Chapter 5: Summary

> Simple interactions can have consequences that are not predictable by intuition based on biological experience alone.
>
> - Lee Segel, 1980

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 produced results similar to *in vivo* experiments. During the course of building the model, the *in silico* investigations suggested that an experiment concerning the connection of retinoic acid and *Hoxa1* would be enlightening. A new experiment was designed to investigate the interaction of these elements *in vivo*, and the corresponding experiment was performed in the model. The resulting data suggested that an implementation decision was incorrect. Based on these results the model was modified to encompass the new data, without losing the fit to the original data set.

In addition, the *in silico* experiments yield intriguing predictions that have yet to be thoroughly examined biologically. For example, the mutation experiments in which 5' RARE is mutated predicts that *Krox20* expression is down-regulated in rhombomere 5 (Figure 3.9C). The simulation also suggests that when *Hoxb1* is mutated, there is an up-regulation of *Krox20* in rhombomere 4, and a down-regulation of *Hoxb2* and *Krox20* in r5 (Figure 3.8C). The formal nature of the model calls attention to these simple test

experiments, and checking predictions will lead to valuable insight into the regulatory network.

If the model predictions are correct, the tool will allow a deeper investigation into the nature of the components and allow researchers to ask more complicated questions about the nature of the interactions. On the other hand, if the model predictions turn out to be incorrect (as was the case in Chapter 4), the experimental data leads to a refinement of the model that incorporates the new results. The revision will then offer different predicted relationships that will stimulate further experiments. 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 the simulation: as the components of the model are given greater support, it can be used to perform *in silico* experiments to identify the *in vivo* experiments that will be the most enlightening.

In addition to serving as an organizational tool for presenting newly established interactions, the model can also be used to investigate hypothesized molecular interactions. This was the case for the *Krox20*/*Hoxb1* connection that was the basis for the experiment described in Appendix A. 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 updating the regulatory network (Figure 3.5) accordingly. An example of this is work currently in progress that seems to suggest *Krox20* contains an auto-regulatory element (P. Charnay, personal communication).

It should also be possible to extend this model in ways that are not only spatial and temporal, but which incorporate more of the known biochemistry of the system. For example, extending the model to include the next segment anteriorly, rhombomere 3, would allow an investigation into the early r3 expression of *Krox20* (Schneider-Maunoury et al., 1993). On the temporal front, it would be instructive to include 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 could include adding more genes, implementation of the mRNA modification and transport steps, and a better characterization of the genes or cofactors. Adding *Hoxa2* is an obvious choice because of the connection to the genes already in the network: it has been shown that *Krox20* is directly involved in the transcriptional activation of *Hoxa2 (Schneider-Maunoury et al., 1997).* New information concerning these genes appears on a regular basis and that provides the information for a better characterization. For instance, it has recently been observed an early low level of *Hoxb2* expression in rhombomere 5 appears to be due to a retinoic acid response element on the *Hoxb1* 3' RARE (R. Krumlauf, personal communication). All of these improvements will allow for a better understanding of the interaction and timing of the events.

There is also reason to 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. It has also been suggested that this may be

influenced by different basal transcription rates between the species (R. Krumlauf, personal communication). Once the mechanisms for *Hoxb2* regulation are in place, it would be possible to use the model to explore this issue. An investigation addressing this would include changing the basal transcription rates, the binding affinity parameters, and experimenting with different transcription factors configurations.

Conclusion

This thesis has shown that a tight coupling of modeling and experimental work provides a valuable framework for investigating biological problems; a framework that will become even more valuable as the amount of data increases. The act of constructing the model identified interesting biology questions, and the answer to one of those questions was used to enhance the model. Once the model was complete, the *in silico* experiments continued to identify potentially interesting biological questions.

The investigation into the early *Hox* genes also shows the success of using a stochastic simulation algorithm to model a gene regulatory network. This is especially important in situations where the fluctuations in the system appear to be a factor, because the stochastic approach is able to incorporate them in a physically intuitive and meaningful way. This investigation has also demonstrated that the SSA methodology has a wider applicability than the previous intracellular investigations. It can be adapted to encompass intercellular interactions, and the use of a priority queue to time order the multi-cellular system is an important addition to the method. The laboratory work stimulated by the model has yielded important biological results. The repression experiment in Appendix A shows that, as it stands, the construct does not successfully

repress *Hoxb1*. The RA perturbation experiment in Chapter 4 suggests that the response of *Hoxa1* to RA is concentration dependant.

It is expected that continued efforts in refining and using these sorts of models will result in a greater understanding of how computer simulations can be used to produce new biological insights. It is hoped that the success of this model will encourage more biologists to investigate the benefits of computer modeling in general, and stochastic simulation in particular. There is evidence that this work is already being noticed in the biology community: the author recently discovered that an article destined for the journal *Developmental Biology* referenced this work.

In a lesson for the mathematicians, this work also demonstrates a common problem with working in biology, one that was addressed in the general comments about modeling in the first chapter. There are too many "right" models, and the available laboratory data does not always allow for the ability to distinguish between them. This was the case with the first incarnation of the model: using a Hill function to produce an activated form of *Hoxa1* was reasonable choice given the information in the literature. Also supporting this choice were the results of the model: the simulation reproduced the wild type expression pattern, and computer perturbations yielded results similar to their laboratory counterparts. When new data were generated that tested this component, it was shown that the original implementation was not correct, and the model was changed to capture the dependence of *Hoxa1* transcription the quantity of transcription factors in a more explicit way. The new model is therefore better in so far as it captures more of the laboratory data. However, as is seen in the similarity between Figures 3.6 and 4.5, the models cannot be distinguished from each other on the basis of the output alone. This

shows the importance of the laboratory work in generating data that clarifies aspects of the model.

Finally, the systems biologists should see this work as a successful example of what they have been preaching: an integrative approach to biology problems will provide insight into how the systems behave. Insight that is not possible from approaching the problem using modeling or laboratory experiments alone. As more such successful interconnected effort appear, it is hoped that both biologists and mathematicians will look beyond the difficulties of interdisciplinary work that is mentioned in the quote from David Botstein at the beginning of Chapter 3, and instead focus on its enormous benefits to both fields.

References for Chapter 5

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,** 1885-95.

Schneider-Maunoury, S., Seitanidou, T., Charnay, P., and Lumsden, A. (1997). Segmental and neuronal architecture of the hindbrain of Krox-20 mouse mutants. *Development* **124,** 1215-26.

Schneider-Maunoury, S., Topilko, P., Seitandou, T., Levi, G., Cohen-Tannoudji, M., Pournin, S., Babinet, C., and Charnay, P. (1993). Disruption of Krox-20 results in

alteration of rhombomeres 3 and 5 in the developing hindbrain. *Cell* **75,** 1199-214.

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 [published erratum appears in EMBO J 1996 Dec 16;15(24):7188]. *Embo. J.* **15,** 5383-96.