dextran has been used to reveal transient, localized calcium "spikes" during the blastula period, which begins shortly after fertilization and is characterized by multiple cycles of near-synchronous cell division followed by cellular stratification.[58] The seemingly stochastic waves of calcium signaling are limited to the outmost embryonic cell layer and are believed to promote the differentiation of these cells into an epithelial-like barrier that protects the embryo.

Embryonic changes in calcium levels can also be detected with a semi-synthetic version of the Ca^{2+}-dependent photoprotein aequorin, in which the natural coelenterazine cofactor has been replaced with a synthetic analog that achieves greatly increased luminescence intensities. By injecting this light-emitting sensor into zebrafish zygotes, Webb and co-workers were able to observe dynamic changes in calcium levels during gastrulation.[59] Calcium signaling occurs in transient non-propagating domains within the blastoderm margin (from which the cells involute to form the mesoderm and endoderm) and in rhythmic waves that traverse the gastrula along either the blastoderm margin or the forming anterior–posterior body axis. The latter signaling events may provide cells with positional information, coordinating their movement and specification.

5.3 Gadolinium-based probes of gene expression

The chemical imaging techniques described above are best suited for optically transparent embryos that develop *ex utero*, and few technologies are available for visualizing pattern formation in opaque or *in utero* embryos. Among the non-invasive techniques developed for clinical use, magnetic resonance imaging (MRI) stands out as having sufficient resolution for observing discrete cell populations ($\sim$10 μm). MRI can distinguish between embryonic tissues by detecting local variations in water concentration and relaxation times, especially if exogenous agents such as paramagnetic ions are introduced to increase image contrast. This sensitivity to contrast agents can be exploited to enable the visualization of embryonic gene expression, as collaboratively demonstrated by Fraser and Meade.[60] By synthesizing a molecule composed of an eight-coordinate chelator for high affinity binding to gadolinium and a galactopyranose residue positioned to fulfill the final coordination site, they created an MRI contrast agent (EgadMe) that does not effectively reduce the relaxation times

of nearby water until the sugar is removed. EgadMe is therefore relatively MRI-inactive in tissues that do not express β-galactosidase, and MRI-active in tissues that do. Using this technology, the Fraser and Meade laboratories were able to selectively detect β-galactosidase-expressing tissues in opaque frog embryos, including cells that cannot be observed by optical methods. Enzyme-activated MRI contrast agents have not yet been used to interrogate embryonic patterning mechanisms; however, this study illustrates the potential of chemical tools to visualize what might otherwise be unseen.

6 Future challenges

Although developmental biology has been largely advanced by embryological and genetic methods, it is now evolving into a molecular and quantitative science that could benefit from chemical insights and technologies. It is notable that many of the molecular approaches described in this review—such as tetracycline-regulated gene expression, tamoxifen-inducible recombination, and negatively charged peptide nucleic acids—were pioneered by biologists with an appreciation for chemical structure and mechanism. As the demand for more complex reagents and technologies increases, scientists skilled in chemistry and organic synthesis will need to assume a correspondingly greater role in developmental biology research. This will require close interactions between traditional developmental biologists and chemists, as well as the training of interdisciplinary scientists who can seamlessly bridge the two communities.

Through this collaborative effort, we will be better situated to tackle the major challenges that remain in the field. For example, current gene expression technologies compare poorly with the dynamics of embryonic patterning, which can involve dramatic spatiotemporal changes in transcript levels within minutes. General methods for modulating gene function with faster kinetics are needed to interrogate these processes. Next-generation caging groups, especially chemical entities that are compatible with two-photon excitation, will be necessary to fully exploit the potential of light-directed embryological perturbations, and new assays for identifying small-molecule modulators that are sufficiently potent and specific for *in vivo* studies are needed, as well as more effective target identification strategies. Finally, chemical tools for visualizing *in vivo* gene expression and function would revolutionize how we investigate fundamental questions in developmental biology, since existing techniques that can detect endogenous transcripts or signaling proteins expressed at physiological levels typically require embryo fixation. Achieving these collective goals will require molecular tools that transcend Nature's building blocks and macromolecules, ensuring that chemical technologies will continue to play an important role in developmental biology research.

Acknowledgements

We gratefully acknowledge financial support from the National Institutes of Health/National Institute of General Medical Sciences (R01 GM072600), the American Cancer Society (RSG DDC-114660), the March of Dimes Foundation (Basil O'Connor Starter Scholar Research Award), the Sidney Kimmel Foundation for Cancer Research, the Astellas USA Foundation, and the Brain Tumor Society/Rachel Molly Markoff Foundation. IAS also gratefully acknowledges predoctoral training support from the California Institute for Regenerative Medicine.

References

1. C. Nusslein-Volhard and E. Wieschaus, *Nature*, 1980, **287**, 795–801.
2. M. Gossen and H. Bujard, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 5547–5551.
3. M. Gossen, S. Freundlieb, G. Bender, G. Muller, W. Hillen and H. Bujard, *Science*, 1995, **268**, 1766–1769.
4. M. K. Shin, J. M. Levorse, R. S. Ingram and S. M. Tilghman, *Nature*, 1999, **402**, 496–501.
5. P. Ridgway, J. P. Quivy and G. Almouzni, *Exp. Cell Res.*, 2000, **256**, 392–399.
6. C. J. Huang, T. S. Jou, Y. L. Ho, W. H. Lee, Y. T. Jeng, F. J. Hsieh and H. J. Tsai, *Dev. Dyn.*, 2005, **233**, 1294–1303.
7. P. J. Kolm and H. L. Sive, *Dev. Biol.*, 1995, **171**, 267–272.
8. J. B. Duffy, *Genesis*, 2002, **34**, 1–15.
9. M. de Graaf, D. Zivkovic and J. Joore, *Dev., Growth Differ.*, 1998, **40**, 577–582.
10. N. A. Liu, H. Huang, Z. Yang, W. Herzog, M. Hammerschmidt, S. Lin and S. Melmed, *Mol. Endocrinol.*, 2003, **17**, 959–966.
11. L. K. Mathew, S. Sengupta, A. Kawakami, E. A. Andreasen, C. V. Lohr, C. A. Loynes, S. A. Renshaw, R. T. Peterson and R. L. Tanguay, *J. Biol. Chem.*, 2007, **282**, 35202–35210.
12. D. No, T. P. Yao and R. M. Evans, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 3346–3351.
13. H. Esengil, V. Chang, J. K. Mich and J. K. Chen, *Nat. Chem. Biol.*, 2007, **3**, 154–155.
14. M. Padidam, M. Gore, D. L. Lu and O. Smirnova, *Transgenic Res.*, 2003, **12**, 101–109.
15. P. S. Danielian, D. Muccino, D. H. Rowitch, S. K. Michael and A. P. McMahon, *Curr. Biol.*, 1998, **8**, 1323–1326.
16. N. Fossat, G. Chatelain, G. Brun and T. Lamonerie, *EMBO Rep.*, 2006, **7**, 824–830.
17. K. Stankunas, J. H. Bayle, J. E. Gestwicki, Y. M. Lin, T. J. Wandless and G. R. Crabtree, *Mol. Cell*, 2003, **12**, 1615–1624.
18. K. J. Liu, J. R. Arron, K. Stankunas, G. R. Crabtree and M. T. Longaker, *Nature*, 2007, **446**, 79–82.
19. L. A. Banaszynski, L. C. Chen, L. A. Maynard-Smith, A. G. Ooi and T. J. Wandless, *Cell*, 2006, **126**, 995–1004.
20. J. Minden, R. Namba, J. Mergliano and S. Cambridge, *Sci. STKE*, 2000, **2000**(62), PL1.
21. H. Ando, T. Furuta, R. Y. Tsien and H. Okamoto, *Nat. Genet.*, 2001, **28**, 317–325.
22. S. Shah, S. Rangarajan and S. H. Friedman, *Angew. Chem., Int. Ed.*, 2005, **44**, 1328–1332.
23. Q. N. Nguyen, R. V. Chavli, J. T. Marques, P. G. Conrad, Jr., D. Wang, W. He, B. E. Belisle, A. Zhang, L. M. Pastor, F. R. Witney, M. Morris, F. Heitz, G. Divita, B. R. Williams and G. K. McMaster, *Biochim. Biophys. Acta*, 2006, **1758**, 394–403.
24. J. Heasman, M. Kofron and C. Wylie, *Dev. Biol.*, 2000, **222**, 124–134.
25. A. Nasevicius and S. C. Ekker, *Nat. Genet.*, 2000, **26**, 216–220.
26. I. A. Shestopalov, S. Sinha and J. K. Chen, *Nat. Chem. Biol.*, 2007, **3**, 650–651.
27. K. A. Urtishak, M. Choob, X. Tian, N. Sternheim, W. S. Talbot, E. Wickstrom and S. A. Farber, *Dev. Dyn.*, 2003, **228**, 405–413.
28. X. Tang, S. Maegawa, E. S. Weinberg and I. J. Dmochowski, *J. Am. Chem. Soc.*, 2007, **129**, 11000–11001.
29. W. Binns, E. J. Thacker, L. F. James and W. T. Huffman, *J. Am. Vet. Med. Assoc.*, 1959, **134**, 180–183.
30. M. K. Cooper, J. A. Porter, K. E. Young and P. A. Beachy, *Science*, 1998, **280**, 1603–1607.
31. D. Huangfu and K. V. Anderson, *Development*, 2006, **133**, 3–14.

32. J. Taipale, J. K. Chen, M. K. Cooper, B. Wang, R. K. Mann, L. Milenkovic, M. P. Scott and P. A. Beachy, *Nature*, 2000, **406**, 1005–1009.
33. J. K. Chen, J. Taipale, M. K. Cooper and P. A. Beachy, *Genes Dev*., 2002, **16**, 2743–2748.
34. M. Frank-Kamenetsky, X. M. Zhang, S. Bottega, O. Guicherit, H. Wichterle, H. Dudek, D. Bumcrot, F. Y. Wang, S. Jones, J. Shulok, L. L. Rubin and J. A. Porter, *J. Biol*., 2002, **1**, 10.
35. X. Wu, S. Ding, Q. Ding, N. S. Gray and P. G. Schultz, *J. Am. Chem. Soc*., 2002, **124**, 14520–14521.
36. J. K. Chen, J. Taipale, K. E. Young, T. Maiti and P. A. Beachy, *Proc. Natl. Acad. Sci. U. S. A*., 2002, **99**, 14071–14076.
37. S. Sinha and J. K. Chen, *Nat. Chem. Biol*., 2006, **2**, 29–30.
38. R. Nusse, *Cell Res*., 2005, **15**, 28–32.
39. G. E. Atilla-Gokcumen, D. S. Williams, H. Bregman, N. Pagano and E. Meggers, *ChemBioChem*, 2006, **7**, 1443–1450.
40. M. Lepourcelet, Y. N. Chen, D. S. France, H. Wang, P. Crews, F. Petersen, C. Bruseo, A. W. Wood and R. A. Shivdasani, *Cancer Cell*, 2004, **5**, 91–102.
41. B. Schmierer and C. S. Hill, *Nat. Rev. Mol. Cell Biol*., 2007, **8**, 970–982.
42. N. J. Laping, E. Grygielko, A. Mathur, S. Butter, J. Bomberger, C. Tweed, W. Martin, J. Fornwald, R. Lehr, J. Harling, L. Gaster, J. F. Callahan and B. A. Olson, *Mol. Pharmacol*., 2002, **62**, 58–64.
43. D. M. Ho, J. Chan, P. Bayliss and M. Whitman, *Dev. Biol*., 2006, **295**, 730–742.
44. P. B. Yu, C. C. Hong, C. Sachidanandan, J. L. Babitt, D. Y. Deng, S. A. Hoyng, H. Y. Lin, K. D. Bloch and R. T. Peterson, *Nat. Chem. Biol*., 2008, **4**, 33–41.
45. B. Thisse and C. Thisse, *Dev. Biol*., 2005, **287**, 390–402.
46. W. R. Jackman, B. W. Draper and D. W. Stock, *Dev. Biol*., 2004, **274**, 139–157.
47. H. Maroon, J. Walshe, R. Mahmood, P. Kiefer, C. Dickson and I. Mason, *Development*, 2002, **129**, 2099–2108.
48. J. R. Spence, M. Madhavan, J. C. Aycinena and K. Del Rio-Tsonis, *Mol. Vision*, 2007, **13**, 57–65.
49. U. M. Fiuza and A. M. Arias, *J. Endocrinol*., 2007, **194**, 459–474.
50. A. Geling, H. Steiner, M. Willem, L. Bally-Cuif and C. Haass, *EMBO Rep*., 2002, **3**, 688–694.
51. C. A. Micchelli, W. P. Esler, W. T. Kimberly, C. Jack, O. Berezovska, A. Kornilova, B. T. Hyman, N. Perrimon and M. S. Wolfe, *FASEB J*., 2003, **17**, 79–81.
52. B. R. Nelson, B. H. Hartman, S. A. Georgi, M. S. Lan and T. A. Reh, *Dev. Biol*., 2007, **304**, 479–498.
53. J. P. Vincent and P. H. O'Farrell, *Cell*, 1992, **68**, 923–931.
54. R. G. Summers, D. W. Piston, K. M. Harris and J. B. Morrill, *Dev. Biol*., 1996, **175**, 177–183.
55. A. M. Goldstein and M. C. Fishman, *Dev. Biol*., 1998, **201**, 247–252.
56. K. R. Gee, E. S. Weinberg and D. J. Kozlowski, *Bioorg. Med. Chem. Lett*., 2001, **11**, 2181–2183.
57. M. Poenie and R. Tsien, *Prog. Clin. Biol. Res*., 1986, **210**, 53–56.
58. E. Reinhard, H. Yokoe, K. R. Niebling, N. L. Allbritton, M. A. Kuhn and T. Meyer, *Dev. Biol*., 1995, **170**, 50–61.
59. E. Gilland, A. L. Miller, E. Karplus, R. Baker and S. E. Webb, *Proc. Natl. Acad. Sci. U. S. A*., 1999, **96**, 157–161.
60. A. Y. Louie, M. M. Huber, E. T. Ahrens, U. Rothbacher, R. Moats, R. E. Jacobs, S. E. Fraser and T. J. Meade, *Nat. Biotechnol*., 2000, **18**, 321–325.