

Modeling a *Hox* Gene Network *in Silico* Using a Stochastic Simulation Algorithm

Jason Kastner,^{*}†,¹ Jerry Solomon,[†] and Scott Fraser[†]

^{*}Department of Applied and Computational Mathematics, and [†]Beckman Institute, California Institute of Technology, Pasadena, California 91125

The amount of molecular information that has been gathered about *Hox cis*-regulatory mechanisms allows us to take the next important step: integrating the results and constructing a higher-level model for the interaction and regulation of the *Hox* genes. Here, we present the results of our investigation into a *cis*-regulatory network for the early *Hox* genes. Instead of using conventional differential equation approaches for analyzing the system, we have adopted the use of a stochastic simulation algorithm (SSA) to model the network. The model allows us to track in detail the behavior of each component of a biochemical pathway and to produce computerized movies of the time evolution of the system that is a result of the dynamic interplay of these various components. The simulation is able to reproduce key features of the wild-type pattern of gene expression, and *in silico* experiments yield results similar to their corresponding *in vivo* experiments. This analysis shows the utility of using stochastic methods to model biochemical networks. In addition, the model has suggested several intriguing new results that are currently being investigated *in vivo*. © 2002 Elsevier Science (USA)

Key Words: computer; hindbrain; *Hox*; model; network; retinoic acid; simulation; stochastic.

INTRODUCTION

Shortly after the closure of the neural tube, the vertebrate hindbrain develops a series of axial bulges called rhombomeres. The segmentation of the hindbrain into the cell lineage-restricted rhombomeres is a crucial process in the proper specification of the developing structures of the hindbrain (Fraser *et al.*, 1990; Lumsden and Krumlauf, 1996). Helping to confer rhombomere identity are members of the *Hox* gene family, a set of transcription factors that play a pivotal role in regulating patterning and axial morphogenesis and which exhibit rhombomere-restricted patterns of expression (Wilkinson, 1993). Retinoic acid (RA) plays an important part in this process as it is able to directly regulate *Hox* family members, and alterations in the RA response elements in the *cis*-regulatory domain of reporter genes significantly change the expression patterns (Gavalas and Krumlauf, 2000). We chose to investigate part of the *Hox cis*-regulatory mechanism in the rhombomeres, involving the earlier expressed patterned genes in the hindbrain, using a mathematical model based on a stochastic simulation algorithm (SSA).

For our SSA investigation into the *Hox* network, we have chosen to investigate the interaction of *Hoxa1*, *Hoxb1*, *Hoxb2*, *Krox20*, and RA in rhombomeres 4 and 5 (r4 and r5). *Hoxa1* is the first of these genes to be expressed in the hindbrain (Murphy and Hill, 1991), and its expression appears to be directly regulated by a retinoic acid response element (RARE) (Frasch *et al.*, 1995; Langston and Gudas, 1992). *Hoxb1* expression appears to depend on RAREs, a 3' element which helps establish early expression (Marshall *et al.*, 1994), and a 5' repressor element which acts in r3 and r5 (Studer *et al.*, 1994) and which appears to start altering gene expression around 8.0 days post coitus (dpc) in the mouse (R. Krumlauf, personal communication). The early expression of *Hoxb1* is also dependent on *Hoxa1* (Studer *et al.*, 1998) with the cofactor pbx (Green *et al.*, 1998; Phelan *et al.*, 1995), but continued expression in r4 is controlled by a strong autoregulatory loop with the cofactors exd/pbx (Popperl *et al.*, 1995) and prep1 (Berthelsen *et al.*, 1998). *Hoxa1* is expressed to a rostral limit in the developing neural tube to the presumptive r3/r4 boundary at 7.75–8.0 dpc, but the expression then regresses, vanishing from the hindbrain by 8.5 dpc. The expression of *Hoxb1* is very similar, except for the continued autoregulatory maintenance in r4 (Maconochie *et al.*, 1996). *Hoxb1*, pbx, and prep1 all have a hand in upregulating *Hoxb2* in r4 (Ferretti *et al.*, 2000; Ma-

¹ To whom correspondence should be addressed. Fax: (626) 584-9527. E-mail: kastner@gg.caltech.edu.

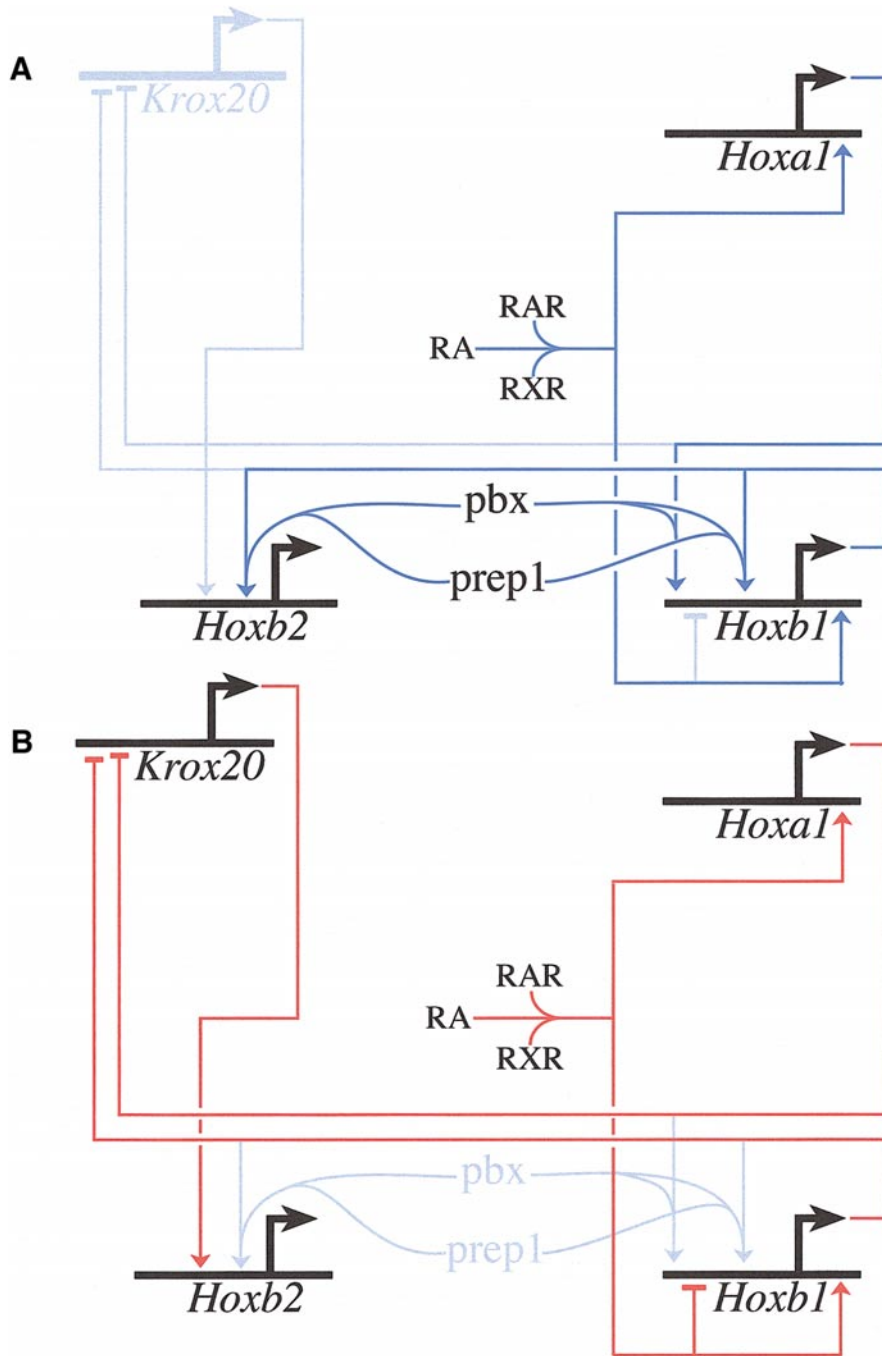


FIG. 1. *Hox* cis-regulatory network in r4 (A) and r5 (B) The network is drawn in a way to emphasize that: (1) each cell contains the entire biochemical network, and (2) certain interactions dominate in a particular rhombomere. Inactive elements are denoted in gray. The activation and repression binding sites are correctly drawn in their relative positions on the chromosome, with the exception of *Krox20* (as it is still unclear how the *Hoxa1* and *Hoxb1* repression mechanism works and where the components are). The horizontal orientation of *Hoxb1* and *Hoxb2* highlights the fact that they appear on the same chromosome, while the vertical orientation of *Hoxa1* and *Hoxb1* highlights the fact that they are paralogs. (A) Starting with RA in the middle of the diagram, the RA binds with RAR and RXR, which can then form a dimer. The dimer can bind as a transcriptional activator to *HoxA1* or *HoxB1* in r4. The *HoxA1* protein, after binding with the pbx/prep1 complex, can then bind as a transcriptional activator to *HoxB1*. *HoxB1*, in conjunction with pbx/prep1 can bind to *HoxB1*, which provides an autoregulatory mechanism. The *HoxB1*/pbx/prep1 complex can also bind as a transcriptional activator to *HoxB2*. (B) The RAR/RXR dimer can bind as a transcriptional activator to *HoxA1* or *HoxB1* in r5, and it can also bind as a transcriptional repressor to *HoxB1*. *HoxA1* and *HoxB1* are hypothesized to be transcriptional repressors of *Krox20*, while *Krox20* is a transcriptional activator of *HoxB2*.

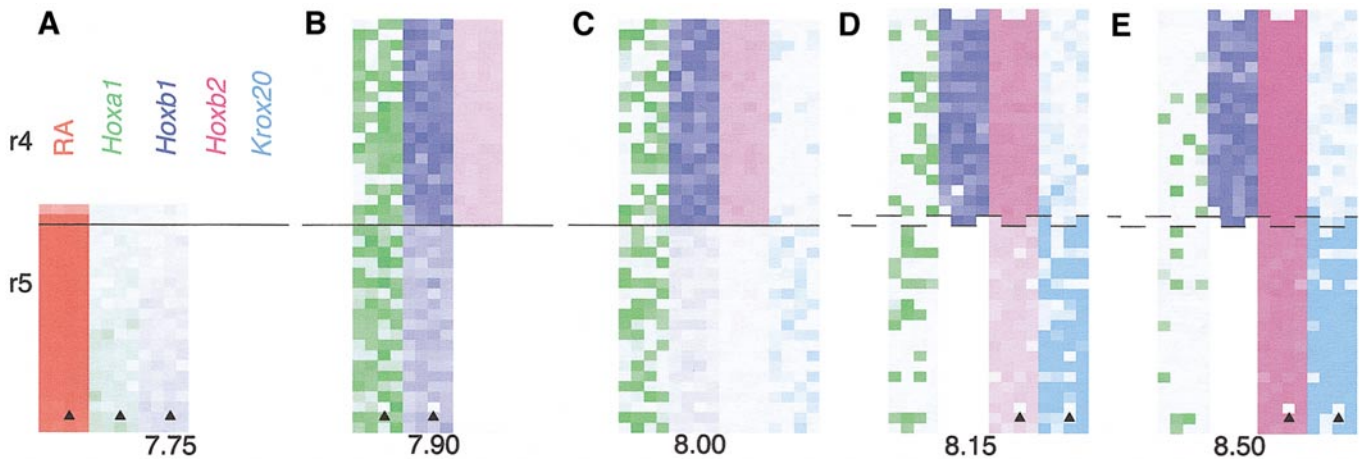


FIG. 2. Simulated wild-type mRNA and RA patterns from 7.75 to 8.5 dpc. (A–E) Selected frames from the computer generated time-lapse movie wt.mov. Four runs of the simulation were required to create this picture, with each run contributing a row of RA, *Hoxa1*, *Hoxb1*, *Hoxb2*, and *Krox20* data for each time point. Notice that sometime between 8 and 8.15 dpc there is a cell division in r5 in the first and fourth data sets. This can be seen most clearly in the *Hoxb2* and *Krox20* data at 8.5 dpc. When a cell divides, its resources are normally distributed between the daughter cells. The data for the marked cell was generated during one of the simulations, and the consequences of this cell misfiring can clearly be seen. (A) At 7.75 dpc, there is an abundance of RA and low levels of both *Hoxa1* and *Hoxb1* expression are evident in the marked cell. (B) The expression of *Hoxa1* and *Hoxb1* fades in this cell by 7.90 dpc, a bit earlier than some of its neighbors. (E) By 8.5 dpc, the cell has failed to initiate its proper expression of *Krox20* and *Hoxb2*. This result suggests that fluctuations are important in the network under investigation.

conochie *et al.*, 1997), while the later r5 expression of *Hoxb2* is regulated by *Krox20* (Nonchev *et al.*, 1996a,b; Sham *et al.*, 1993). In r5, *Krox20* appears to be repressed by *Hoxa1* and *Hoxb1*, and expression of *Krox20* occurs in r5 after they retreat from the hindbrain around 8 dpc. By 8.5 dpc, expression of *Krox20* and *Hoxb2* can be detected in r5 (Barrow *et al.*, 2000; Wilkinson *et al.*, 1989). Thus, we can draw the mouse *cis*-regulatory network as in Fig. 1.

While most of the *cis*-regulatory studies have been carried out in mice, chick has proven to be a useful system for investigation of RA distribution. RA has long been thought to be a diffusible morphogen that is able to pattern the hindbrain (Gavalas and Krumlauf, 2000; Maden, 1999) and recent studies of RALDH-2 and CYP26, enzymes important in RA synthesis and degradation, reveal expression patterns that continue to support this view (Berggren *et al.*, 1999; Swindell *et al.*, 1999). In addition, a RALDH-2 knockout shows effects similar to vitamin A deficiency (Niederreither *et al.*, 1999). More direct tests of sensing this gradient in mouse or chick have been challenging; there has been no conclusive evidence (Gavalas and Krumlauf, 2000). Despite this lack of direct evidence for a gradient, circumstantial evidence for it continues to accumulate. Most recently a study of RAR blocking by an antagonist has shown that the establishment of hindbrain boundaries is dependent on RA concentration (Dupe and Lumsden, 2001). Thus, even if there is not an actual RA gradient, there may be a graded response to retinoids, possibly involving other factors in the system that help modulate the ability of the cell to respond

to RA. Taken together, the evidence is suggestive that a differential of some sort, perhaps through RA concentration, or through the temporally modulated ability to respond to RA, helps establish the *Hox* gene patterns.

The molecular studies of the hindbrain have offered sufficient details to assemble a model for the interactions important in regional control of gene expression. Differential equations have been the preferred method to construct and analyze biochemical network models. The reasons for these are many, but by far, the most important one is that the approaches based on differential equations produce results that are in extremely good agreement with the data (c.f. Hynne *et al.*, 2001; Poolman *et al.*, 2001). In addition, differential equations come with a wide range of analysis tools that allow for a detailed investigation of the model properties.

Despite their prevalence in modeling pure chemical processes, stochastic approaches have not been widely used for modeling biological problems. Instead, some differential equation approaches attempt to capture stochastic effects by adding a “noise” term to their otherwise deterministic treatment (c.f. Meinhardt and de Boer, 2001). Stochastic simulations have proved enlightening and useful in situations where the small volumes and moderate concentrations of reacting species makes the fluctuations an important part of the system (McAdams and Arkin, 1998). Compared with differential equations, stochastic models in biology are still in a relative infancy and only now are generalized tools for constructing and analyzing stochastic

simulations starting to appear. More attention has been focused lately on stochastic effects in biology, and evidence shows that stochastic effects play major roles in gene expression (Greenwald, 1998; Ko, 1992; Zlokarnik *et al.*, 1998). On the simulation side, stochastic methods have already been used to model phage λ -infected *Escherichia coli* cells (Arkin *et al.*, 1998), Notch/Delta lateral inhibition in *Drosophila* (J.K. and J.S., unpublished results), and calcium wave propagation in rat hepatocytes (Gracheva *et al.*, 2001).

The goal of this report is to explore the utility of SSA approaches in modeling of gene networks. The direct coupling of the SSA implementation of a network and individual molecular events would seem to lend itself to both the analysis and logical organization of the ever-growing data on the control of *Hox* genes in the developing hind-brain. The analysis presented here shows that the approach captures the timing, patterning, and variation in *Hox* gene expression without the need for an artificially injected noise. The tests against some of the available experimental perturbations suggest that the SSA will have predictive value and allow researchers in the laboratory to identify and focus attention on the most fruitful experiments.

METHODS

Implementation

An extensive review on the use of stochastic processes in chemical kinetics can be found elsewhere (McQuarrie, 1967), but it is worthwhile to summarize some of the important results. Papers appearing as early as 1940 (Delbrück, 1940; Kramers, 1940) began to examine stochastic processes in chemical reactions, but it wasn't until the early 1950s that it became clear that, in small systems, the Law of Mass Action breaks down and even small fluctuations in the number of molecules may be a significant factor in the behavior of the system (Singer, 1953). Soon after, it became evident that some processes in biological cells fell into this category and that a proper mathematical formulation of the chemical reactions in a cell must be based on stochastics (Bartholomay, 1958).

The SSA was formulated to describe the time evolution of a chemical system (Gillespie, 1976, 1977b). It considers the time evolution as a kind of random walk process governed by the master equation, a differential-difference equation that tracks all of the molecules of interest in a system. Reactions are treated explicitly as stochastic processes and any given reaction is a discrete event. Because this method is the basis for our simulation, we give a brief overview here.

The method is based on the quantity $P(\tau, \mu)$, which is the probability that a particular reaction R_μ occurs in the next time interval τ . The chance that a particular combination of molecules will react is defined by c_μ , a discrete version of the classical continuous rate coefficient k_μ . In order to obtain the proper expression for the underlying probability of each reaction, we must also account for a combinatorial factor, denoted h_μ , which arises due to the number of ways reactant molecules can combine. The joint probability $P(\tau, \mu)$ can be factored into two disjoint probabilities as follows:

$$P(\tau, \mu) = P(\tau) \cdot P(\mu|\tau) \quad [1]$$

where $P(\tau)$ is the probability that a reaction occurs in the next time interval τ , and $P(\mu|\tau)$ is the conditional probability that the specific reaction R_μ occurs given an elapsed time interval τ has passed. These individual probabilities are found to be

$$P(\tau) = a_0 \cdot \exp(-a_0\tau) \quad [2]$$

and

$$P(\mu|\tau) = \frac{a_\mu}{a_0} \quad [3]$$

with

$$a_\mu \equiv h_\mu \cdot c_\mu, \quad a_0 \equiv \sum_{\mu}^M h_\mu \cdot c_\mu \quad [4]$$

where M is the total number of reactions that can occur. a_μ is then the probability that a certain reaction will occur given that the system is in a particular state at time τ .

Each reaction that occurs changes the quantity of at least one reactant. When this happens, the combinatorial factors h_μ change and it is necessary to recalculate the a_μ values. In general, only a small number of the a_μ will actually have to be updated and an efficient implementation needs to take advantage of this fact. Other improvements to the method have also been introduced, including an extension for diffusion which we have incorporated into our simulation (Stundzia and Lumsden, 1996). Two important points should be noted about the SSA: the solution of a system of coupled chemical reactions by this method is entirely equivalent to the solution of the corresponding stochastic master equations (Gillespie, 1976, 1977b; McQuarrie, 1967), and in the limit of large numbers of reactant molecules, the results of this method are entirely equivalent to the solution of the traditional kinetic differential equations derived from the Law of Mass Action (Gillespie, 1977a). An example of the SSA applied to the Michaelis-Menten basic enzyme reaction can be found at <http://www.bioimaging.caltech.edu/hox.html>. The pages also contain a list of the chemical reactions and the parameters used in the model.

Because the model is built on, and driven by, the underlying biochemistry of the system, the reactions can be translated directly into the discrete events of the simulations. In this investigation, some of the steps of the system were deliberately omitted. For example, instead of creating explicit reactions for the transcription of nuclear RNA, the splicing into mRNA, and the exporting of the mRNA to the cytoplasm, the simulation instead creates mRNA as a primary transcript. This is not unreasonable as long as the rate parameters c_μ are adjusted accordingly. In an effort to bring more realism to the model, we plan to add the above steps at a later date.

Using Fig. 1 as our network of interest, we have implemented an SSA, using the C programming language, for the *Hox* network. The model currently contains 59 chemical events that can occur in each cell. They can be classified into 5 main categories: binding (including activation, repression, dimerization, and *Hox*/pbx/prep1 complex formation), unbinding, transcription, translation, and decay (of mRNA, dimers, complexes, proteins, and receptors). The two remaining events that do not fall into these categories are diffusion and division.

The function for binding RA to RAR is a second order reaction and takes the form $a_\mu = c_\mu s_1 s_2$, where s_1 is the number of RA molecules, and s_2 is the number of RAR molecules. The functions

TABLE 1
 K_d Values for Various Chemical Events

Event	K_d (nM)	Reference
RA binding to RAR	0.5	(Allegretto <i>et al.</i> , 1993)
RA binding to RXR	2	(Allegretto <i>et al.</i> , 1993)
RAR/RXR dimerization	17	(Depoix <i>et al.</i> , 2001)
Dimer binding to <i>Hoxa1</i>	3.8	(Mader <i>et al.</i> , 1993)
Dimer binding to <i>Hoxb1</i>	5.3	(Mader <i>et al.</i> , 1993)
Hox/pbx/prep binding to DNA	2	(Pellerin <i>et al.</i> , 1994)

for RA binding to RXR, and the dimerization of the bound RAR and RXR are similar, as is formation of the Hox/pbx/prep complexes. The activation and repression functions are implemented by using a Hill function (Hill, 1910), a typical way to represent cooperative

binding, and takes the general form $a_\mu = c_\mu \frac{f^h}{\kappa_\mu + f^h} f \cdot g$, where f is the number of molecules of a particular transcription factor, κ_μ is a modulation factor, and g is the number of molecules of a gene available. If a gene is currently unbound, the value of g is 1, while if it is bound by a transcriptional activation or repression factor, the value of g is 0. The exponent h is called the Hill coefficient and it affects the speed of the response. The Hill function is an empirically derived expression, used in differential equation models, that yields the observed kinetics in these situations. Thus, in the stochastic reaction approach, we treat the complete Hill function expression as simply another rate coefficient for the purposes of converting it to the appropriate probability of occurrence of the corresponding reaction. Others have used a similar method in their stochastic description of gene transcription (Arkin *et al.*, 1998).

As mentioned previously, there is compelling evidence that a variation of some sort is an important component in patterning the hindbrain, and we have chosen to model this as an RA gradient. We considered many possible functions for modeling the RA distribution, with one of our key criteria being a mechanism where RA would diminish over time. We finally chose to use a Rayleigh function to model the diffusion source term for RA from the posterior of the embryo. The general form is given by $RA(\tau) = RA_0 \tau e^{-\alpha \tau^2}$, where α controls the decay time of the source. We used the Rayleigh function because we believe the behavior captures two important features of a real biological source: a fairly rapid and smooth ramp-up, and a long diminishing tail. The remaining functions are modeled as first order reactions with $a_\mu = c_\mu s_1$, where s_1 can be the number of mRNA available to be turned into proteins or the number of proteins that are available to be decayed, or other similar reactions. In implementing the repression of *Hoxb1*, the simulation started this mechanism around 8.0 dpc because of the current understanding that the repression starts later than the activation (R. Krumlauf, personal communication). The *Hoxa1* and *Hoxb1* repression for *Krox20* is also started at around 8.0 dpc to ensure the establishment of *Hoxa1* and *Hoxb1* before the *Krox20* expression.

Parameters

Using appropriate values for the model parameters is an important component in modeling the system behavior. Fortunately, several key parameters are known, but many of the important

parameters for the model have not been assayed directly in experiments on the developing hindbrain. Estimates of many of their values can be made from data obtained in other systems, and were used in selecting parameters here (Table 1). For example, the K_d value for RAR/RXR dimerization has been determined in HeLa cells. We do not expect that the model results will be significantly different when we incorporate newly measured parameters and concentrations in place of the values we were forced to estimate. A sensitivity analysis, in which the model is rerun with systematically varied parameters, shows that the results remain qualitatively unchanged. As examples, altering the number of RAR and RXRs in each cell by 20% does not change the qualitative behavior of the wild-type system, nor does changing the binding affinity of RA to RAR by 20% (results not shown). This is encouraging, as we would not expect the overall biological system to be sensitive to moderate changes in the concentrations or rates.

The half-lives for mRNA can range from minutes to hours, and values for the *Hox* mRNA have not been measured. We have chosen a typical half-life of around 15–20 min, numbers that are in line with other values in early embryogenesis (Davidson, 1986). The half-lives of the proteins in our network have not been measured and we chose values between 15 and 30 min. These numbers are again in an acceptable range for transcription factors (A. Varshavsky, personal communication). Similar values were used for the turnover of the receptors and complexes. With respect to the number of RARs and RXRs, we have chosen a value of around one thousand of each type as being reasonable (Lauffenburger and Linderman, 1993). No distinction is made between the α , β , and γ forms. In this version of the simulation, the cofactors pbx and prep1 are treated as a single molecule, which the *Hox* proteins can bind with on the DNA. In future versions, we will include a more complete characterization of the cofactors.

RESULTS

The early *Hox* genes first appear around 7.75 dpc (head-fold) and the patterns of *Hoxa1*, *Hoxb1*, *Hoxb2*, and *Krox20* stabilize by 8.5 dpc (~10 somites). Using the network shown in Fig. 1, we sought to capture this wild-type expression and have run the model for a simulated time of 18 h. The model is one-dimensional along the rostralcaudal axis of the embryo. Running the simulation with different random number seeds shows that the model is not overly sensitive to the initial seed value. In our figures, a number of these independent runs are assembled side-by-side to construct a two-dimensional sheet of cells that resemble the tissue (with a mediolateral dimension). This offers insights into the expected two-dimensional pattern of gene expression in the hindbrain and displays the variability in the results.

To display the results of the simulations, we have scaled the raw data (the number of molecules of each type in each cell) to numbers between 0 and 1 by dividing by the maximum value in that data set. This allows us to create a color shading so that differences in levels of molecules are clear. We have displayed the results in an easy to understand format: a virtual dynamic *in situ*. Because the maximum value used to scale the data is typically on the order of a couple hundred molecules, the color variations that are

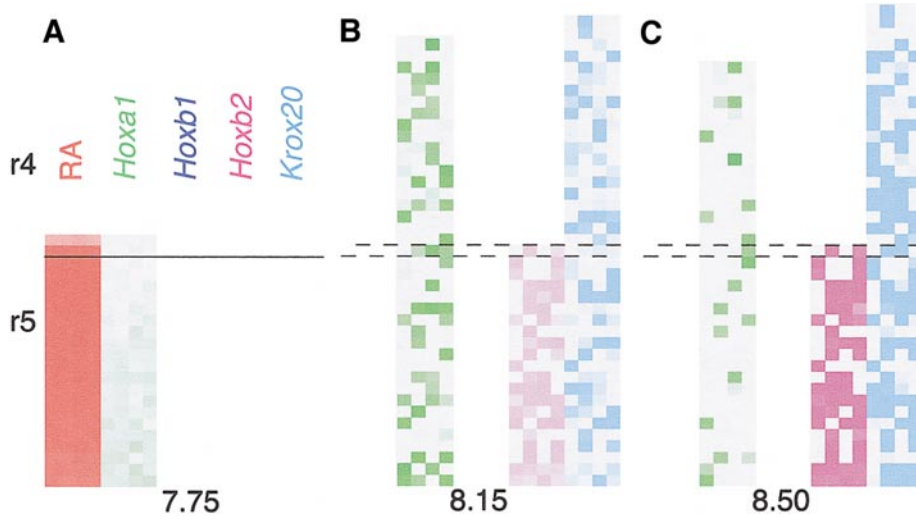


FIG. 3. Simulated *Hoxb1* mutant mRNA expression patterns. (A–C) Selected frames from the computer-generated time-lapse movie *Hoxb1mutant.mov*. This data set shows cell division having occurred in both r4 and r5. Besides affecting the *Hoxb2* expression in r4, the *Hoxb1* mutant also has an effect on *Hoxb2* and *Krox20* in r5. (B) The levels of *Krox20* are lower at 8.15 dpc than in the wild-type (Fig. 2D). (C) By 8.5 dpc, the levels of *Krox20* and *Hoxb2* are noticeably lower than the wild type (Fig. 2E).

seen in the figures and the movie may in fact be too small to distinguish in a laboratory setting using conventional *in situ* staining.

Wild-Type

In Fig. 2, we present the dynamics of the model concerning the emergence of *Hoxa1*, *Hoxb1*, *Hoxb2*, and *Krox20*, over time from approximately 7.75 to 8.5 dpc. The figure presents single frames from the movie *wt.mov*. Along with all the other movies referenced in this paper, *wt.mov* can be found on the Web site <http://www.bioimaging.caltech.edu/hox.html>. The movie offers a dynamic view of the mRNA and RA in the developing hindbrain. Each rhombomere starts out with 20 cells, and the presumptive boundary is clearly marked. Because the simulation encompasses 18 h of developmental time, we have included a rudimentary mechanism for cell division and this is why the presumptive boundary sometimes shifts. We see the low levels of *Hoxa1*, *Hoxb1*, and *Hoxb2* mRNA in r4 and r5 soon after the simulation starts when the RA sweeps across the cells (Fig. 2A). After the mRNA is translated into protein and subsequently forms a complex with pbx and prep1, it can subsequently bind to the DNA. We then see the effects of the *Hoxa1* binding site on *Hoxb1* and the *Hoxb1* autoregulatory loop, namely the higher levels of *Hoxb1* in r4 (Fig. 2B). By 8 dpc, the RA has long since vanished from the hindbrain, and consequently, the RAR/RXR dimers are no longer being created. This is the main reason that *Hoxa1* starts to vanish from the hindbrain. The lack of available dimers also contributes to *Hoxb1* vanishing from r5, as does the late repression mechanism (Fig. 2C). Now that *Hoxa1*

and *Hoxb1* no longer repress *Krox20* in r5, its expression rises and subsequently brings up *Hoxb2* in r5. At about this time, *Hoxb2* has appeared in r4 due to the up-regulation by *Hoxb1* (Fig. 2D). The ending expression pattern of the simulation at 8.5 dpc (Fig. 2E) is very similar to reported patterns (Lumsden and Krumlauf, 1996).

We know from laboratory data that cells sometimes “misfire,” and using this simulation, it is possible to see the consequences of such misfirings. In Figs. 2A, 2B, and 2E, the cell marked with an arrow deviates from its normal fate and ends up not expressing any genes. At the same time, there are other cells that appear to misfire early, exemplified by low levels of expression, but later recover. Both of these events are known to happen in biological systems, and we are encouraged to see this behavior in our model, as these events are not captured with conventional modeling methods. This result suggests that fluctuations are a factor in the network under investigation.

In Silico Experiments

The versatility of the computer simulation also allows for the possibility of performing *in silico* experiments. We report on two experiments here and show that the results are similar to their corresponding *in vivo* experiments. In the investigation of the cross-regulation of *Hoxb2* by *Hoxb1* in r4 (Maconochie *et al.*, 1997), the authors showed that the up-regulation of *Hoxb2* in r4 is lost in *Hoxb1* mutants. Duplicating this experiment *in silico* requires a minimum number of changes to the model, and the input parameters used were the same as in the wild type. In stills taken from the movie *Hoxb1mutant.mov*, we see that, as in the wild-

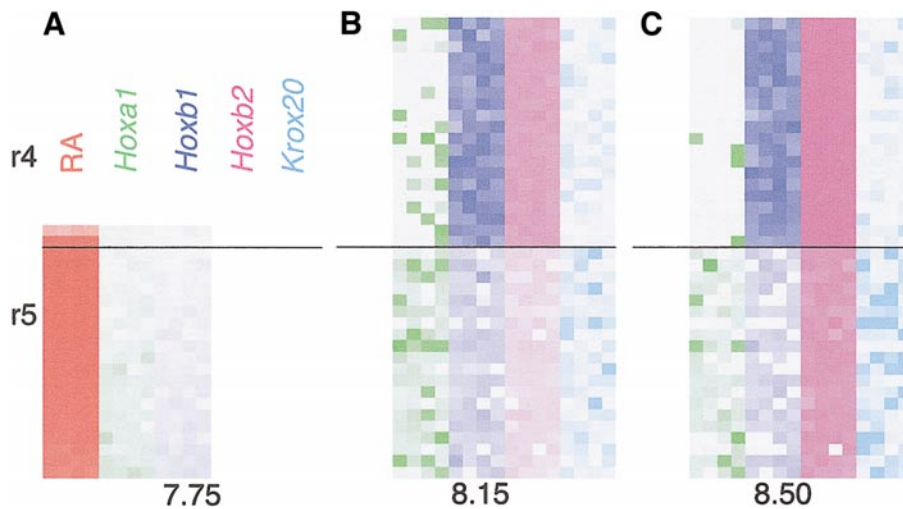


FIG. 4. Simulated expression patterns after inactivation of the 5' *Hoxb1* RARE. (A–C) Selected frames from the computer generated time-lapse movie RAREmutant.mov. By turning off the 5' RARE, we see a change in the levels of *Hoxa1* expression in r5. This occurs because the 3' and 5' RAREs are in effect fighting for the RAR/RXR dimers. This intriguing result needs to be more fully investigated. As in the wild type, it is easy to see downstream effects from cells that have misfired, most notably the patches where *Hoxa1* or *Hoxb1* are continuing to repress *Krox20*. (A) The behavior of the system mimics the wild-type at 7.75 dpc because the 5' RARE does not kick in until 8 dpc. (B) By 8.15 dpc, the expression of *Hoxb1* is still noticeable in r5, but the levels are low enough to allow *Krox20* expression to take hold. (C) The levels of *Krox20* in r5 are higher than in the wild-type (Fig. 2E). The effects of the *Hoxb1* RAREs not having to compete for the dimers are clear by 8.5 dpc as evidenced by the higher levels of *Hoxa1* as compared with the wild type (Fig. 2E).

type, the RA comes through the hindbrain at 7.75 dpc and induces the expression of *Hoxa1*, and because the *Hoxb1* gene is “turned off,” we do not see any *Hoxb1* expression (Fig. 3A). Later, we see that, as reported in the literature, *Hoxb2* is absent from r4. We also see that *Krox20* fails to be well repressed in r4 (Fig. 3B). By 8.5 dpc, *Hoxb1* expression is still absent and high levels of *Krox20* are firmly established in r4 (Fig. 3C). This last result has yet to be thoroughly investigated.

The effects of a selected deletion in the *Hoxb1* 5' RARE showed that the RARE plays a role in the r4 restricted expression of *Hoxb1* (Studer *et al.*, 1994). In this work, the authors showed that, if the construct lacked the 5' RARE, the reporter expression spread to r3 and r5. Further study suggests that the r3/r5 repressor region that contains the RARE is activated later than the 3' enhancer element (R. Krumlauf, personal communication). Duplicating this experiment using the model is again a simple matter, and is accomplished by not turning on the repressor. The stills from the movie RAREmutant.mov show that the expression pattern looks normal at 7.75 dpc (Fig. 4A). However, at 8.0 dpc, we do not turn on the repression mechanism, and by 8.15 dpc, the expression of *Hoxb1* in r5 is still strong (Fig. 4B). By 8.5 dpc, the *Hoxb1* expression has faded in r4 somewhat due to the lack of available RAR/RXR dimers, but is still noticeable (Fig. 4C). In addition, we once again see a change in the pattern of *Krox20*, but this time we see lower expression levels in r5 (Fig. 4C). This is due to the continued repression effects of *Hoxa1* and *Hoxb1*.

DISCUSSION

The stochastic simulation algorithm model captures the timing of several *Hox* gene expression patterns in wild-type animals, and *in silico* simulations performed as a check of key interactions produce results similar to *in vivo* experiments. In addition, the *in silico* experiments yield intriguing predictions that have yet to be thoroughly examined biologically. For example, the mutation experiments in which either *Hoxb1* or the 5' RARE are mutated predict that *Krox20* expression is altered, and in Fig. 3C, it appears that *Hoxb2* expression is affected when *Hoxb1* is mutated. The formal nature of the model calls our attention to these simple test experiments, and checking our prediction will lead to valuable insight into the regulatory network and the veracity of the model. If the model predictions are correct, we will have a tool that will allow us to delve deeper into the components and ask more complicated questions about the nature of the interactions. On the other hand, if the model predictions turn out to be incorrect, the experimental data will allow us to refine the model and incorporate the new results. The revision will then offer different predicted relationships that will stimulate further experiments. Either way, this investigation will ultimately lead to a better predictive tool for the next round of experiments. Indeed, this is one of the great strengths of our simulation: as the components of the model are validated, it can be used to perform numerous *in silico* experiments to identify the *in vivo* experiments that will be the most enlightening.

Our model simulations suggest that a transitory early release of RA may be sufficient to initiate the *Hox* genes. During our investigation of functions for modeling the RA source, it became clear that initiation of the network only required the RA source to stay on for as few as 3 min. All that was needed was enough RA to bind the receptors in r4 and r5. Once this occurred, the network was adequately established and we had proper r5 expression of *Krox20*. Our refinement of the RA gradient hypothesis fits well with recent work on blocking RAR with a chemical antagonist in which the authors made a careful study of concentration and time dependent effects of the blocking agent by using morphology and gene expression as assays. Chick embryos treated with the agent at HH stage 6 (Hamburger and Hamilton, 1951) do not express *Krox20* in r5, but treatment at HH stage 7 permits r5 expression (Dupe and Lumsden, 2001). Thus, the *Krox20* insensitivity to a later change in RA fits well with our model predictions.

In addition to serving as an organizational tool for presenting newly established interactions, the model can also be used to investigate hypothesized molecular interactions. Using it for this purpose will allow researchers to explore the consequences on the network as molecular connections are added or removed. The simulation itself is designed in a way to make modifications easily, and adding new pieces is a modular process. This will inevitably need to occur as new data are presented which require us to update the *cis*-regulatory network (Fig. 1) accordingly. Because the speed of the SSA is linear with respect to the number of reactions, adding new reaction channels will not greatly increase the runtime of the simulation. The speed of the SSA depends more on the number of molecules, and as this population increases, the runtime will scale quadratically. Recent improvements to the algorithm, including a method that does not require the probabilities to be updated after every reaction, are helping to keep the runtime in check (Gibson and Bruck, 2000; Gillespie, 2001). As currently implemented, a typical run of this simulation (without the aforementioned speedups) consists of over 23 million events, and takes less than 6 min on a computer with a 2-GHz Pentium 4 processor.

We believe that the model also can play an important role in explaining differences between species; for example, *Hoxb2* expression in r3 and r5 is much lower in chick than in mouse (Vesque *et al.*, 1996). The differences may be due to regulatory sequences that have yet to be fully characterized, and which can be easily updated in the model once they are known.

We plan on extending this model in ways that are not only spatial and temporal, but which also incorporate more of the known biochemistry of the system. For example, we plan to extend the anterior-to-posterior region of the model to include r3 to investigate the early expression of *Krox20* and we have recently learned that an early low level of *Hoxb2* expression in r5 appears to be due to the *Hoxb1* 3' RARE (R. Krumlauf, personal communication). On the temporal front, we plan on adding the proper mechanisms

to capture later events such as the progressive down-regulation of *Hoxb2* in r3 by 10.5 dpc (Maconochie *et al.*, 1997). Biochemical improvements include implementation of the mRNA modification and transport steps, and a better characterization of the cofactors. All of these improvements will allow for a better understanding of the interaction and timing of the events.

Even in its relatively early state, our model shows the power of using a stochastic simulation algorithm approach for modeling *cis*-regulatory networks, especially in situations where the fluctuations in the system appear to be a factor. We expect that our continued efforts in refining and using the model will result in a greater understanding of how computer simulations can be used to produce new biological insights.

ACKNOWLEDGMENTS

We thank Robb Krumlauf, Paul Kulesa, Rusty Lansford, and Helen McBride for helpful criticisms and suggestions. J.K. is supported by the California Institute of Technology's Initiative in Computational Molecular Biology, a Burroughs Wellcome funded program for science at the interface.

REFERENCES

- Allegretto, E. A., McClurg, M. R., Lazarchik, S. B., Clemm, D. L., Kerner, S. A., Elgort, M. G., Boehm, M. F., White, S. K., Pike, J. W., and Heyman, R. A. (1993). Transactivation properties of retinoic acid and retinoid X receptors in mammalian cells and yeast. Correlation with hormone binding and effects of metabolism. *J. Biol. Chem.* **268**, 26625–26633.
- Arkin, A., Ross, J., and McAdams, H. H. (1998). Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* **149**, 1633–1648.
- Barrow, J. R., Stadler, H. S., and Capecchi, M. R. (2000). Roles of *Hoxa1* and *Hoxa2* in patterning the early hindbrain of the mouse. *Development* **127**, 933–944.
- Bartholomay, A. (1958). Stochastic models for chemical reactions. I. Theory of the unimolecular reaction process. *Bull. Math Biophys.* **20**, 175–190.
- Berggren, K., McCaffery, P., Drager, U., and Forehand, C. J. (1999). Differential distribution of retinoic acid synthesis in the chicken embryo as determined by immunolocalization of the retinoic acid synthetic enzyme, RALDH-2. *Dev. Biol.* **210**, 288–304.
- Berthelsen, J., Zappavigna, V., Ferretti, E., Mavilio, F., and Blasi, F. (1998). The novel homeoprotein Prep1 modulates Pbx-Hox protein cooperativity. *EMBO J.* **17**, 1434–1445.
- Davidson, E. H. (1986). "Gene Activity in Early Development." Academic Press, Orlando.
- Delbrück, M. (1940). Statistical fluctuations in autocatalytic reactions. *J. Chem. Phys.* **8**, 120–124.
- Depoix, C., Delmotte, M. H., Formstecher, P., and Lefebvre, P. (2001). Control of retinoic acid receptor heterodimerization by ligand-induced structural transitions. A novel mechanism of action for retinoid antagonists. *J. Biol. Chem.* **276**, 9452–9459.
- Dupe, V., and Lumsden, A. (2001). Hindbrain patterning involves graded responses to retinoic acid signalling. *Development* **128**, 2199–2208.

- Ferretti, E., Marshall, H., Popperl, H., Maconochie, M., Krumlauf, R., and Blasi, F. (2000). Segmental expression of Hoxb2 in r4 requires two separate sites that integrate cooperative interactions between Prep1, Pbx and Hox proteins. *Development* **127**, 155–166.
- Frasch, M., Chen, X., and Lufkin, T. (1995). Evolutionary-conserved enhancers direct region-specific expression of the murine Hoxa-1 and Hoxa-2 loci in both mice and *Drosophila*. *Development* **121**, 957–974.
- Fraser, S., Keynes, R., and Lumsden, A. (1990). Segmentation in the chick embryo hindbrain is defined by cell lineage restrictions. *Nature* **344**, 431–435.
- Gavalas, A., and Krumlauf, R. (2000). Retinoid signalling and hindbrain patterning. *Curr. Opin. Genet. Dev.* **10**, 380–386.
- Gibson, M. A., and Bruck, J. (2000). Efficient exact stochastic simulation of chemical systems with many species, and many channels. *J. Phys. Chem. A* **104**, 1876–1889.
- Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403.
- Gillespie, D. T. (1977a). Concerning the validity of the stochastic approach of chemical kinetics. *J. Stat. Phys.* **16**, 311–319.
- Gillespie, D. T. (1977b). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361.
- Gillespie, D. T. (2001). Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* **115**, 1716–1733.
- Gracheva, M. E., Toral, R., and Gunton, J. D. (2001). Stochastic effects in intercellular calcium spiking in hepatocytes. *J. Theor. Biol.* **212**, 111–125.
- Green, N. C., Rambaldi, I., Teakles, J., and Featherstone, M. S. (1998). A conserved C-terminal domain in PBX increases DNA binding by the PBX homeodomain and is not a primary site of contact for the YPWM motif of HOXA1. *J. Biol. Chem.* **273**, 13273–13279.
- Greenwald, I. (1998). LIN-12/Notch signaling: lessons from worms and flies. *Genes Dev.* **12**, 1751–1762.
- Hamburger, V., and Hamilton, H. (1951). A series of normal stages in the development of the chick embryo. *J. Morphol.* **88**, 49–92.
- Hill, A. V. (1910). Possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *J. Physiol.* **40**, iv–viii.
- Hynne, F., Danø, S., and Sørensen, P. G. (2001). Full-scale model of glycolysis in *Saccharomyces cerevisiae*. *Biophys. Chem.* **94**, 121–163.
- Ko, M. S. (1992). Induction mechanism of a single gene molecule: stochastic or deterministic? *BioEssays* **14**, 341–346.
- Kramers, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**, 284–304.
- Langston, A. W., and Gudas, L. J. (1992). Identification of a retinoic acid responsive enhancer 3' of the murine homeobox gene Hox-1.6. *Mech. Dev.* **38**, 217–227.
- Lauffenburger, D. A., and Linderman, J. J. (1993). "Receptors." Oxford Univ. Press, Oxford.
- Lumsden, A., and Krumlauf, R. (1996). Patterning the vertebrate neuraxis. *Science* **274**, 1109–1115.
- Maconochie, M., Nonchev, S., Morrison, A., and Krumlauf, R. (1996). Paralogous Hox genes: Function and regulation. *Annu. Rev. Genet.* **30**, 529–556.
- Maconochie, M. K., Nonchev, S., Studer, M., Chan, S. K., Popperl, H., Sham, M. H., Mann, R. S., and Krumlauf, R. (1997). Cross-regulation in the mouse HoxB complex: The expression of Hoxb2 in rhombomere 4 is regulated by Hoxb1. *Genes Dev.* **11**, 1885–1895.
- Maden, M. (1999). Heads or tails? Retinoic acid will decide. *BioEssays* **21**, 809–812.
- Mader, S., Chen, J. Y., Chen, Z., White, J., Chambon, P., and Gronemeyer, H. (1993). The patterns of binding of RAR, RXR and TR homo- and heterodimers to direct repeats are dictated by the binding specificities of the DNA binding domains. *EMBO J.* **12**, 5029–5041.
- Marshall, H., Studer, M., Popperl, H., Aparicio, S., Kuroiwa, A., Brenner, S., and Krumlauf, R. (1994). A conserved retinoic acid response element required for early expression of the homeobox gene Hoxb-1. *Nature* **370**, 567–571.
- McAdams, H. H., and Arkin, A. (1998). Simulation of prokaryotic genetic circuits. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 199–224.
- McQuarrie, D. A. (1967). Stochastic approach to chemical kinetics. *J. Appl. Prob.* **4**, 413–478.
- Meinhardt, H., and de Boer, P. A. J. (2001). Pattern formation in *Escherichia coli*: A model for the pole-to-pole oscillations of Min proteins and the localization of the division site. *Proc. Natl. Acad. Sci. USA* **98**, 14202–14207.
- Murphy, P., and Hill, R. E. (1991). Expression of the mouse labial-like homeobox-containing genes, Hox 2.9 and Hox 1.6, during segmentation of the hindbrain. *Development* **111**, 61–74.
- Niederreither, K., Subbarayan, V., Dolle, P., and Chambon, P. (1999). Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. *Nat. Genet.* **21**, 444–448.
- Nonchev, S., Maconochie, M., Vesque, C., Aparicio, S., Ariza-McNaughton, L., Manzanares, M., Maruthinar, K., Kuroiwa, A., Brenner, S., Charnay, P., and Krumlauf, R. (1996a). The conserved role of Krox-20 in directing Hox gene expression during vertebrate hindbrain segmentation. *Proc. Natl. Acad. Sci. USA* **93**, 9339–9345.
- Nonchev, S., Vesque, C., Maconochie, M., Seitanidou, T., Ariza-McNaughton, L., Frain, M., Marshall, H., Sham, M. H., Krumlauf, R., and Charnay, P. (1996b). Segmental expression of Hoxa-2 in the hindbrain is directly regulated by Krox-20. *Development* **122**, 543–554.
- Phelan, M. L., Rambaldi, I., and Featherstone, M. S. (1995). Cooperative interactions between HOX and PBX proteins mediated by a conserved peptide motif. *Mol. Cell. Biol.* **15**, 3989–3997.
- Poolman, M. G., Ölçer, H., Lloyd, J. C., Raines, C. A., and Fell, D. A. (2001). Computer modelling and experimental evidence for two steady states in the photosynthetic Calvin cycle. *Eur. J. Biochem.* **268**, 2810–2816.
- Popperl, H., Bienz, M., Studer, M., Chan, S. K., Aparicio, S., Brenner, S., Mann, R. S., and Krumlauf, R. (1995). Segmental expression of Hoxb-1 is controlled by a highly conserved auto-regulatory loop dependent upon exd/pbx. *Cell* **81**, 1031–1042.
- Sham, M. H., Vesque, C., Nonchev, S., Marshall, H., Frain, M., Gupta, R. D., Whiting, J., Wilkinson, D., Charnay, P., and Krumlauf, R. (1993). The zinc finger gene Krox20 regulates HoxB2 (Hox2.8) during hindbrain segmentation. *Cell* **72**, 183–196.
- Singer, K. (1953). Application of the theory of stochastic processes to the study of irreproducible chemical reactions and nucleation processes. *J. R. Stat. Soc. B* **15**, 92–106.
- Studer, M., Gavalas, A., Marshall, H., Ariza-McNaughton, L., Rijli, F. M., Chambon, P., and Krumlauf, R. (1998). Genetic interac-

- tions between *Hoxa1* and *Hoxb1* reveal new roles in regulation of early hindbrain patterning. *Development* **125**, 1025–1036.
- Studer, M., Popperl, H., Marshall, H., Kuroiwa, A., and Krumlauf, R. (1994). Role of a conserved retinoic acid response element in rhombomere restriction of *Hoxb-1*. *Science* **265**, 1728–1732.
- Stundzia, A. B., and Lumsden, C. J. (1996). Stochastic simulation of coupled reaction-diffusion processes. *J. Comp. Phys.* **127**, 196–207.
- Swindell, E. C., Thaller, C., Sockanathan, S., Petkovich, M., Jessell, T. M., and Eichele, G. (1999). Complementary domains of retinoic acid production and degradation in the early chick embryo. *Dev. Biol.* **216**, 282–296.
- Vesque, C., Maconochie, M., Nonchev, S., Ariza-McNaughton, L., Kuroiwa, A., Charnay, P., and Krumlauf, R. (1996). *Hoxb-2* transcriptional activation in rhombomeres 3 and 5 requires an evolutionarily conserved cis-acting element in addition to the *Krox-20* binding site. *EMBO J.* **15**, 5383–5396. [Published erratum appears in *EMBO J.* **15**, Dec. 1996].
- Wilkinson, D. G. (1993). Molecular mechanisms of segmental patterning in the vertebrate hindbrain and neural crest. *BioEssays* **15**, 499–505.
- Wilkinson, D. G., Bhatt, S., Cook, M., Boncinelli, E., and Krumlauf, R. (1989). Segmental expression of *Hox-2* homoeobox-containing genes in the developing mouse hindbrain. *Nature* **341**, 405–409.
- Zlokarnik, G., Negulescu, P. A., Knapp, T. E., Mere, L., Bures, N., Feng, L., Whitney, M., Roemer, K., and Tsien, R. Y. (1998). Quantitation of transcription and clonal selection of single living cells with beta-lactamase as reporter. *Science* **279**, 84–88.

Received for publication November 16, 2001

Revised March 7, 2002

Accepted March 8, 2002