6 The Bimodal Model-Building Strategy

As ISB develops as a science, it continues to face a classical methodological problem that has been present since its inception: how to manage and integrate wet-lab experimentation with model-building. As I have discussed in chapter 5, the enterprise of ISB is to model large-scale biological systems using modern computational and modeling techniques. In that undertaking, modelers are dependent on bioscientists for data to build models, for experimentation to validate models, and for biological expertise to understand the possibilities and limitations of model-building. ISB is far from settled on what are the best methodological practices or on what are the best modes of research organization to further its model-building practices (see, e.g., Calvert 2010; Calvert and Fujimura 2011; Nersessian and Newstetter 2013; O'Malley and Dupre 2005). As a field in development, it has some flexibility to experiment with respect to its methodological practices. The labs we have investigated belong to the area within ISB in which the lack of data of sufficient quantity or quality are among the foremost problems. Contrary to the widespread 'omics rhetoric, these researchers rarely have easy or sufficient access to high-throughput, time-series data. In chapter 5, we saw how this data problem runs through the range of constraints within which ISB modelers have to work. There I cast the process of managing the complexity of the model-building task as "adaptive problem-solving." Adaptive problem-solving ranges from specific practices of individual researchers to strategies by lab directors to organize their labs in the configurations they consider most effective to facilitate model-building. In this chapter I examine how the lab C director has organized her lab primarily around what we have called the "bimodal strategy": modelers are trained to conduct the wet-lab experiments they need to build their models.

Lab G was organized to comprise modelers, with backgrounds in engineering or applied mathematics, who collaborate with experimentalists external to the lab. The adaptive problem-solving in lab G relies totally on what highly mathematically and computationally skilled modelers can do to build a computational model. What we call the "unimodal strategy" is the typical organization of ISB model-building. As we saw, in unimodal model-building, adaptative problem-solving takes the form of such practices as reducing the scope of the problem; taking advantage of the affordances of simulation to build out pathway networks and the model, as well as to validate the model; and using and developing sophisticated algorithmic techniques to fit parameters in the absence of data.

Most often the modelers are collaborating with bioscientists in different labs, but there are some labs that are organized to have both within them. With this organization, the hope is that collaboration might run more smoothly, with data produced in a timely manner. The lab G director was skeptical. He thought this organization unlikely to work unless you had many experimentalists per modeler, and large amounts of funding to support them. He pointed to the well-known lab of Douglas Lauffenburger at MIT, which has around fifty members to accommodate modelers and experimentalists in productive collaboration as a prime exemplar of such an effective lab organization. Further, he contended, the biological systems investigated would be limited by the constraints of the experimental setup. We did not investigate this kind of lab. When we learned that lab C had a wet lab, we anticipated it to be that kind of lab, so we were surprised to discover its novel bimodal strategy. We decided to continue with it because it seemed a potentially important attempt at hybridization in ISB, despite the fact that it meant our project would not be studying pure experimentalists.

The lab C director was trained in combined research labs, where she was the only modeler in her PhD lab and one of several modelers and experimentalists in her postdoc lab. As I detail in the next section, she had unusual training as a "hybrid" researcher in ISB, having learned, sequentially, first to do modeling and then to do wet-lab experimentation. To organize her own lab C, she decided on an unusual form of adaptive problem-solving in which modelers were to be trained concurrently to conduct their own biological experimentation as part of the model-building process. The "coupling" of modeling and experimentation builds a kind of distributed cognitive-cultural system with some epistemic affordances and limitations

different from those of lab G. Although both labs build mesoscopic models, lab C, as we will see, builds models more "bottom up" by accumulating parts and dynamics of systems through the interaction of modeling and experimentation to solve problems. In contrast, lab G uses the averaging power of power laws and the parameter flexibility of the models it builds to close "downward" on a satisfactory representation.

6.1 Lab C: Redox Systems Biology

Lab C had been in existence less than three years when we entered. Its lab director was, then, an assistant professor, who, as with the lab G director, described for us an original and serpentine route to becoming a systems biologist, which was still not an established field when she began, even though she was much younger than he. By the time we met her, she had developed into the experienced kind of hybrid researcher the BME program was envisioning. However, even from the perspective of the senior faculty in that program, as well as from the perspective of her developing field of ISB, her vision to build a lab that trained graduate students to perform research that required they simultaneously learn to build systems models and to conduct wet-lab experiments to investigate complex biological processes was a high-risk undertaking.[1] In her own training, she had learned to do these sequentially, and she felt strongly that path had held her back: "*I tell my students to never do this because you should always do these things in parallel. It kind of delayed my graduation date because I ran into all the learning curve issues that [my] early graduate students face, only here, I was 4.5 years in and starting from scratch.*"

As the director recounted her learning trajectory, she had set out in college to be a biomedical researcher, having decided not to be a doctor, and so began as a biology major. Because she had tested out of freshman biology through her high school AP exams, her college education started with the required physics (thermodynamics), calculus, and chemistry courses. The thermodynamics course was full of premed students, which most of the biology majors were at her institution. She found what she experienced as a "*cutthroat environment*" of premed culture, where "*it's a big game of 'if I didn't get an A on the first test, I drop the class.'*" This was off-putting, so she decided to change her major. During a summer internship at a medical institution with a physicist who conducted research on protein structures,

she discovered you "*don't have to do a bio major to do bio research*." In what she described as "*flipping through the course schedule . . . in a kind of process of elimination*," she discovered that the nuclear engineering department had a radiological sciences track, which required cellular physiology and included other biology selections a student could take as part of their degree requirements. She called it a "*general engineering degree*" that has a curriculum "*very close to today's BME majors*." With that major she was able to take enough biology courses for a biomedical engineering minor (there was no major yet). For her undergraduate research, she joined a lab that was building computational models to investigate the effects of X-rays passing through soft tissues and did a modeling thesis.

She decided to focus on computational modeling with a bioengineering PhD. In the lab she selected, she "*wound up with a professor that wasn't a modeler at all. . . . He's a jack of all trades*." Her supervisor conducted research on muscle physiology, mainly with NMR spectroscopy. She described herself as "*the only modeler in the group—everyone else was an experimentalist*." Her research focused on building models of the temporal dynamics of supply and demand of phosphorous (P-31) metabolites in muscle contraction based on the data the lab collected. She described her modeling (ODE) work as requiring "*a lot of applied mathematics*" that used "*third party software*." The final step in her development toward being a bimodal researcher came when she found that to test her model, she "*had utilized all the literature possible*" and that "*there were certain things we couldn't measure with a magnet [by NMR]*." She stated that she realized that the only way to put her model "*through stringent tests to see if some of the things that were emerging as properties of the system actually happen was if I did some of the experiments myself*." So, at the point where she should have been graduating, she began a whole new line of wet-lab research with mouse muscles. Although she characterized this as a setback (quoted above), she also saw it as a "*opportunity*," since her adviser was an expert experimentalist and she "*hadn't taken advantage of that aspect of his skillset*."

She conducted postdoctoral research in a lab that specialized in modeling protein signaling networks and "*where a lot of mathematical and modeling advances were being made*." Signal transduction, basically, is a process that communicates a message to a cell to grow, divide, alter metabolism, and so forth, in which proteins are the main actors. That lab comprised both modelers and experimentalists who worked in collaboration. She decided, in the

context of this lab, to work only as a modeler on an immunology project using T cells, in part because "*it would be kinda cool, because I didn't know anything about it*." To figure out how to model such systems, she had to learn a lot of immunology and develop a new, "*more of a statistical*," modeling method based on partial least squares, both of which would prove important for the research projects in her own lab. Her two model-building experiences with quite different biological systems also made her aware that there was a "*disconnect*" between metabolic pathway and signaling pathway research, even though in a biological system metabolic and signaling processes are integrated in vivo. She felt her training offered "*a way of integrating those things together*"—a way of building models to "*understand the way all these things are regulated*." The more she investigated, the more she both "*realized, wow, no one studies this because it's just too complicated*" and that this is "*all the more reason why you need to do these computational approaches in parallel with the experimental while you were making progress*." So, at the outset of establishing a research program, she saw the "*parallel*" approach as both a necessary and a justified risk, because the novelty of her project meant that the kind and quantity of data required for the model to do the "*integrating*" could not only be found through searching the literature and databases, but also would need to be determined and collected as the models were being built. She stressed in her interviews, as well as several of her research presentations we attended, that her lab's wet-lab research was in the service of model-building.

That said, every presentation the director made of her lab, including the numerous visual representations she made of the lab's research, put the biological dimension of the research at the center, rather than the model-building, whereas the director of lab G always gave the model pride of place. This difference in emphasis, which reflects how they conceptualized their research, as well as organized their labs, can be seen quite clearly in the representation each drew when we asked them to "draw us a picture of your lab, that is, the problem-space of the lab's research." Unfortunately, the lab G director finished his first and showed it to the director of lab C, so she chose the same format. But it turned out to be a good thing, because it made for a striking contrast. His depiction (figure 6.1a) placed the model and methods for model-building at the center of the lab's activities, while hers (figure 6.1b) placed the biological context there. Hers also depicted the technologies the lab was developing to carry out their experimental

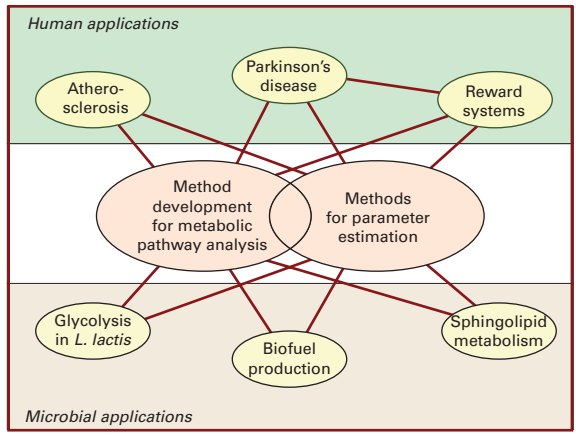


Figure 6.1a
The lab G director's schema of his research problem space. He placed the lab's goals to develop methods for model-building at the center, with various kinds of biological systems feeding into those developments. Lab G's experimental collaborators were in university, medical school, and industry experimental labs locally and worldwide.

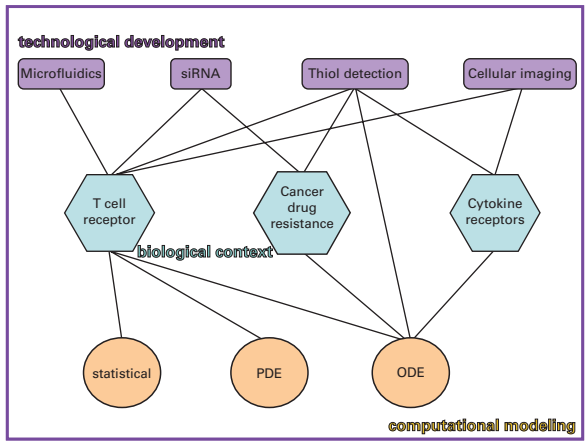


Figure 6.1b
The lab C director's schema of her research problem space. She placed the "biological context" of the lab's research at the center, while noting that they use various kinds of model-building formats and develop experimental technologies. Lab C had experimental and engineering collaborators in different fields within her own university, including chemists and electrical engineers, and at a nearby medical school. They both listed the names of researchers associated with each of the systems or methods, which I have removed.

research, as well as the different kinds of models they built, most using off-the-shelf packages. In addition, their recommendations for reading to help orient us on their research also underscored this difference between the labs. The lab G director recommended a graduate text on modeling biochemical systems using the BST framework and the lab C director, an undergraduate immunology text.

As mentioned in chapter 1, these lab directors were not collaborators, though they had significant interaction, and occasionally served on the dissertation committees of each other's students. The students in these labs did not interact much. Our research project introduced the only occasion we know of in which they presented their research to one another in a joint lab meeting. Since we were studying both labs, the directors thought it might be a good idea for their students and their research projects to be introduced to one another, and so they arranged two joint lab meetings for these purposes. The students did, of course, attend one another's presentations of research in departmental community forums and, often, dissertation defenses. In working with our group, the lab directors developed and taught together a graduate-level introductory systems biology modeling class.

As I noted in chapter 5, the lab G director stated there was a "*philosophical divide*" between ISB modelers who typically take the unimodal approach of his lab, with collaborators at a distance, and those who work more closely connected to experimental research, and he especially noted the divide with respect to the bimodal approach of lab C. Importantly, lab G and lab C have differences in their epistemic agendas. As we saw in chapter 5, the lab G director was quite emphatic about the importance of his lab's aim to advance and enrich mathematical theory and mathematical analysis to further investigation of complex biological systems. Sophisticated biosystems modeling requires rigorous and novel mathematical analyses that can capture a wide range of nonlinear behaviors within a tractable formalism that keeps complexity under control. To produce these kinds of models and the tools for mathematical analysis, however, requires high levels of applied mathematical and computational skills that, in the lab G director's view, cannot be achieved in combination with doing one's own experiments. As he saw it, the trade-offs are, "*If you do the experiment yourself, you know what the data are like; you know how reliable they are. You know the kind of assumptions you made in order to produce the data. . . . So, you get a better idea about the whole context. On the other hand, life is complicated, and to do good modeling is a full-time job;*

to do experiments is a full-time job. And if you don't want to do two full-time jobs, then something will suffer from it." From his perspective, lab C's modeling "*suffers*" in the respect that "*the models that are being developed . . . are . . . by and large, off-the-shelf type modeling approaches that are, not always, but that are often rather simplistic. . . . Our models are at least going into much more depth.*" But, as we saw, parameter-fitting was a significant, overarching problem in lab G research, because adequate data and data of the right kind for the system they are modeling are often missing. Lab C modelers rarely experienced these problems. Further, as we saw, the need for unimodal modelers to collaborate with experimentalists is fraught with difficulties that increase the complexity of the model-building task. As we will see in this chapter, lab C's epistemic agenda tracked more closely with contributing to medical research and molecular biology through getting a grip on targeted biological systems.

The "*philosophical divide*" the lab G director referred to might well reflect deeper attitudes in the field as a whole. As O'Malley and Dupré (2005) have argued, there are, in organization and in research strategy, divisions over the practices and aims of systems biology, even though they might not always be debated openly. Some systems biologists are relatively pragmatic and aim to use modeling as a tool to further develop molecular biology, while largely pursuing that field's theoretical agenda; others have a strong systems-theoretical agenda to advance the role of mathematics in systems analysis and to promote the development of a mathematical theory of biological systems. In the case of lab C and lab G, these divisions are acknowledged explicitly and are expressed in the form of laboratory organization that each director has chosen, and in the kinds of model-building practices they favor.

As mentioned earlier, lab C research is driven by a specific theoretical agenda: to determine whether and how particular biochemical systems play a key part in the regulation of cell signaling[2] and metabolic processes. The director sees this issue as critical to the advancement of molecular biology and physiology: "*So they are almost like two different camps in cell biology—all enzyme-based people cared about was how it got into the cell, all the metabolites that are involved, and didn't really think all that much about what was controlling the expression of these proteins or how the muscle was responding to any other cues from its environment or anything . . . and all these signaling people were interested in is what's connected to what and how this receptor is, you know, causing changes in gene expression, but they don't think about basic things like the energy supply in the cell. . . . So, there was this disconnect between the two*

camps." She claimed that she "*was thinking about this in a really different way*," likely because of her unusual hybrid training, which also provided her with "*a way of kind of integrating those things together*." Thus, as with the other lab directors we have discussed, she staked her claim to be conducting innovative, frontier research. But she also noted that the problem was recognized, if not addressed, in the field: "*It's not like there are no other people out in the world thinking this, otherwise I wouldn't have gotten these ideas. But I realized no one was trying to take this [integrative] approach*." She framed her overarching research goal as to make a contribution to molecular biology, but distinguished her lab as importantly different from those of molecular biologists: "*What we really specialize in—what can we do that no one else can do—is put it in terms of the context with respect to the rest of the network*."

This goal likely explains the director's clear preferences for working with more tractable and experimentally accessible systems than those of lab G. The models lab C builds tend to be smaller-scale, more mechanistic, models that help to demonstrate how specific mechanisms operate within the system. The modelers' preference for Michaelis-Menten or simple mass-action models of interactions engages with common representational techniques of molecular biologists.[3] Unlike in lab G, we heard few complaints in lab C about parameter estimation, because wet-lab experimentation by the modelers serves to keep the unknown parameters of their models mostly under control. As the director noted, with the bimodal approach, when faced with "*this issue of having more parameters than we're capable of fitting with the data, we have to say, 'ok, what data do I need to collect just to fit the model to these parameters?'*" Her goal was to train the modelers in her lab to be able to design and "*run the experiments under these conditions, do the analysis, and then plug these values into the model*."

The overarching epistemic goal of lab C is to understand cellular oxidation (a metabolic process) in regulating immunological and cancer cell signal transduction.[4] Such understanding has the potential to be used to develop interventions for numerous diseases associated with the oxidation state of cells (atherosclerosis, HIV, Parkinson's disease, lupus, cancer, and so forth) and can, especially, be utilized in personalized medicine. In particular, the lab conducts research on the impact of the redox environment on proteins. "Redox" is an abbreviation of reduction-oxidation, which is a chemical reaction that changes the oxidation state of atoms. Under normal physiological conditions, cells maintain a reduced oxidation environment. However, oxidizing molecules and free radicals are produced by cells as part of

physiological processes, or they can enter them, and can react with cellular components, including DNA, cell membranes, and proteins. Such reactions have been implicated in several diseases. Cells use enzymes to counteract these oxidants and proteins to mitigate the effects of oxidation. The communities that investigate redox and oxidative stress have only recently begun to appreciate the need to understand the interplay of these processes in order to determine the mechanisms of disease and to conduct therapeutic interventions.

Lab C's research focuses on the impacts of oxidants on proteins, which are part of signaling pathways. As we saw above, the lab is a pioneer in seeking to integrate these phenomena, as they are in the in vivo biological system. As she explained it, "*Oxidative stress is coupled with metabolism. . . . I realized that there is a way that the byproducts of oxidative stress, which are these reactive oxygen species, can bind to signaling proteins and affect the way they operate. For me, that's the kind of missing link, 'cause oxidative stress is controlled by metabolite levels. . . . I saw this as a really different perspective on the traditional signaling cascade.*" Modeling provides the means to study the dynamics of these metabolic and signaling processes in an integrated manner. However, given the novelty of this biological problem, the modelers need also to conduct experiments under specific conditions to obtain much of the data they need to build out the integrated pathway and find the parameters required to fit the model. In the period of our investigation, the lab's specific modeling problems within their redox agenda were both generated from the interests of the lab director and brought to her by experimental researchers outside of the lab.

Lab C is located in a new building that was designed to facilitate collaboration among labs. The lab spaces are largely open, with dividing walls surrounded by a wide corridor in which the expensive technologies in common use are housed. Lab C comprises a large wet lab where experiments are conducted, in which there is a walled-off cell culturing room and a dedicated space for conducting western blot assays, and a "grad cave" with cubicles where the graduate students do their modeling work on laptops and store their stuff. The wet lab has the typical accoutrement of a molecular biology lab, which includes pipettes, centrifuge, test tubes, a biohazards waste bin, a cryogenic freezer for the immortalized cancer cells they purchase in bulk, incubators, and wall pegs for hanging the clean white lab coats all members don when entering the space. Importantly, the wet lab was the center of

social as well as experimental activity in lab C. As one researcher reported, "*In doing experiments you just sort of gain an understanding of how the lab [lab C] runs. . . . I'm not even talking about just the technical skills that you gain—I'm talking about more of the social aspect of the lab. . . . A lot of stuff happens in the [wet] lab area not in the desk area where I do my computing*."

Cells of various kinds figure in the experimental research of the lab members. The immortalized cells line they purchase are, primarily, HeLa (cervical adenocarcinoma), JurKat (acute T-cell leukemia), and Caco-2 (epithelial colorectal adenocarcinoma). Given the sensitivity and cost of primary cell types, such as T cells and neutrophils, they use the lab members own cells from freshly donated blood drawn in the health center when these are needed, which they can maintain for a short period. There was considerable joking among the researchers around their donations and the characteristics of their donated cells. All these types of cell lines provide what the director called the "*model-systems*" of the research, because they are used "*in substitution of what may actually be occurring in normal [in vivo] cells*." She explained that in biological research, such cell lines are called "model-systems" because the processes of maintaining them alter them in some ways, and, further, some cell lines, for instance cancer cell lines, are not "*even normal to begin with*." However, all in vitro cell types still provide a "*fairly good representation of what's happening in the real cell*."

The lab absorbed the considerable expense to purchase its own Bio-Plex machine for immunoassays, which sits on a dedicated counter space, because the researchers use it so frequently for studies of temporal dynamics on the primary cells, which are difficult to acquire and age rapidly in vitro. The director thought the investment worthwhile, given that it can run eight different time points on one sample. As she commented with considerable enthusiasm when she demonstrated to us how it works, "*My machine can do these things simultaneously and then it's like 'Wow you got eight different measurements with one single time point, you got sixty-four measurements with this one input!' So, it's a way of generating the data that can supply our models*." The last statement is indicative of the way this lab differs, significantly, from a customary molecular biology lab, even though it has the look of one: the data are collected by modelers who need it to "*supply,*" or "*feed*" (a commonly used expression), their models.

The grad cave space has several whiteboards the students use for work, leaving reminders, playing games, and joking with one another. Unlike lab

G, the students were in the lab space much of the time, even the one dedicated modeler, and "the lab" had the feel of an active community. They often ate lunch together and reported gathering for social activities outside the lab. (Lab G researchers also reported such gatherings.) There were weekly research meetings with one or two presenters to update lab members on research or to troubleshoot experimental or modeling problems. There were also weekly "journal club" meetings in which the lab members discussed pertinent papers in the literature, selected by the members on a rotating basis. The lab director was frequently in the wet lab, where she conducted her own experimental research as well as supervised the student projects. So, she was usually on-site when students ran into difficulties or had questions. None of the students had conducted experimental research prior to entering the lab, nor had any developed biosystems models. Not all of the students were bimodal modelers when we entered, but by the time we ended our investigation all had followed that path. The process of developing skills in both modeling and experimentation had become central to what made one a part of the cognitive-cultural system of lab C.

When we arrived, the lab consisted of the director, three PhD students, nine undergraduates (all in the new BME major), and a research technologist, who had an MS in biology and who carried out the responsibilities of lab manager as well as conducted some experimental research in collaboration with the director and grad students. Although the lab membership expanded while we were conducting our research, because our time and resources were limited, we focused on "founding" members. The lab members were a diverse group internationally, spanning four continents. The grad students had undergraduate degrees in electrical engineering (C10), materials science and engineering (C9), and biotechnology (C7). Because there was no PhD degree in ISB, their current degree programs were located in electrical engineering, biomedical engineering, and bioinformatics, respectively, while the lab and the director were in the newly formed BME department. Interestingly, just as with the director, the students all characterized their degrees as "*general engineering*" in their initial interviews. C10, for instance, contrasted her degree program with electrical engineering programs in the United States, such as her current department, by saying, "*They learned us to learn, not to learn something*." I suspect that the fact that the broad-based engineering programs they came from had required less rigorous education in the high-powered applied mathematical skills than

we saw lab G members possessed might have attracted them to the research agenda of lab C and predisposed them to be more open toward learning biological experimentation methods.

The research technologist's undergraduate and MS degrees were in biology, and she had worked for several years in the biotech industry before joining lab C. We interviewed her and followed her research, as we did with the PhD students. She provided an interesting case as the only person trained as a biologist and, initially, solely engaged in experimentation. Near the end of our research, she transitioned to a PhD student, and we did continue to follow her progress for a case study on learning even after we had stopped intensive data collection. She intended at first to do only experimental work, but ended up building models as well, after she took the introductory biosystems modeling class we helped the lab G and C directors develop for the PhD program. I say more about her experience in chapter 7.

C10 and C9 were bimodal researchers and had been in the lab two and three years, respectively. C7 had just joined, did only modeling, and maintained he would not be doing experimentation, but, as our study ended, he, too, began experimental research because his modeling project needed data that were not available in the literature. To carry out her research, C10 needed to collect high-throughput data, and so her first project was to collaborate with members of an engineering lab to design and fabricate a microfluidics lab-on-a-chip device that she could use to collect sufficient experimental data for her modeling project. Since our primary interest was to understand the epistemic affordances and limitations of the bimodal strategy for research, as well as the challenges it poses for learning, we developed detailed, longitudinal case studies of the practices of C10 and C9, and a briefer one of C7.

A primary research contribution of all of the undergraduate researchers was to conduct western blot assays, which identify specific proteins and measure their amounts in a sample, for the graduate students to which they were assigned. The students assigned to C10 and C9 assisted in other aspects of their research projects as well. One undergraduate, who had joined when the lab started, had his own research project supervised by the lab director. The director had collaborations with researchers (engineers, chemists, biochemists, and medical researchers) located within her department and at a nearby medical school, some of whom we were able to interview.

As with lab G, lab C researchers identified functionally in accord with their epistemic practices rather than as systems biologists. To us, they identified primarily as modelers. When we asked the bimodal researchers, including the director, how they identified themselves to other researchers, they all responded that it depended on the person with whom they were talking: sometimes as a modeler, sometimes as both modeler and experimentalist. The students noted how their engineering education provided them with the skills for model-building and technology design. They expressed confidence in their ability to use and modify the third-party software, as well as do limited de novo coding when needed, but they also noted, explicitly, their technical skills were not at the level of the modelers in lab G—whom they called "*theoretical modelers*." The director and all the graduate students noted the importance of concepts, theory, and methods from control engineering, in particular, in their model-building practices. Control engineering, itself, could be called a "general engineering" area, since it is an interdisciplinary mix of various engineering fields, including electrical, mechanical, telecommunications, computer engineering, and product engineering.

All the graduate students said that learning to conduct experiments presented a significant challenge. Their research required specialized in-depth knowledge of a specific system, so they, too, felt that taking a number of biology courses, beyond those that their degrees required, would not be useful. Biological concepts and experimental techniques were the subject of most journal club sessions. With respect to experimentation, the standardization of molecular biology techniques and the availability of prepared assay components from vendors were significant factors that contributed to their ability to carry out experiments. They often commented that experimentation mainly comprised skills one needed to practice repeatedly, such as "*pipetting to make sure it was accurate*" and "*following a recipe*." The latter comment would elicit vehement objections from the research technologist, who had the uncanny ability to hear these and other comments she considered disparaging of the complexity and sophistication of biological research even when uttered quietly across the room from her.

As in the BME labs, anthropomorphic language figured prominently in discussions about the cells and their behavior in this lab. Lab members talked about seeing things from the "*perspective of the cell*" and of the need to keep cells "*happy*," especially so that they do not "*commit suicide*" (apoptosis). They often took the perspective of the cells when discussing their

behavior, using phrases such as "*if I were a cell*." The researchers exhibited the same kind of affective engagement with their cells that we saw in the BME labs. A vivid example is the explanation one researcher offered about why they need to stimulate the cells in an interview that turned into a three-way discussion with another researcher, who concurred: "*You have to stimulate them . . . so they are happy—and so it basically, because it is stimulated—it tells her [the other researcher], 'oh, I am useful, so I cannot commit suicide because somebody needs me.'*"[5] We interpreted such language as an expression of how their intensive interaction with the cells developed cognitive partnerships with them, as in the BME labs.

To examine the bimodal strategy, I first take a brief look at C10's design of the microfluidic device to demonstrate that an important dimension of lab C's ability to conduct their own experiments is the members' ability, as engineers, to create technologies for experimentation, as were depicted by the director in figure 6.1. Once the microfluidic device is built, it changes experimental biological practice and, in this instance, makes it possible for an engineer to more easily collect her own "*gold-standard*" time-series data, since it replaces a complex series of experimental manipulations. The data from the device allows the researcher to build detailed models that can make more accurate predictions than those built on scant data. Such predictions often lead to novel experimental manipulations, which create still more data for modeling, thus generating a positive feedback spiral in the direction of a more accurate model representation and deeper understanding of the system.

We followed the iterative and incremental design process for this device, as it was constructed and tested in experimentation with cells, through to its completion. We were not able to follow her research through to the phase where C10 conducted the experiments needed in the course of her modeling project. In this case study, we detailed the nearly three-year process through which C10 designed and fabricated a high-throughput device to automate the experimentation she would need to conduct to build a model of T-cell senescence. In addition to our coding of it as a part of the research in the lab, we recoded our data for the project, separately, as a case of engineering design (for a detailed analysis, see Aurigemma et al. 2013). The rest of the chapter focuses on an in-depth examination of how C9 used the bimodal strategy to couple experimentation and model-building in her investigation of the differential sensitivities of cancer cells to a chemotherapeutic drug, and its implications for our research themes.

6.2 "You Need Very Precise Stimulation at Very Precise Time Points": Turning Experiments into Devices

When we were initially told about the research to design a *microfluidic device* so the lab could carry out their own experimentation on T-cell senescence, we, of course, thought it might be something along the lines of the simulation devices in the BME lab, where we first heard the terminology of "device." However, once we understood that this was a high-throughput data collection technology and not an in vitro simulation device, we asked the director about the term. She explained that engineers use "device," generically, to mean "*the man-made object that's being used to manipulate cells and change them in some fashion. In our case, we are breaking them open at some point. . . . We're able to treat them, break them open, and collect all the outlet proteins that come out of it [microfluidic device].*" Thus, in vitro simulation, as carried out in labs A and D, is just one way in which a device can manipulate and change cells.

The lab director explained that they wanted to understand the possible role of redox processes in T-cell senescence (the aging of the cell that leads to an inability to replicate), especially with respect to a clinical application. Clinicians were starting to develop procedures to boost a person's immune system by harvesting their T cells, multiplying them in vitro, and then returning them to the patient. Although they had some success with the procedure, the rapid aging of the cells in vitro presented a considerable obstacle. The lab director suspected redox processes were the culprit, and C10's project was to investigate these processes in T-cell signaling. However, there were little data available on their behavior, so lab C needed to collect their own in order to build a model of this system. They needed primary T cells for the research, and the fact that they could obtain only small quantities of these, coupled with the fact that T cells age rapidly, presented problems for the research. They needed to be able to collect multiple data time points rapidly—more than were feasible by hand. High-throughput data collection technologies are designed to overcome these problems. Engineers develop what they call "lab-on-a chip high-throughput devices" (hereafter, LOC) to bring together in a single apparatus—and replace—a complex series of experimental activities that would need to be executed by researchers in biology labs. This technology has been a major contributor to the development of computational systems biology, since vast amounts of time-series data can be collected quickly and efficiently from a population

of cells (thus, the designation "high-throughput"). These LOC devices improve the accuracy and quality of measurements, as well. C10 decided to build a specific microfluidic LOC device to conduct her experimental research on T cells.

Initially, C10 came to a microfluidics lab in an electrical engineering department to do research on LOC design for a MS degree. Microfluidics engineering builds devices to precisely control and manipulate fluids that are geometrically constrained to a small scale. As C10 explained, "*When I arrived, I didn't know anything about microfluidics, about biology, and about research.*" When we met her two years later, after she had completed that degree, she had developed expertise in microfluidics and was learning T-cell biology. She had also transitioned to being a member of lab C, as well as of her original lab—indeed, she served as the bridge in a collaboration between them. She came to know lab C because the director had become interested in LOC devices. To develop computational models of cell signaling requires time-series data that are difficult to collect in benchtop experimentation, because signaling events happen rapidly (sometimes within twenty seconds of stimulating a cell). The LOC device can automate the stimulation of the cell and the collection of cell samples at different time points simultaneously. This automation improves data collection, particularly for early signaling events that occur right after stimulation, and signaling events that occur in quick succession, thus providing cleaner and richer data for modeling.

The problem C10 wanted to investigate using the device was to quantify senescence in T cells, specifically, to determine which biomarkers change in correlation with age. This research required primary T cells, and those she planned to use were collected freshly from human donors, which were available in limited quantities. These cells also immediately begin to age rapidly, and so can be used for experiments for only a few days. One of the advantages of the LOC is that C10 would need only a limited number of cells to collect a significant amount of data as compared to benchtop methods. To investigate senescence with this approach, the T cells need to be stimulated (mixed well with a reagent), which causes different proteins to form in the cells (as a result of the signaling process), and then measured at many time points, ranging from twenty seconds to twenty minutes after stimulation. With the data from the LOC, these measurements can be done at both the population level (a certain number of cells) and at the single-cell level. The measurement of proteins, itself, is not done in the LOC, but

separately, with biological instruments. The device freezes the cells' internal state at different time points by quenching the biochemical reactions in the cell. This is done by lysing the stimulated cells (adding a reagent that breaks open the cells, which creates population samples) and fixing them (adding formalin, which creates single-cell samples). Proteins, whose internal states are frozen by the LOC at different time points in the signaling process, can then be measured.

The LOC design process, in general, requires the designer to translate the goals and actions of an experimentalist executing a complex lab routine into mechanical procedures that can be accomplished by the device, which is only a few centimeters in size. In this case, the device C10 was building needed to automate three processes: (1) stimulate the cells (by mixing with stimulant); (2) freeze the cells internal state by lysing (by mixing with lysis buffer) half of the samples and fixing (by mixing with formalin) the other half; and (3) do this at precisely the right moments (as defined by the desired time points). In the initial stage of the design, she only considered lysing and added fixing toward the end of the design. One of her early design decisions was to have a modular design, one module for the mixing process and another for the freezing process. The two modules would be connected by tubing of different lengths, so that the liquid (stimulated cells in media) in each tube would take different amounts of time to reach the second module, where the biochemical reactions in the cells are then quenched. The varying tube lengths thus function as an analogue for different time points, turning time into space. C10 built the device in PDMS, which is a 3D CAD (computer-aided design) software program, using soft lithography.

Although I will not detail it here, the device design involved complex iterative and incremental processes through which C10 distributed her cognition across various kinds of representations she constructed as she created what would become the final LOC, which she would use to build a computational model of T-cell senescence (see Aurigemma et al. 2013 for a detailed analysis of these processes). These processes involved numerous interactions among components of a D-cog system that comprised C10's mental models, computational models that simulated design possibilities for various geometries for the modules, sketches, fabricated device prototypes, different cell lines, and visualizations, which included those created by tagging cells, computational visualizations, and sketches, as well

as numerous problem-solving sessions with lab mates and experimental collaborators in the microfluidic lab and lab C. The LOC device needed to accommodate and integrate engineering and biological constraints. C10 encountered numerous problems, and each time she would develop a computational model of the LOC (in MATLAB) to simulate the effects of design possibilities. The most difficult problem was to develop a solution for the geometrical configuration of the pressure drop chamber (PDC).

The PDC was designed, initially, as rectilinear channels folded in a rectangular zigzag pattern, which she thought a good solution to the engineering constraint that the PDC needed to be long and thin, but fit into the footprint of the device. When the device was tested with fluids, it worked perfectly. But, when she tested with JurKat cells, which they used in the design process because they are plentiful and longer lasting, they became stuck in the corners and were "*getting stressed*." As C10 stated her frustration, "*If you don't have cells, it's almost perfect. You put cells, nothing works anymore*." She tried various zigzag configurations with fewer turns, but the problem became even worse when she tried with primary T cells, which she discovered were larger than the JurKat T cells. The final design solution came at a lab meeting where she once again discussed the problem with the cells still getting stuck. C11 hit upon the idea of circles and then, echoing one another rapidly, the lab members proposed it could be "*like a spiral*," which they elaborated could be drawn from the center, "*the way a seashell is made . . . like nature*."[6] C10 balked at the idea because it would be "*painful to draw one—drawing a circle in AutoCAD is painful*," to which they responded "*but you only have to draw it once*." In the end, the spiral design solved the problem. C10 fabricated the final version of the LOC device and was ready to begin collecting data to build and test models of T-cell signaling processes when we finished our data collection.

This C10 case shows how the lab could use the affordances of their engineering skills to obtain the experimental data they needed to carry out, and manage the complexity of, the bimodal model-building strategy. In the next section I develop a case in which we followed how a researcher, C9, used the bimodal strategy of coupling experimentation and model-building to manage the complexity of modeling a biological system about which she and the lab director had formulated a novel hypothesis.

6.3 "As I'm Building the Model, Things [about Experiments] Are Popping Up in My Head": Investigating Cancer Cell Drug Sensitivities

C9's research provides an example of a distributed cognitive-cultural system with epistemic affordances different from those we have considered thus far in ISB. An examination of the path C9's research took over the course of her PhD is central to understanding how the bimodal strategy works as a form of adaptive problem-solving, which aims to manage the complexity of biosystems model-building. The bimodal strategy directed and determined C9's investigative possibilities. Through this strategy she was able to leverage affordances of both in silico simulation and wet-lab experimentation as an effective means to handle her complex problem-solving task. As is likely common in bimodal research, her problem-solving process took a circuitous path, driven by how the dynamics of her interactive methodological system, in particular, generated novel relevant phenomena. C9 was the first graduate student to enter the lab, and when we began our research, she was in her fourth year. However, until that time she had been building models with data obtained from the literature and from a large, unused data set provided by an experimentalist whom she called "*a mentor*" to both her and the lab director. So, we were able to follow her use of the bimodal strategy from start to finish.[7]

By the time we first interviewed her, C9 clearly saw herself as distinct in terms of the kind of research she undertook and the kind of researcher she was. She was investigating a problem about cancer drug sensitivities and thought "*it was very interesting because no one had approached it that way before*." She clearly saw herself and her lab as on the frontiers of research. In her undergraduate education, she had not taken any "*hard core biology*" courses, but only chemistry "*with sprinklings of biology*." She had also taken applied math and done some modeling with MATLAB software as part of her engineering degree. She stated that what had drawn her to biosystems engineering was the "*interesting*" idea of "*using math to describe biology*." She thought her lack of intensive training in molecular biology had allowed her to begin model-building without the typical experimentalist biases against modeling: "*So coming in I might not have had those biases, you know, that some experimentalists might have—so I had, maybe I was more of an open canvas for accepting modeling*." At the same time, she also distinguished herself from

what she called "*theoretical modelers*" (of the lab G type): "*We don't just come up with ideas and then just shoot them out there and wait for people do to them [wet-lab experiments].*"

The overall aim of C9's research was to try to explain different sensitivities in cancer cell lines to the chemotherapy drug doxorubicin (Dox). A clinical researcher at a medical school had brought this intriguing problem to the attention of the lab director. The director hypothesized that the sensitivities are somehow related to signaling functions of ROS, such as hydrogen peroxide within cells. She ventured this hypothesis on the basis of two plausible assumptions, which would inform C9's research. The first assumption is that this signaling system is sensitive to drugs like Dox, which generate hydrogen peroxide, and the second is that this signaling system modulates pathways relevant to cell apoptosis (self-initiated cell death) and proliferation. C9's research goal was to figure out whether redox systems play a role in Dox metabolism and Dox-mediated cell signaling and what that role might be. They hoped that this research would contribute to understanding the mechanism behind the differential sensitivity of cancer cells to Dox and, thereby, make a contribution to personalized cancer therapy. C9 saw this research as very much a joint project with the lab director, and often shifted between "*my*" and "*our*" when talking about it.

Over the course of her research, C9 constructed four models (which we have labeled Model 1 to Model 4). These were constructed consecutively and form the problem-solving tasks around which her research was organized. Although C9 carried out all the model-building and experimentation for her project, as the research progressed, she had extensive discussions about how to interpret what she was finding and how to proceed with the director and a senior biochemist from outside the lab who, as mentioned above, had become an informal mentor to the director and later joined C9's dissertation committee. The lab G director was also on her committee, although they did not have as much interaction. She was well into the work of building Model 2 when we arrived, and she had recently defended her dissertation proposal. So, I present abbreviated descriptions of Model 1 and 2 from her retrospective accounts, and focus on the latter models, which required her to do wet-lab experimentation, and for which we have concurrent data in the form of progress interviews, field observations, presentations, journal club discussions, and dissertation.

6.3.1 Phase 1: From Local Simulation to Global Simulation

C9's research began with the task of modeling a specific pathway thought to be an important instance of how the redox environment of a cell (the balance of oxidants and antioxidants, or of oxidized and reduced chemical agents) affects signaling processes within the cell. The particular signaling pathway with which her research began is that of the activation pathway NF-κB (nuclear factor kappa B), which is a transcription factor that regulates genes responsible for immunity and is involved in programmed cell death or apoptosis. C9 framed the "*working hypothesis*" of the research in the following way: "*So, our working hypothesis has always been that, some cells are preferentially resistant to Dox because Dox does something that leads to signal transduction with the cell that leads to, you know, anti-apoptotic transcriptions or something like that. And we know in the literature, also, that there are certain points in the NF-κB pathway that are ROS-regulated. So, then it didn't take too much to say, 'ok if you have this drug that induces ROS, it is a possibility that the ROS that's induced can affect this pathway within this cell that might lead it to be pro-survival.'*"

They suspected the NF-κB transcription factor might be relevant in this respect to the response of cancer cells to redox environments. She claimed this line of thinking was a new "*perspective*" on NF-κB, the pathway for which was well-known, because it had never been approached in terms of "*the underlying mechanisms that control*" calcium fluxes that influence NF-κB. C9 spent her first year constructing an accurate topology of the NF-κB pathway by searching through the literature and looking up or determining rate constants and chemical concentrations, building an ODE model on the basis of these, and then testing the model simulations against published data. By the end of the first year, she and the director reasoned that they had a "*pretty good model*" that simulated the interaction of the products of ROS processing with NF-κB and the regulation of these processes.

She reported that her presentations of this model at conferences drew reactions that took two forms. On the one hand she received encouragement for the basic concept the model seemed to illustrate: NF-κB is redox regulated. On the other hand, she encountered resistance to the fact that the model represented such a small fraction of the in vivo physiological process, and so the conditions used to build the model were "*very far from what occurs physiologically*." At this point their biochemist mentor encouraged them to shift their attention from the small NF-κB model to the whole system of redox regulation itself—that is, from the entry of ROS, such as

hydrogen peroxide, into the cell through to the processes by which ROS are processed and cycled. She could then situate the NF-κB model within this larger global system and would have a more accurate and realistic understanding of the smaller system through simulating, in particular, the environmental factors that influence the NF-κB model's various inputs or control points. The biochemist had been interested in how redox buffering and regulation could explain more general disease phenomena, and had collected relevant experimental data on cardiovascular disease, which he gave them for the general model.

The process of building Model 2 required C9 to build out the biological pathway from the existing literature and the biochemist's data. The pathway diagram from which she built the simulation model appears in her dissertation defense presentation, announced as the "*first ever comprehensive account of the mammalian antioxidant system*." Model 2 was an ODE model that used a simple model of enzyme kinetics (Michaelis-Menten) commonly assumed by modelers of metabolic systems (as we saw in lab G) to describe the changes in concentration of cellular redox buffering components. It contained four branches or pathways of H_2O_2 elimination. The model follows the entry of H_2O_2 into the cell and the processes of redox buffering that eliminate the incoming hydrogen peroxide. Her simulations with Model 2 replicated the basic dynamic data for key proteins in the network, glutathione (GSH) and thioredoxin (NADPH). The pathway/model does not mention NF-κB, but its modular structure would allow for Model 1 to be incorporated, as was her original intention in building the global model.

Although the model reasonably accounts for the structure of the system and its participating elements, the move to a "whole cell" perspective and a general model of redox buffering multiplied the number of components, which in turn multiplied the number of rate constants and concentrations she needed to unearth. In all, there were twenty-two kinetic parameters, and C9 spent nearly two years foraging through the literature for these parameter values. In this process she encountered the kinds of problems I discussed in the general description of lab G modeling practices. She was just finishing up this process when we met her. C9 was able to draw on successful simulations with Model 2 to make a number of inferences, including a sensitivity analysis of parameter responses and an analysis of the relative burdens of different proteins in peroxide removal. In the former instance, she analyzed membrane permeability as the factor that produced

the greatest network response in the model. In the latter instance, she found the protein thioredoxin (Trx1) does much greater antioxidative work than the protein glutathione (GSH). In possession of a global model of how a cell deals with oxidative stress, C9 was now in a position to move on to the problem of how cells might respond to different parameter configurations or, more specifically, how DOX might alter redox environments.

6.3.2 Phase 2: Building the Dox Bioactivation Network Model

In building Model 2, C9's task had been more or less to assemble the information in the form of a dynamic model, since the structural features of the system had already been well-established in the literature. In the next phase of her research, C9 faced the situation of needing to derive, experimentally, unknown features of the system. Once she began wet-lab experimentation, the literature took on a different role: she mainly went to the literature when something unexpected came up. In the process of experimentation, she discovered that some of her expectations, based on the literature and her previous models, about how the system should behave were wrong and that the essential mechanics of some of the components of the system were in fact unknown. She had to localize and isolate these inaccuracies and unknowns. This part of her discovery process started with Model 2, the global model of ROS. As she explained her problem-solving strategy up to that point, "*So, one of the main [questions] that came up was how does the cell . . . deal with these increases in oxidative stress? No one really knew about that so . . . we need to answer that question before we can move on to how is this drug able to alter the redox environment—to explain how something modulates that redox environment—why or how that works*." C9's plan, after she had built Model 2, was to move on to formulate a mechanism that could confirm their hypothesis that it was redox regulation of Dox that up-regulated NF-κB activity, thus helping insensitive cells survive the drug. She planned to model this mechanism by perturbing the values of hydrogen peroxide entering the Model 2 system, given that Dox was known to raise its levels, and then feed the outputs into the redox-sensitive points of the pathway for Model 1. But, to "*get [that] model up and running*," she would need to do "*a lot of [wet-lab] experimentation . . . to go in and get some of the rate constants and concentrations*."

Knowledge of how Dox functions to cause toxicity was in fact limited, although it was thought to intercalate DNA and thereby lead to apoptosis. One of its main known side effects is, though, the production of ROS,

which made it a candidate for C9's analysis. From the clinical researcher, who had introduced them to the problem, she knew that different cell lines have different responses or sensitivities to Dox. C9's goal was to explain these sensitivities: "*If you ask in the field how exactly does doxorubicin work, they would not be able to give you a set answer. Specifically, particularly since in different cell types you have different responses and no one understands why that is. So hopefully, by the time I am done with my dissertation, we can shed some light on whether or not the ROS generation portion of the doxorubicin story is something that leads to differences in the sensitivity of cells to—to this drug.*"

The clinical researcher provided two patient cell lines, EU1 and EU3, for her wet-lab experiments. EU1 is Dox-insensitive and was retrieved from a patient who had not responded to treatment. EU3's patient, however, had had a good response that led to remission. As mentioned previously, when she began the wet-lab experiments, C9 planned to combine her global model (Model 2) with the NF-κB pathway (Model 1). But then a significant problem emerged. It happened early in the course of putting together the findings from her initial experimental research to draft a paper she planned to submit to an experimental journal of her research linking Dox to the NF-κB pathway.[8] As she posed the problem, "*The issue that I think we are having or, I don't even know if it's an issue, not really sure yet 'cause we, we don't know what's going on, is that, the cell lines are what would be logically expected based on literature and what not, with regards to the cell that's insensitive to Dox should have more NF-κB activation—and that's what our model is sort of predicting, but experimentally, we are seeing kind of the opposite and we're not really sure how reconcile that.*"

This observation from her initial experimentation led C9 to an extended novel investigation in which she went back and forth between modeling and experimentation. Her experiments with Dox were producing data that were the opposite of what her model was giving and of what would be expected from the existing experimental literature. It is unclear (from our records) on which model she based her expectations, but they were most likely based on a combination of the Model 1 and Model 2, by extrapolating from different levels of ROS input to NF-κB levels. Those levels, according to her models and the experimental literature, were expected to be more active or up-regulated in the case of the insensitive EU1 cells, and down-regulated in the case of the sensitive EU3 cells. Her wet-lab experiments had, however, found more up-regulation in the EU3 cells. The extra ROS generated by Dox should regulate NF-κB, but at this point she was beginning to doubt whether NF-κB was even

"*a link in the story*"—or at least a link in the way she had anticipated. As with all the research undertaken in the bioengineering labs, C9 often encountered impasses or failures. As she recounted, "*Some days I come into work and I don't feel like doing anything. . . . I mean, if your models aren't working, experiments are failing, you need that extra push to keep you going*." What helped to motivate her were "*the camaraderie*" with the other researchers who were also "*struggling*" and also thinking about "*the patients that could benefit from it*." She also noted that, now, "*talking about it*" in our interviews "*reinforces what you're doing, so, I can go back and feel motivated*."

To isolate the conflict between her model output and her experimental data, she conducted further experiments with the cell lines, while at the same time she planned to use Model 2 to "*try and explain what we are seeing experimentally*." Her experimental study of ROS in those cell lines did confirm something expected; namely, there was more ROS in EU1 cells due to Dox, which ruled out that the problem was something in the general ROS or NF-κB mechanisms of the cells. She then found a paper that associated the generation of the toxic form of Dox with a ROS-reducing enzyme, NADPH, which she and the lab director thought to be an important experimental result. Her further experiments showed that the NADPH levels were low in EU3 and high in EU1. This suggested that both toxicity and the extra ROS emerged from the mechanisms by which Dox is activated into its toxic or reduced form. As a result, the lab director decided that an additional step was required in their problem-solving process: C9 needed to model the production of ROS by Dox, that is, to build a new Model 3, rather than simply input the estimated amount of ROS produced straight into the global Model 2, as she had planned. The new model-building would serve to open up an area of Model 2, otherwise black-boxed, in order to look for differences between EU1 and EU3. Instead of combining "*all these reactions into one single arrow [black box] and then just have an estimate of what the culmination of all of these reactions would be—we realized that there were areas where there are differences between the EU1 and EU3 cells particularly with their NADHP*."

Once again, she was turned away from her intentions to follow the downstream aspects of Model 2 and return to the role of NF-κB. Instead, she focused on the processes involved in the reception of Dox into the cell and the production of ROS. To build Model 3 she relied on a two-pronged strategy. In the first part, she constructed "*hypothetical mechanisms*" in close conjunction with her NADPH and other wet-lab experiments on different

interactions needed to "*build out the model*." For instance, she investigated the role of superoxide (SOD), O_2, and CPR (an enzyme that catalyzes redox reactions) when feeding Dox into the different cell lines and compared the results. As she went along, she simulated the hypothetical mechanisms she constructed and checked the simulation outputs against controlled wet-lab experiments on the different interactions. Finally, she ran simulations in interaction with experiments in order to fix parameters. The model contained ten parameters and ten initial conditions (concentrations) for what she called "*a relatively simple network*." Only two parameters needed to be estimated. Through this interactive process she built Model 3 as she constructed a bimodular pathway structure for Dox metabolism (figure 6.2).

In the second part of her strategy, once a model was up and running, C9 tested it by running simulations to perturb the model and performing corresponding controlled wet-lab experiments, which she cast as having "*physically experimented on the model*." For instance, she conducted computational simulation experiments to rule out other possible causes, such as

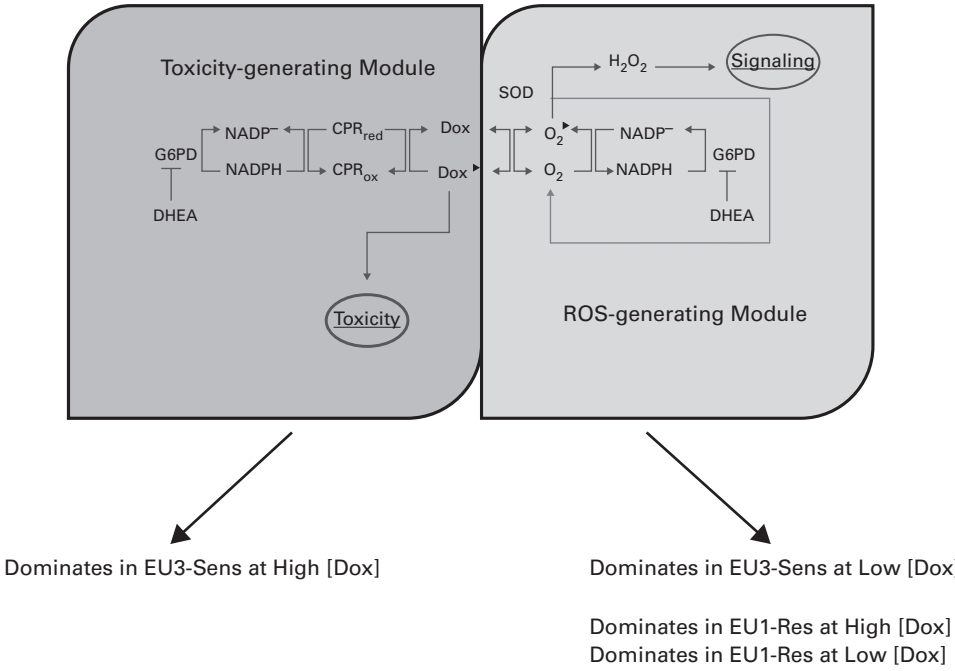


Figure 6.2
C9's proposed bimodular pathway of Dox metabolism in leukemia cells for Model 3.

efflux of Dox from the cell at different rates in the different cell lines. She also "*experimentally perturb[ed]*" the model to test it outside standard clinical ranges by simultaneously inhibiting the key elements of the cycles that had been determined by the model, namely the enzyme G6pd (important to the production of NADPH) and superoxide (SOD). In parallel with these simulations, C9 conducted wet-lab experiments on the two cell lines using chemical agents to inhibit these elements physically and determine the outcomes.

In accord with her experiments, the Model 3 simulations predicted that, in the case of low NADPH, redox cycling takes place. When NADPH is high, however, it absorbs the oxidative species that would otherwise be reduced by toxic Dox ($Dox^{\cdot}$ in figure 6.2) leaving toxic Dox to go free. The reductive conversion/toxic Dox production model is represented by the pathway module on the left in figure 6.2 and the redox-recycling model is represented by the module on the right. C9 stated that the agreement between her model and her experiments "*is again proving that our model has stood these different perturbations and interventions and it's still predicting what's happening. So, it further validates the fact that we might have actually gotten it right.*" Because of the successful outcome of the close interaction between computational model simulation and wet-lab experimentation, she was able to assert with a high degree of confidence that they had discovered the mechanism behind the relative sensitivity of the EU1 and EU3 cell lines. She had been able to trace back the cause of sensitivity to the levels of G6pd possessed by each line and, thus, to whether or not the line could replenish stocks of NADPH quickly enough to keep cycling Dox. This result had immediate potential clinical relevance, because G6pd is measured regularly by clinicians, and so the study indicated that its level could be used as a signal as to whether Dox treatment is likely to kill a patient's cancer cells or not.

There is a final twist in this story, however. C9 and the lab director wrote a paper on this research and sent it first to an experimental journal that rejected it because the reviewers thought a two-cell-line study would likely not be generalizable, and therefore would "lack impact." They then sent it to a well-known computational biology journal. In this case, they were surprised when the reviewers complained that the levels of Dox that they were using to study their cells were higher than those used clinically, even though she and the director maintained to us that they had seen numerous experimental papers that used their levels. So, C9 went back and

used lower values with the cell lines. These experiments produced radically different behavior, which—somewhat to their surprise—the model reproduced. She considered the fact that her model was able to reproduce her experimental observations as a powerful validation of her discovery of a new mechanism.

6.3.3 Phase 3: Wrapping Up: More Surprises, More Discoveries, Model 4

Even though her research thus far had led to the discovery of a novel mechanism to explain Dox sensitivity, with, of course, some help from the literature, she still found herself dealing with unexpected and unknown behaviors in the final phase of her research that again required interactive experimental and modeling work. In this phase, she had planned to return to her original problem, namely, the redox regulation of transcription factor NF-κB. The questions of how EU1 cells handled both the extra oxidative stress these cells generated under Dox treatment and how they survived whatever toxic Dox they generated still remained. The answers, they thought all along, had to lie in the redox regulation of the NF-κB pathway. She declared she felt she had now "*gone full circle*," and had finally done "*the preliminary stuff I needed to do in order to answer this question.*" Of course, that she had needed to build those models and do those specific experiments had not been clear to her at the outset, but had only emerged as she tried to formulate and tackle pieces of the problem.

C9 had already built a validated computational model for NF-κB (Model 1) based on the literature, so the issue now was to establish a connection between it and the other models and make whatever modifications might be required. To do this, she started a new line of experimentation. First, she established that Dox treatment in EU1 cells created higher levels of hydrogen peroxide. Second, she showed that Dox is correlated in these cells with increased NF-κB production. When she introduced antioxidants at the same time, NF-κB went down again, which established the relation between Dox-induced ROS and NF-κB levels. She continued in this vein to establish that adding antioxidants to cut the production of NF-κB pushed cell survival rates down. C9 told us that she thought this level of experimental detail or "*fine resolution*" is necessary to convince other researchers of these causal relations. It enabled her to show that the causes she hypothesized in her models were robust: "*There's nothing written in stone about the steps you take. You need to sort of say to yourself, 'ok, how fine of a resolution am I*

comfortable with, or how fine of a resolution do I actually need . . . to get other people to believe this is actually what's happening?'"

Once she established these relations experimentally, it was time to get into the "*nitty gritty*" details of the NF-κB pathway itself. Returning to Model 1, the question was which pathway elements would be modulated by the increase in ROS; in other words, what are the potential points in the NF-κB network that govern redox regulation (see figure 6.3, left pathway). With the model as the basis, she planned to work through wet-lab modulations of the different components using a specific antioxidant (NAC: N-acetyl cystine). She intended, initially, to use an experimental procedure of soaking up oxidants to see whether she could confirm model predictions about which pathway modulations are due to excess oxidants caused by Dox and which are not.

When she went through elements of the NF-κB pathway, such as NEMO (an essential NF-κB modulation gene), with a more detailed wet-lab examination adding Dox, she discovered one protein, IKK-β, whose s-glutathionylation levels changed when Dox was added. S-glutathionylation modification is caused by oxidation of the protein and thus provides an indication of the sensitivity of that protein to oxidative agents. She isolated the IKK complex (which binds IKK-β) as a ROS-sensitive component, circled in the NF-κB pathway above, and confirmed this in the wet lab by adding NAC and seeing IKK-β levels drop accordingly. However, when she tested with this antioxidant more expansively, she discovered that the effects of IKK-β varied non-monotonically with the levels of NAC in the system. This was unexpected and non-intuitive for her and the lab director, and put them again in the situation of having discovered complexity that they had not anticipated. Once again they would need to draw on C9's interactive experimentation and modeling strategy to untangle the knot. They decided C9 needed to build out Model 1 by adding a model of the IKK complex s-glutathionylation (Model 4) to see if they could explain the wet-lab findings. C9 built this model by working quite closely to the chemical details, targeting the "*mechanisms by which a protein such as IKK-β can be s-glutathionylated in the presence of an ROS promoting agent like Dox*."

C9's literature search yielded ten potential biochemical reaction mechanisms by which IKK-β s-glutathionylation could occur. She used both the conditions under which those mechanisms were observed in the literature in comparison to the mechanism she postulated, and what conditions were

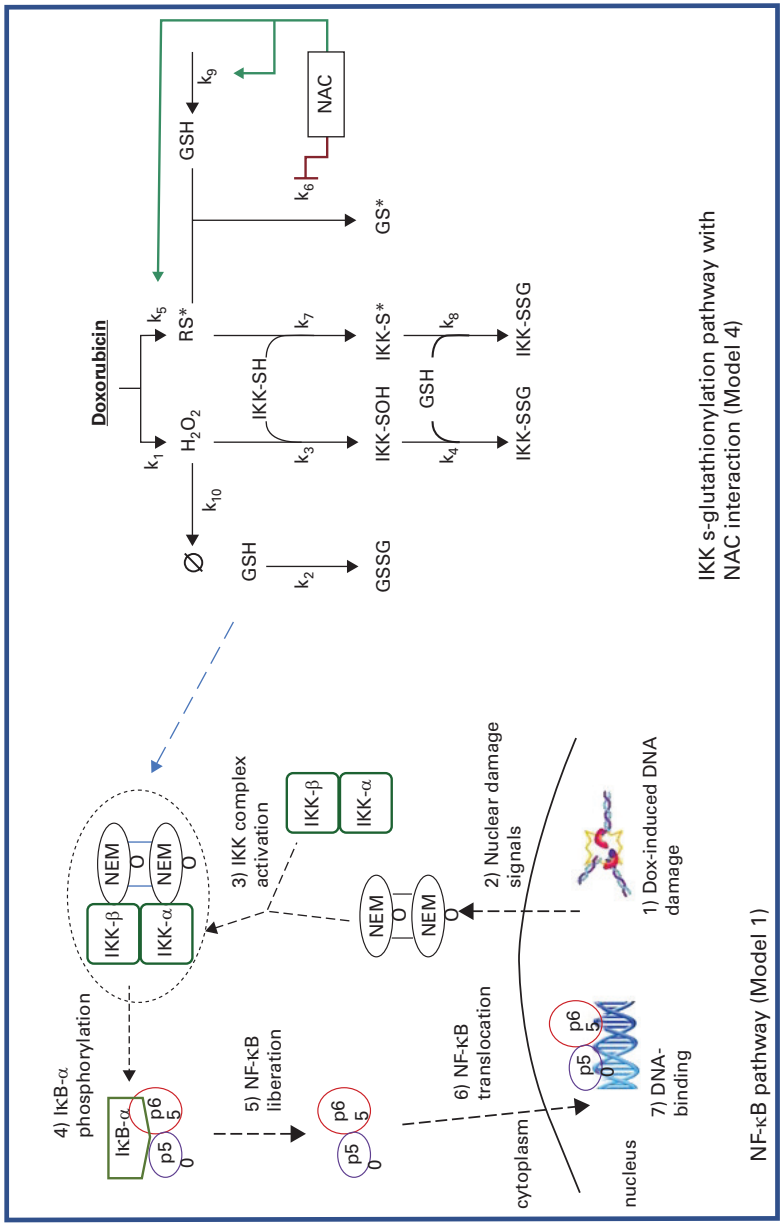


Figure 6.3
The left pathway is the one C9 built of NF-κB pathway-Y (Model 1). The right pathway is for the IKK s-glutathionylation system with the action of NAC built in, where IKK-β is the active component of the IKK complex, circled on the left pathway. On the right pathway, the values of fixed parameters k_5, k_6, and k_9 were used to help reveal the action of NAC. The values of IKK predicted by the right pathway in response to Dox and antioxidation could be used to modulate IKK in the NF-κB pathway.

possible under the treatment conditions she employed in the wet-lab experiments, to narrow the ten down to three possible mechanisms. She used simulations of different concentrations of NAC to build the models of these candidate mechanisms and to fit them to the experimental conditions in order to choose the best of them. The fitting process revealed certain modified parameters that suggested NAC has both pro- and antioxidant effects (see figure 6.3, right pathway). Through more literature searches, simulations, and parameter fitting, she was able to postulate a mechanism by which NAC operates on the glutathionylation process. Her model reproduced the nonlinear response of the IKK-β protein. Above certain concentrations it has pro-oxidant effects. In clinical terms, her findings indicate that it might be counterproductive to use NAC to treat cancer in certain instances.

When C9 began what she thought would be her final examination of the NF-κB pathway, she had not intended to build a model to describe NAC effects, but found herself compelled to in her attempt to explain the counterintuitive effects she saw in her experiments. She separated the use of a model in this instance from other uses she had made in the first two phases, where she had needed the model at the outset of her investigation to develop a basic understanding of the phenomena. Here, she had presumed she could do without it and rely on a black-boxed relation (an "*arrow*") for the input of antioxidants into the system, but had found herself short: "*It was doing an experiment, seeing that it was crazy, and having to build a model. . . . With the other two I started with the model and used the model to try and inform experiment. And in this, I did the experiment not expecting to build a model, but had to build the model to explain what I saw experimentally. So that's the relationship there.*"

She found this final model-building process quite difficult, and it was the one time she expressed the wish that she had been able to work with the director of lab G to build a better model. But time was short to complete her dissertation within the allotted time. The model applies a relatively straightforward set of functions to the relationship between NAC and the rates ('k's in Model 4, figure 6.3) of various protein syntheses in the ROS network. The process of building Model 4 was tightly reasoned on the basis of chemical reactions she found in the literature. The model was able to capture their nonlinear interaction and to replicate the nonlinear effects observed in her wet-lab experiments. In all, although not an exhaustive representation, C9 thought "*it's a good enough approximation to get things*

done." As she explained what she felt she had achieved in this part of her research, "*Since the experiments were good and done, I think we accomplished what we needed there, but . . . [in] the generation of the model—I wouldn't necessarily call it a full-blown model, it was more of a mathematical framework to try and understand what was going on*." This framework could provide a basis for another graduate student or postdoc to work from. She also felt, "*in hindsight*," that if she had collaborated more with the lab G director, she could have gotten more "*feedback from him to tweak the model and make it a little more robust and applicable. . . . But I feel I've done enough in my doctoral career*." So, although it might have been possible to negotiate more time to her degree, given its interdisciplinary nature, she also felt she had spent enough time as a student.

This work proved to be the final research of the dissertation, and the focus of her third paper. However, as she discussed in her final post-defense interview, the primary aim of discovering how manipulation of NF-κB affects cell viability through ROS signaling remained an open and unresolved problem. When looked at as a whole, she saw the result of her PhD research as reinforcing the idea that "*redox mechanisms do play a role in chemotherapy administration and more attention should be paid to those mechanisms*." She characterized her research as having opened up the discussion without having come to precise conclusions that could be used clinically. In acknowledging the limitations of their research into Dox and NF-κB, she noted, "*I would say [it is] even more far removed from a clinical setting because we only looked at one anti-oxidant, a particular range, and we didn't even look at the effect of altering NF-kB signal transduction on viability. All we said is this does alter NF-kB. There is a huge question that's left to be answered that is: once you have found out that it alters NF-kB what does that mean for viability?*"

In the end, her research never managed to make the final connection between the NF-κB pathway and the insensitivity of EU1 cells. Overall, her findings definitely were novel and important, but as with most PhD research, time constraints, and the feeling of having "*done enough*," at this point shared by the student and the committee, in the end led C9 to not continue on to the "*viability*" question. She settled (or satisficed) for having shown that Dox did interact with certain parts of the system without having determined the mechanism by which it produces its ultimate effects. She had made substantial progress, though, and based on her research, a future member of the lab, or someone else in the field at large, could pick

up the research from where she left it. The complexity of the biological system was greater than C9 and the lab director had anticipated. As the director would say, "*we started out doing this project, and realized we kind of had to back up a few steps*."

6.4 Epistemic Affordances of Coupling Simulation and Experimentation

The foregoing analysis of C9's investigative pathway tells the tale of an intensive research process that ranged over different systems and took turns and detours on the way, meeting unexpected obstacles while making valuable discoveries. Our reconstruction and analysis of her model-building process offers important insights into the ways in which the bimodal strategy can provide researchers in ISB another means to manage the complexity of modeling biological systems through creating a kind of distributed cognitive-cultural system different from the typical one, which requires collaboration between modelers and experimentalists. C9 built a D-cog system that, through their intensive interaction, coupled experimentation, computational modeling, and mental modeling in her model-based reasoning processes. The epistemic affordances of this system enabled her to undertake the challenge of building models of a system for which there was little understanding, and which required new experimental information and experimental testing to succeed.[9] As a result of being able to collect the data she needed as she was building the model, she was able to find most of the parameters required to fit it and to test its simulation outputs, so there were considerably fewer arbitrary features in the models.

The system-nature of her research practices also informed and shaped the ways in which she was able to make progress on her epistemic project. Chief among the affordances of the bimodal strategy is the ability to do her own experiments when and how she determined, which facilitated, significantly, her ability to fit and test her models. As we saw, she was able to design experiments to target the specific data she needed. As she described her process, "*I like the idea that as I'm building my model things are popping up in my head: 'oh wow this would be a good experiment'. I plan out the experiment myself and then go into the lab and I do it*." Further, those processes were much more efficient when compared with modelers who have to rely on experimental collaboration alone—often significantly delayed, as we saw in lab G. For one thing, the bimodal strategy solved the considerable

problems pure modelers face when trying to convey their experimental needs to pure experimentalists: "*I personally think [my approach] is better only because—I could tell someone what I needed, but—I think, not really understanding the modeling aspect, they can't accurately come up with a good enough experimental protocol to get what it is I need*." Importantly, doing both modeling and experimentation enables the researcher to "*understand the links between them*."

In collaborative relationships, experimenters and modelers interact with respect to a model, but as we have witnessed, and been told numerous times, there are typically many inefficiencies involved in such relationships. On the one hand, experimentalists often do not understand what kinds of experiments and data modelers require, because they do not understand model-building. On the other, modelers do not understand the nature of wet-lab experimentation sufficiently to know how to frame a request appropriate to the affordances and constraints of experimentation or, more precisely, those of their specific collaborators. Experimental collaborators usually conduct experiments to test hypotheses derived from the models only after the modelers are finished, not during the building process. A failed collaboration within lab C of C7 and C11 bears out this point. C11 tried to conduct the experiments C7 needed in the course of building his model, but the lab director ended up having to take over the experimental work to get at what the modeler needed. Only much later, while she was taking the biosystems modeling course, did C11 realize what was needed by C7. As she formulated it, "*You know, I wish I had taken this class two years ago . . . and it would have been very helpful for me to understand what kind of data he needed, to understand what kind of questions he should be asking of me. . . . [I] didn't have an insight to what exactly—what kind of data would be useful for him*." She also noted that C7 had begun conducting his own experiments, so he had "*figured out the same thing from his end. It's easier if you have more knowledge on the other side*." Chapter 7 discuss insights from our experiences in these labs on ways to facilitate collaboration in bioengineering sciences and other interdisciplinary practices.

C9, of course, was not confined by such problems. As a result, she was able to coordinate her modeling and experimental activities efficiently and, most importantly, to make sure that her experimentation was well-adapted to obtain the information she required, within the constraints of her experimental abilities. This coordination gave her the ability to run

experimentation and modeling (computational and mental) as an effective coupled system. Here, as with the coupling between model and modeler I discussed in chapter 5, the overall effect of the back-and-forth exchange between these components of the D-cog system is to extend human inferential powers so as to facilitate the modeler in building out the pathway while building out the model and improving its parameter fit, as well as to provide model validation. Such a coupled system has all the epistemic affordances of a modeler-model coupled system we discussed in chapter 5. However, adding wet-lab experimentation to the coupling provides significant new epistemic affordances in that it enables the modeler to discover and extract relevant and needed experimental information from a complex jumble of biochemical interactions, as well as to control and direct the information flow in the course of building models.

Further, wet-lab biological experimentation, in general, has important epistemic affordances that "google biology" lacks. Experimental engagement with target systems provides another, different kind of epistemic access to the target system from that possessed by those who do only modeling. For one thing, in developing the ability to design and run biological experiments, C9 was able to develop her own sense of what is "reasonable" biologically, which we saw was a constant question for lab G modelers. Her embodied engagement with the biological entities informed her mental models and, undoubtedly, gave her an understanding different from unimodal modelers, for whom these are abstract entities, known only through literature, pathway logic, and simulation.[10] C9, too, talked about getting a "feeling" for the behavior of her models, and, as with the lab G modelers, the system properties and dynamic behaviors of the biological system were only available to her through the model simulations. However, her "feeling for the system" was not based solely on the model and literature, but also on her engagement with the phenomena as she designed and conducted experiments. As we saw with lab G, aspects of developing a "feeling"—for the model or system—in the context of research include a growing insight into and understanding of the target system, belief in the credibility of the model, and affective engagement in the research.

6.4.1 "I Did Them at the Same Time": The Coupled System

To discern the epistemic affordances of this kind of coupled system, I start from an overview of C9's modeling practices. As with other model-based

reasoning, once she built an initial computational model, it became the main cognitive apparatus through which C9 interpreted the experimental data and advanced her research by her choices as to what to investigate, experiment on, and model further. She relied on the behavior of the model itself to inform her on the properties and functions of its parts, particularly to reveal the existence of missing parts. The model provides information only available at the system level, such as the relationships between indirectly related variables. As she ran simulations and produced visualizations of the resulting variable and parameter relations, each model provided information otherwise inaccessible to wet-lab experimentation. For instance, the model provided measures not possible to determine in experiment, and simulations helped to reveal which pathways bear most of the burden in a system or which points are relevant to control a network. Determining network control points enabled C9 to black-box mechanisms that did not appear to affect control, as well as to target her experiments. In comparing experimenting in building the model (simulation) and in the wet lab, she pointed out that she could "*tease apart*" things in experiments with the model that she could not in the wet lab and claimed, "*You have access to everything in the model . . . in that sense, I can actually see deeper into the experiment in my model than I can in the real experiment.*" However, she had more "*trust*" in what she could "*see*" in the wet-lab experiment, which "*obviously is showing you what actually happens.*" Of course, what she could "*see*" there is what "*happens*" in the context of an experimental—in vitro—situation.

As we saw, in many instances C9 conducted experiments while in the process of building and running simulations with her models. She called her research process "*iteration*" between them. When we asked which was driving her results—models or experiments—she replied that she was doing them nearly simultaneously: "*You can say that I did my model first, but I don't . . . I don't see it as I finished my model and then I did my experiments. I see it as kind of like I did them at the same time.*" She referred to their interaction as "*synergistic.*" We detailed the aspects of this coupled interaction by looking at the reciprocal roles model-building played in directing her experimentation, and conversely the role experimentation played in her model-building. Figure 6.4 provides our schematic of the way her model-building and experiment interacted.

C9's usual strategy was, in the first place, to construct a diagram of the topology of the biological pathway based on her extensive review of

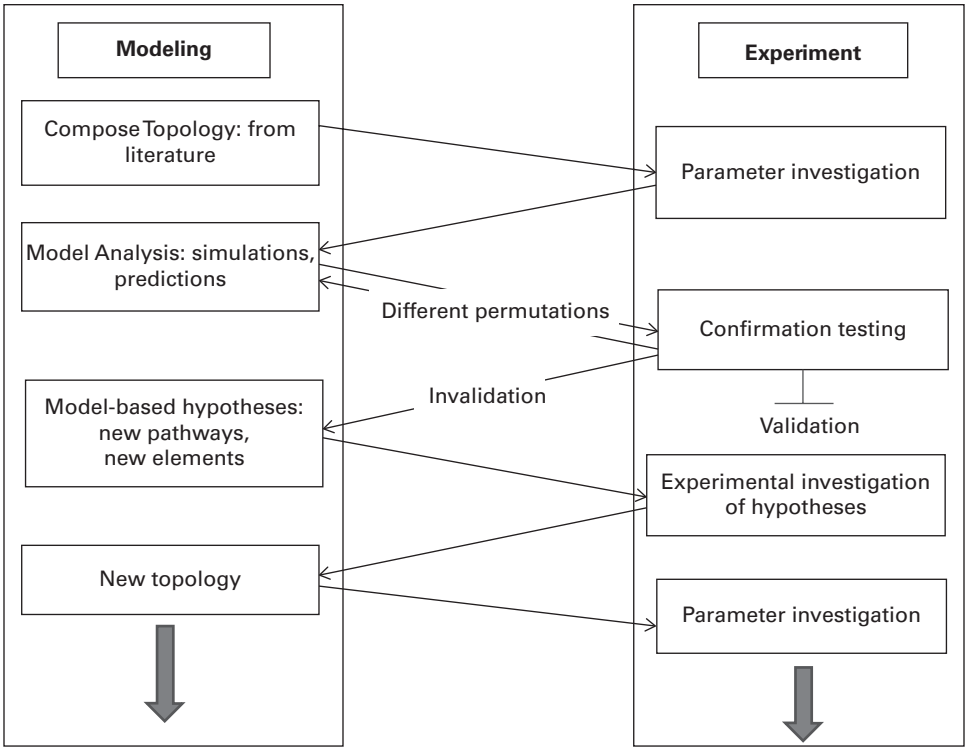


Figure 6.4
The diagram illustrates the coupling of modeling and experimental procedures in the course of C9's model-building as an iterative interactive process of discovering the correct topology and parameter values for the networks on which she was conducting research. The process ends when validation is achieved. Experimental results that depart from the model simulations ("invalidation"), require the generation of new hypotheses, which are derived from the existing models or their modification/elaboration.

the literature. This process is much the same as what we saw in lab G, that is, she would trace and piece together parts of the network from different sources in the literature and add or alter components on the basis of simulations of the preliminary model. In our interviews on how she was building the pathways, she describes an iterative and incremental interactive process that contributes to building her model-based understanding of the system. She noted the specific value of the visual nature of these pathway representations, that are themselves "*models*" that are "*just visual now. They*

don't have any math affiliated with them. I have to visually see what's happening before I can start writing the equations." As I have discussed, the pathway representation provides both a conceptual and a visual model of a causal chain of interactions within the biological system that lead to the observed data. C9 further explained her process of "*visually seeing*" to us as, "*I have a notepad where I do a lot of my calculations, thinking, sketching, different things. I have notes and sometimes, like right now with the NK-κB model, I have to come up with some of the connections. I have to hypothesize for some of those because it's not known in the literature. So, I'm going back and forth drawing different things*." These processes of diagramming, sketching, and calculating on paper while thinking and making inferences contribute to coupling her mental models with the pathway model and the computational model.

As with lab G modelers, to build the topology, she also needed to collect as much parameter information from the literature as she could. However, after she had gathered and pieced together whatever information she could find through those searches, there was an important difference: her next move was to rely on her own wet-lab experimentation to obtain pathway and parameter information that was not available in the literature. Once she put together this information in the form of an ODE computational model, she could begin the process of running preliminary simulations to get a feel for the model's behavioral dynamics and to build out the pathway, as well as to produce predictions that she could investigate and possibly verify experimentally. The bimodal strategy enabled her to run simulations and experiments nearly simultaneously, as we saw in the drug Dox (Model 3) case. She could try to replicate interventions and perturbations she made in the models in the physical systems as, for instance, she did with the aid of known blocking or suppressing agents.

The systematic and organized manner in which she coupled experimentation and model-building had a range of epistemic affordances. A prime affordance was the ability to evaluate, experimentally, parts of the model as she was building it. When simulations did not replicate the experimental data, she could rely on the elements of the model she was relatively confident about and could probe those parts she was not. In this way, her models and their limitations provided resources from which she could make inferences as to how to direct her wet-lab experimental activity, as she tried to track down solutions in the form of new pathways and new functional metabolic elements that could fill out the structure of the models. The salience

of these elements would not have been apparent on the basis of a program of experimental investigation alone, but required the system-level perspective provided by the model to be identified. For instance, when building Model 1, she discovered the role of cysteine by virtue of the model: "*I sort of reached that IKK . . . needed to be disulfide bonded in order to be active. I reached that through a conjecture because I was like 'something is missing.. When I was drawing my pictures, I was like 'okay, there is a big gap here. How is it that IKK has to be oxidized to be active?' And I was like 'There must be a redox sensitive cysteine.'*" She noted at another time that, "*our model can help us tease out that information . . . like when the model helped us tease out that thioredoxin dependent proteins are basically reacting more speedily with hydrogen peroxide than like glutathione dependent proteins. That's something we couldn't have gotten at without the mathematics*."

In other instances, when things turned out unexpectedly with her experiments, she could build a model to figure out the problem. When, for example, she discovered that her experiments provided results that were in opposition to her model-based (Model 1) expectations that NF-kB levels would be up-regulated in EU1 cells and down-regulated in the EU3 cells, she built a new Model 3 of the production of ROS by Dox. As we showed in the case analysis, that model informed which questions to ask and experiments to run, and which "*arrow*" in the pathway to pursue as the avenue to best resolve the discrepancy between her models and the data. She again needed to build a model (Model 4) when she discovered, contrary to her expectations, the non-monotonic behavior of IKK-β with respect to NAC levels. She talked repeatedly about how she used the models as explanatory tools, saying, for instance, in this case, "*We have to go to the model to explain what we are seeing because we can't do it experimentally . . . and going to the model and kind of pairing it with the literature evidence, we were able to come up with an explanation that we believe sort of encompasses the changes we saw. . . . I couldn't explain what I saw without the model. It just wasn't possible for me to do that*."

The repeated coordinated interaction of experimentation and simulation created a coupling between these processes and her mental models. Through this coupled system, C9 could construct the topology of these complex systems by going back and forth to gather data, measure parameters, and make new hypotheses. Her simulations produced novel behavior that would tell her to "*pay more attention to this [novel model behavior]*" on the "*experimental side*." Experimentation provided a source of ongoing

epistemic validation for the model and reinforced its role as a platform for hypotheses. It also informed the mental models through which she reasoned about how to improve her computational models, what specific experiments to undertake, and the dynamics of the phenomena. Further, as Wimsatt has perceptively and correctly noted, "*processes of validation often shade into processes of discovery*—since both involve a winnowing of the generalizable and the reliable from the specific and artifactual" (Wimsatt 2007, 56; emphasis added). The discovery dimension of coupling is highly significant in C9's case. By running controlled experiments and simulations side by side, C9 was able to discover reliable pathway structures, novel biochemical mechanisms, and parameter values with precision, as well as to build robust, validated models.

Through coupling experimentation and modeling, C9 was able to manage the complexity of the problem-solving tasks she faced, because the strategy provides an efficient and effective means by which to explore and constrain the large possibility space of her tasks. As we saw in chapter 5, ISB modelers, because of limited data or data of the wrong kind, face large possibility spaces that are difficult to search through because the systems are nonlinear, and so it is much more difficult to determine what pathway structure or parameter hypotheses might be in the neighborhood of a solution. They often have to test a large number of alternatives. The bimodal strategy, however, provides advantages for narrowing down the search through a parameter space. For example, C9 was able to diagnose and localize uncertainties in her models and to discover relevant sources of information to improve them by simulating and comparing parts of them against controlled experiments. Such localization afforded her an ability to posit and to test, experimentally, tractable hypotheses about uncertain mechanisms in order to extract relevant information she lacked.[11]

In particular, C9 used perturbations of the model and experimental perturbations of the biological system to derive strengths and weaknesses of the models she had built from the available literature. She could isolate parts of the models, simulate those interactions, and check them against experiments that isolated these relations physically in order to establish the accuracy of those parts or to measure their parameter values. When, for instance, her experiments showed differences in NF-κB activation for EU1 and EU3 from what she had expected from her model and the literature, she was able to trace the problem to "*a single arrow*" in the pathway

through a combination of wet-lab experimentation and model simulation. She could also use her models to identify sources of biological information necessary to improve the model. In general, as the result of a coordinated probing of the model by means of perturbation and controlled simulations and of probing of the physical biological systems with targeted wet-lab experimentation, C9 was able to establish, robustly, the parts of the model that were accurate, as well as localize the problematic parts or parts that were relevant to further development of the model. Once localized, she could conduct experiments to run through various model-based hypotheses about the interactions of those components, as she did when she formulated a mechanism for the IKK-β s-glutathione system and the antioxidant NAC. Her further wet-lab perturbations to the biological system, such as by changing chemical inputs and using other chemicals to suppress specific reactions, could be checked against the dynamics predicted by the model and used to fix pertinent model constraints. This gave her, in particular, the ability to fine-tune the specific parameters of the reactions and interactions she was experimenting on without having to do the kind of large-scale algorithmic fitting we saw in lab G. For instance, rather than having to model each of the ten possible mechanisms for NAC and trying to infer a best fit, C9 used controlled experimentation of their mechanisms to narrow the model-building task down to three.

This case analysis provides an exemplar of how the bimodal strategy creates a coupled problem-solving system that serves to manage the complexity of model-building in ISB. C9 relied on information embodied in her models, derived from experiments and culled from the literature, to triangulate the location of inaccuracies and missing elements. This bimodal strategy enabled her to build confidence in the parts of the models as she built them through wet-lab testing and perturbation, which allowed her to localize problems and locate and reduce uncertainties. She was able to narrow down avenues for solving model-building problems to a limited set of possibilities, for which she could make hypotheses and test them through controlled experimentation. The bimodal strategy, in particular, enabled her to avoid the challenge faced by unimodal modelers that, given limited experimental data, they have to sort through pathway structure alternatives and parameter values algorithmically, knowing these often will need to be refit. However, this strategy does have limitations, which include significant ones with respect to what and how biological systems can be modeled.

These limitations underlie what the lab G director called a "*philosophical divide*" about how best to build models, as well as train modelers, in ISB.

6.4.2 Limitations of the Bimodal Strategy for ISB

Although some kind of interaction between experimentation and simulation is required to model complex biological systems, the three-way coupling (modeler–model–experiment) of the bimodal strategy is atypical in the current landscape of ISB practice. For the most part, interaction does not extend beyond that of an experimentalist supplying initial data and then data to verify or falsify the model's predictions, as we saw with the G12 case. Such a "collaborative" strategy is often fraught with difficulties, as we have witnessed and heard described by lab G researchers. In some instances, it can even lead to the abandonment of a line of research. It is not easy for a modeler-experimentalist team to overcome a range of barriers created by specialization and create an efficient and effective problem-solving system. The bimodal modeler does have an advantage in this respect. However, one should not take away from the C9 case that a bimodal strategy is necessarily the best, or most effective, epistemic route to building models ISB. There are trade-offs associated with each approach. Further, as we saw, the bimodal approach does not mitigate, completely, the epistemic interdependence I noted to be inherent in ISB. First, modelers using the method still rely on the experimental literature to build and validate models, and, second, since they have not been trained as biologists, they still need to rely on the expertise of bioscience advisers or of collaborators who might be on their project, even though they conduct all the experiments themselves. They are, of course, better equipped to take advantage of the expertise and to collaborate than dedicated modelers. All the experimentalists we interviewed in connection with research in lab C or Lab G were aware of the practices of both labs and uniformly expressed their opinion that the bimodal strategy was the better approach, since it was likely to enhance collaboration between the fields, and expressed the hope for this "*new breed of students*" to become dominant in ISB.

Importantly among the trade-offs, there are limits to the scale of the system a modeler can manage reliably by a bimodal strategy. The strategy worked well in C9's case because she was dealing with relatively small-scale systems, with a manageable number of unknowns. This fact enabled her to contain her experimental work and direct her model-building to keep them

within reasonable constraints. Because she had sufficient data to develop detailed models of the biochemical interactions, the need to use mathematical averaging techniques and fitting algorithms was reduced, along with the potential for error these bring with them. So, she did not require the level of mathematical and computational sophistication we saw with lab G to build useful models of her system. Even so, she did not get as far on her problem as she had initially thought she would, and as would be needed to solve the clinical problem. She, herself, felt she could have benefited from more sophisticated modeling skills. By contrast, as we have seen, a dedicated modeler can handle a significantly larger system either because there are sufficient data available or experimentalists with whom to collaborate or because they possess a high level of mathematical and computational skills that they can use to try to get around the lack of data. As we have seen, the rhetoric surrounding ISB promotes its promise to build models of large-scale systems, and so favors a unimodal approach.

Another way to get around the lack of experimental data for large-scale model-building, and to provide it in a timely manner approaching the efficiency of bimodal strategy, might be to have a lab organization that comprises both dedicated modelers and dedicated experimentalists. This was the kind of lab in which the lab C director did her postdoctoral research. The lab G director felt this strategy would require supporting a huge lab financially, which is not usually practicable. He stated that the ideal ratio is, as he has "*said it for twenty years, you need ten experimentalists for every one modeler*." This is because experimentation and modeling work on different, asynchronous, time scales. Wet-lab experiments needed for specific data to build the model or to check the outcome of a model simulation can take many months. We often saw lab G modelers waiting around for data from collaborators to continue their modeling work, while in the meantime they worked on algorithm development. On the flip side, building a robust, stable model takes a long time, too, and, as experimentalists told us, they had often "*moved on to working on something else*" by the time the modeler gets back to them with hypotheses to test, and they are unwilling to redirect their time and money.

As to modelers doing their own experiments, the lab G director expressed worries about the risk of experimental bias or the unwitting manipulation or interpretation of results to fit the model: "*If you produce the data to validate your model, implicitly or explicitly . . . there is a lot of room for interpretation*." He felt the best way to avoid this risk was to have someone else do

the experimental validation research, particularly at some arm's length (in different labs). On the other hand, his concerns contrast with what the lab C bimodal modelers told us. For instance, C7, who had considerable experience building models from data collected by others before he started to conduct experiments, claimed, "*I feel like I have more confidence in the data if I'm doing it on my own part rather than having someone else do it.*" Of course, confidence does not mitigate the possibility of bias. Another plus for bimodality, though, is as C9 told us and as was confirmed by experimentalists, it is often difficult for a modeler to convey—or for an experimentalist to determine—what experiments will satisfy their needs.

Perhaps most importantly, for the question as to whether the bimodal approach will become more common, is that to become skilled in both experimentation and model-building takes significant time and can involve compromise in the level of skill developed, unless, perhaps, one opts for intensive sequential training, either by taking a longer route to the PhD or in postdoctoral research. The lab C director, however, felt she had "*lost time*" by training sequentially, but clearly, she was more accomplished modeler and experimentalist than the students in her lab would become in the customary five-year PhD research period. The lab G director thought the sequential approach to training was best if one wanted to adopt the strategy. Backing up the lab G director, the postdoc who was then training in his lab to be a modeler after getting an experimental PhD told us he would not organize his own lab to have graduate students use the bimodal strategy: "*What I get from my personal experience is that I lose time going from one side to the other. . . . Both fields have their own problems, you know. . . . A student needs to have his brain focused—and performing experiments is an art, just like modeling.*" Modelers in ISB tend to come from applied mathematical or engineering backgrounds and have little if any knowledge of experimentation. But they also need to be trained in the "*art*" of systems biology model-building, which requires that they not only learn how to apply and further develop their mathematical and computational skills to model biological systems, but also develop new skills for how to search for the data in literature and databases (conduct "google biology"), as well as for how to develop a basic grasp of the nature of the system. It is no easy task to learn these skills, especially given that today one might be modeling a yeast system and tomorrow a cancer-producing system. And, we would add from our investigations, modelers need to invest in developing

collaboration skills much more than is currently done. Researchers who adopt the bimodal strategy need to develop all of these skills, but in addition they need to learn experimental design and benchtop skills.

As noted, with respect to bimodal training, the "*philosophical divide*" extends to whether to learn experimental and modeling practices concurrently or sequentially and, in the latter instance, whether to do so in the context of a PhD, or first do a modeling PhD and then learn experimentation as a postdoc or vice versa. Concurrent training risks the consequence, as the lab G director noted, of ending with "*modeling lite and experimenting lite*," but sequential bimodal training puts the researcher in the position, as the lab C director noted, of basically "*starting over*" after several years of education. As we saw, C9 did express some regret, in our final, post-PhD defense interview, that she had not been able to develop her modeling skills more, especially that she had not taken sufficient advantage of working with the lab G director. She felt that having those more sophisticated skills might, in particular, have helped her to work better with the complex nonlinearity of the system in the NAC case. But she also said she would not have given up the ability to couple experimentation with her model-building to rely on collaboration instead.

6.5 Summary: Getting a Grip with/on Bimodal Model-Building

The bimodal strategy can be understood as a particular response to features of the problem-solving contexts of systems biology: namely its complex problems, lack of theoretical starting points, lack of data, and other constraints. With the bimodal strategy, experimentation and simulation are closely coupled in the model-building process, not only to validate model-building steps, but also to provide an effective means of limiting search spaces and triangulating on a good representation. The bimodal strategy has both advantages and limitations when compared to the unimodal strategies we have studied. It does offer unique epistemic affordances.

In the central phases of her research, where C9 needed to use the bimodal strategy, she gave much more weight epistemically to wet-lab experimentation than to simulation. The uncertainty over the structure and properties of her systems prevented her from building a model by starting from accepted assumptions and filling in numerical details of an established modeling framework. She was never in a position where she could rely

on a model to conduct only simulation experiments on her systems. The limitations of the models, on the other hand, proved to be resources that directed her experimental activity to track down solutions in the form of new pathways elements and new functional metabolic elements that could fill out the structure of the models. The interaction of experimentation and simulation provided continual epistemic validation for the model-building steps along the way and reinforcement of the model's role as a platform from which to both build understanding of network dynamics and make hypotheses about the behavior of the biological target system.

As we have seen throughout this book, methodological innovations enable scientists to create or enhance their native cognitive powers by building distributed cognitive-cultural systems appropriate to their complex problem-solving tasks. Most often, philosophers and cognitive scientists, when they address cognition and complexity in science, have focused on heuristics that help narrow down search spaces for complex problems (e.g., Newell and Simon 1972; Wimsatt 2007). Cognition is interpreted in terms of its constraints, which heuristics help overcome. We have seen, for instance in the case of mesoscopic modeling, ISB model-building methods can, indeed, be developed with cognitive limitations in mind. The bimodal strategy, however, is not just a means to employ heuristics to deal with cognitive limitations. The bimodal strategy enables the modeler to actively distribute her cognition and build her cognitive powers through coupling simulation and experiment. In the C9 case, running simulations provided a method by which she could calculate and visualize dominant network patterns, which, in turn, helped her develop and simulate mental models of the system dynamics. She was able to infer from the understanding obtained through these simulations how to direct experimentation, manipulate the biological materials, and interpret the results, thus turning experimentation into a sharper investigative tool that could help her efficiently and intelligently search through the space of network and parameter possibilities. The whole problem-solving system served to augment her cognitive capacities to investigate the complex biological systems.

Direct experimental engagement with the biological systems also served to reduce the risk of error. In particular, she developed sophisticated strategies to localize inaccuracies in her model, and thus localize quite precisely where wet-lab experimentation was required. She would do this by running sets of controlled simulation experiments (fixing particular sets of

parameters and varying others) against controlled biological experiments that replicated the model controls physically. In this way she could work through the model to pick out specific problematic relationships and investigate them further. This fine-tuned operation gave her a specific ability to handle the complexity of her problem-solving task, not available to unimodal modelers. Unimodal modelers are usually forced to explore complex large parameter and structure spaces algorithmically or with Monte Carlo methods to discover approximately accurate structure and parameter values that fit the whole model to the data. C9 could cut down the space significantly with well-targeted experimentation on specific relationships. Once she localized errors to specific structural hypotheses or parameter values, she could conduct wet-lab experiments to run through sets of hypotheses about the interactions among those components and measure their parameter values. When it came to parameter-fixing, C9 declared that she had always at most two to three parameters to fix. This compares with the upward of twenty that most lab G members report.

In sum, the bimodal strategy tightly integrates the two modes—modeling and experimentation—into a system to generate and validate information about complex biological systems. This *coupled methodological system* provides a novel means to manage the complexity and uncertainty of biological systems, but its challenges and limitations make it not yet a widespread methodological choice in ISB. It is not evident whether bimodality will be the choice of a "*new breed*" going forward; however, one thing is clear from our investigations: modelers and experimentalists need to become more conversant with one another's epistemic practices, norms, and values than is the current situation.

Interdisciplinarity in the Making

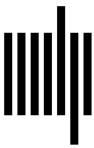
Models and Methods in Frontier Science

By: Nancy J. Nersessian

Citation:

Interdisciplinarity in the Making: Models and Methods in Frontier Science
By: Nancy J. Nersessian
DOI: 10.7551/mitpress/14667.001.0001
ISBN (electronic): 9780262372275
Publisher: The MIT Press
Published: 2022

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open



The MIT Press

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Nersessian, Nancy J., author.
Title: Interdisciplinarity in the making : models and methods in frontier science / Nancy J. Nersessian.
Description: Cambridge, Massachusetts : The MIT Press, [2022] | Includes bibliographical references and index.
Identifiers: LCCN 2021061880 (print) | LCCN 2021061881 (ebook) | ISBN 9780262544665 | ISBN 9780262372268 (epub) | ISBN 9780262372275 (pdf)
Subjects: LCSH: Biotechnology—Methodology. | Bioengineering—Methodology. | Biotechnology—Research—Case studies. | Bioengineering—Research—Case studies. | Biotechnology laboratories. | Scientific surveys. | Interdisciplinary research.
Classification: LCC TP248.24 .N47 2022 (print) | LCC TP248.24 (ebook) | DDC 660.6—dc23/eng/20220720
LC record available at https://lccn.loc.gov/2021061880
LC ebook record available at https://lccn.loc.gov/2021061881