

An Introduction to Simulation and Visualization of Biological Systems at Multiple Scales: A Summer Training Program for Interdisciplinary Research

Rajan Munshi^{1*}, Rob D. Coalson², G. Bard Ermentrout³, Jeffrey D. Madura⁴,
Hagai Meirovitch¹, Joel R. Stiles⁵, and Ivet Bahar^{1*}

¹*Department of Computational Biology, University of Pittsburgh, Pittsburgh, PA 15261*

²*Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260*

³*Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260*

⁴*Department of Chemistry and Biochemistry, Duquesne University, Pittsburgh, PA 15282*

⁵*Mellon College of Science & Pittsburgh Supercomputing Center, Carnegie Mellon University, Pittsburgh, PA 15213*

*To whom correspondence should be addressed

Tel: 412-648-7799

Fax: 412-648-3163

E-mail: rajan@pitt.edu

1. ABSTRACT

Advances in biomedical research has necessitated the need to train a new generation of researchers having both a strong background in the life and physical sciences, and a knowledge of computational, mathematical, and engineering tools for tackling biological problems. The NIH-NSF Bioengineering and Bioinformatics Summer Institute at the University of Pittsburgh (BBSI @ Pitt; www.ccbb.pitt.edu/bbsi), is a multi-institutional ten-week summer program between the University of Pittsburgh, Duquesne University, the Pittsburgh Supercomputing Center, and Carnegie Mellon University, and one of nine Institutes throughout the nation currently offering this unique training opportunity. Each BBSI focuses on a different area; the BBSI @ Pitt, entitled “Simulation and Computer Visualization of Biological Systems at Multiple Scales”, focuses on computational and mathematical approaches to understanding the complex machinery of molecular-to-cellular systems at three levels, namely, molecular, subcellular (microphysiological), and cellular. We present here an overview of the BBSI @ Pitt, the objectives and focus of the program, and a description of the didactic training activities that distinguish it from other traditional summer research programs. Furthermore, we also report several challenges that have been identified in implementing such an interdisciplinary program that brings together students from diverse academic programs for a limited period of time. These challenges notwithstanding, presenting an integrative view of molecular-to-system analytical models has introduced these students to the field of computational biology, and has allowed them to make an informed decision regarding their future career prospects.

2. INTRODUCTION

There have been fundamental changes in biological and biomedical research methodology and topics in recent years due to spectacular advances in genomics/proteomics as well as in computational power and technology. With the wealth of information now being accumulated in biological databases, there is a critical need for novel theories and advanced tools to decipher the implications of these data, and devise methods of controlling/modifying biological function. New initiatives are now taking shape in the form of structural genomics, functional genomics or proteomics. There is now a shift in emphasis – from sequence to structure, from genes to proteins and their complexes, from interacting pairs to interaction networks. There is also a change in the scale of the explored processes, from atomic/molecular to supramolecular, cellular and systems levels. This field, broadly (and interchangeably) referred to as “computational biology” or “bioinformatics”, encompasses a wide range of topics, from molecular modeling and protein dynamics to large-scale analysis of genome/proteome data. To successfully train a new generation of interdisciplinary investigators in this field, individuals who will develop and implement theories, methods and tools must be taught comprehensively. For potential researchers in the field, it is essential to acquire the quantitative background and skills to advance the field, and be in a position to appreciate the potential, strength and limitations of computational, mathematical, and engineering tools for tackling biological problems (*1*).

Due to this emergence of interdisciplinary research, boundaries between traditional fields are gradually disappearing, especially at the level of graduate education. This is

evidenced by the large number of multidepartmental graduate programs in biomedical science that have been established over the past several years. However, similar undergraduate programs are quite rare, even though the need to introduce interdisciplinary curricula – particularly in undergraduate biology programs – has been well documented and acknowledged (2, 3).

Undergraduate research, such as summer Research Experience for Undergraduates (REU) programs, offers students the opportunity to prepare and pursue careers in science (4). In addition to improving and strengthening coursework fundamentals (5, 6), students who participate in these programs demonstrate a greater interest in careers in science, engineering, and mathematics (7). These programs have been shown to expand a student's educational potential and pave the way for students wishing to pursue graduate education (8), and to develop better research and communications skills (9). Furthermore, undergraduate research also contributes to retention of students from underrepresented minority groups and increases the rate of graduate education in this pool of students (10).

As important as traditional REU programs are, they have two limitations: (i) Student participants begin a research project with minimal preparatory training. Thus, knowledge acquired during the program is often “project-specific”. (ii) For the most part, only those students who already possess an appropriate background can join and successfully complete a research project. In other words, chemistry REU programs typically attract students majoring in chemistry; similarly, mathematics majors are generally drawn to

mathematics REU programs. These issues notwithstanding, REU programs are a valuable educational tool for students and researchers alike.

Computational biology and bioinformatics are relatively new fields, with degree-granting programs being offered only recently. Due to the inherent nature of these new disciplines, they draw heavily upon the physical and biological sciences, and also on mathematics, engineering, and computer science. A summer research training program in these areas requires that the interdisciplinary nature of these fields be addressed. Furthermore, students wishing to conduct research in computational biology and bioinformatics would need to be taught the fundamental principles of the fields prior to being able to attempt any research activity.

We report here a new interdisciplinary summer research program for undergraduate and graduate students, which attempts to address both of the above-mentioned issues. This program, the Bioengineering and Bioinformatics Summer Institutes (BBSI), is an NIH and NSF-funded initiative designed to train a new generation of multidisciplinary investigators in the fields of computational biology, bioengineering, and bioinformatics. Currently, there are nine such NIH-NSF BBSI programs in the nation, each with a specific research concentration (Table 1). The primary focus of the BBSI described herein is computational biology. Entitled ‘Simulation and Computer Visualization of Biological Systems at Multiple Scales’, it is a joint program offered by the University of Pittsburgh (lead institution), the Pittsburgh Supercomputing Center, Duquesne University, and Carnegie Mellon University.

Table 1: The NIH-NSF Bioengineering and Bioinformatics Summer Institutes (BBSI)

| Institution(s) | BBSI focus |
|-----------------------------------------------------------------------------------------------------------------------------|--------------------------------------------|
| University of Pittsburgh (lead), Pittsburgh Supercomputing Center, Duquesne University, Carnegie Mellon University | Computational Biology |
| California State University at Los Angeles | Bioinformatics |
| The Pennsylvania State University | Biomaterials and Bionanotechnology |
| Clemson University | Biomaterials Science and Engineering |
| Harvard University - Massachusetts Institute of Technology | Biomedical Optics |
| University of Minnesota | Bioinformatics |
| New Jersey Institute of Technology | BioMEMS (micro-electro-mechanical systems) |
| Virginia Commonwealth University | Bioinformatics and Bioengineering |
| Iowa State University | Bioinformatics and Computational Biology |

3. PROGRAM DETAILS

3.1. Overview

The BBSI at the University of Pittsburgh (BBSI @ Pitt) is a ten-week interdisciplinary program organized into coursework, laboratory research, and student presentations. Students receive intensive didactic training in the form of comprehensive course work that covers the theoretical and computational aspects of computational biology as applied to investigating molecular-to-cellular systems dynamics. At the same time, students conduct research in a laboratory of choice for the duration of the program. A major goal of this program is to identify talented students, increase their awareness of the quantitative and computer science methods in life sciences, and encourage them to consider career opportunities in the field of biomedical computing by providing them

with an integrative team-based hands-on research experience at an early stage of their studies. Furthermore, the program strongly emphasizes professional development in the form of weekly seminars (which include research and career seminars, and also an ethics forum), career workshops, and student presentations.

3.2. Intellectual merit

The program focuses on current computational and mathematical approaches for deriving information on molecular and cellular systems dynamics, and thereby function, using biomolecular structure, biological pathways, and other biochemical and biomedical data, much of which is continually incorporated into databases. An increasing number of research groups are now engaged in coordinated experimental studies within the scope of structural and functional genomics and proteomics. In parallel to these efforts, there is a need for trained individuals who can effectively pursue the related computational and theoretical aspects of systems biology, cell biology, and neuroscience, to devise new theoretical models and approaches that increase our understanding of the molecular basis of observed physiological phenomena. In particular, dynamics is not as widely and clearly understood as structure or state. The BBSI @ Pitt gives students an overview of existing tools, and an introduction to the type of background and expertise needed for tackling current complex biological problems with computation and visualization methods; students become acquainted with the opportunities and challenges in computational biology, which can potentially stimulate their interest in future graduate studies and research in the field.

3.3. Objectives

The vision of our BBSI Program is to bring together students majoring in biological sciences, computer sciences, engineering, mathematics, and physical sciences and provide them with an integrated, multidisciplinary education and research experience. The program's goal is to increase the number of students considering careers in computational biology, bioengineering, and bioinformatics as well as to initiate the development of talented young scientists with the necessary skills to make an impact in these fields. It also aims at presenting to the students in an early stage of their studies an integrative view of molecular-to-system analytical models, and the fundamental physicochemical, statistical mechanical, and kinetic principles, which are needed for predictive theoretical and computational studies/research in the field of computational biology, bioengineering, and bioinformatics.

3.4. Focus

The dynamics of biological processes at multiple scales is the focus of the present institute. While computational and theoretical methods for analyzing the equilibrium behavior of biological molecules or systems have been well established, methods for exploring the dynamics have not been as comprehensive and available due to the complexity of biomolecular and microphysiological interactions. An expertise is beginning to build in the fields of computational molecular and cellular biology as well as in mathematical (systems) biology, which is completely unknown to most students at the undergraduate level, and may be advantageously utilized in the broader community of future biological/medical researchers. This program intends (i) to provide the students with an overview of the existing models, methods and tools for exploring the dynamics of

biological processes, (ii) to broaden the vision and awareness of the students on current post-genomic computational or mathematical research problems in the field of molecular, cellular, and systems biology, and (iii) stimulate their interest in the newly evolving field of computational biology while increasing their awareness, and thus motivate them to further pursue career opportunities in the field. An important aim is also to enhance communication and interaction between the participating institutions and faculty, which is expected to have a synergistic effect on on-going research activities.

3.5. Targeted students

Consistent with the goals and desired impact of the program, student participants are selected from a diverse academic pool that includes basic life sciences or physical sciences, mathematics, computer science or engineering. Due to the comprehensive coursework designed for the program, students must have completed their sophomore or junior year and should have demonstrated a strong interest in cross-disciplinary studies through course work outside of their major field. While a majority of the participants are undergraduates, the program also provides an opportunity for graduate students entering their first or second year of graduate school to participate and further develop and define their proposed or ongoing research projects. Furthermore, special attention is given to attract students from underrepresented groups and from institutions where research programs are limited. In accordance with the guidelines of the NIH and NSF, undergraduate students receive a stipend of \$300/week (total \$3,000); graduate students, \$500/week (total \$5,000).

3.6. Organizational Structure

The overall organization is designed to maximize the interaction between the students and the faculty, and to integrate the education and research activities. A group of six faculty members from the participating institutions form the Core Instructors of this program and also serve as Research Mentors for the students. In addition, 17 other faculty from the participating institutions are also available as Research Mentors during the research component of the program. Prior to the start of the program and also during the program period, a Program Coordinator oversees the organization of all activities to ensure satisfactory implementation of the goals and vision. At the conclusion of the program, student evaluations are assessed by the Core Instructors and the Program Coordinator in preparation for the following year's program. All decisions and changes regarding the program are unanimously agreed upon prior to implementation.

3.7. Didactic Training

3.7.1 Coursework: A distinguishing feature of the BBSI program is the didactic training offered in the form of comprehensive coursework that covers the theoretical and computational aspects of computational biology as applied to molecular-to-cellular systems dynamics. This coursework, taught by the core instructors, is divided into three sessions (Table 2), and is briefly summarized below. A detailed description of the coursework is outlined in Appendix I.

Knowledge of a number of fundamental topics is required to understand the computational and mathematical models and algorithms developed for investigating the

dynamics of complex biological systems. In session I, basic mathematical, numerical and probabilistic methods are presented, along with a review of biochemistry, cell biology, and protein structure. Since the program brings together students from diverse academic backgrounds, this session is essential to program participants with minimal coursework in mathematics and biology. Session II teaches visualization tools and fundamental concepts of statistical mechanics and kinetics, and in conjunction with Session I intends to provide the necessary interdisciplinary theoretical background for the simulations of biological molecules and systems. Session III introduces the current problems and trends in computational biology, and models and methods are described for simulations at three different scales, namely, biomolecular, subcellular (microphysiological), and cellular. The students thus see the applications of the concepts and methods discussed in Session II to the simulations of biological molecules and systems.

Table 2: Organization of BBSI coursework sessions

| Session | General Topic | Inclusive Topics | *Hours |
|---------|--------------------------------------------------|------------------------------------------|--------|
| 1 | Review of Mathematical and Biochemistry Concepts | Linear Algebra | 3 |
| | | Differential Equations | 3 |
| | | Probability and Stochastics | 3 |
| | | Biochemistry Review | 4.5 |
| 2 | Computer Visualization | Data Visualization | 4.5 |
| | | Molecular Visualization | 4.5 |
| | Fundamentals of Computational Biology | Thermodynamics and Statistical Mechanics | 3 |
| | | Kinetics: Theory and Methods | 3 |
| | | Protein Folding | 3 |
| 3 | Multi-scale Simulations | Molecular | 6 |
| | | Subcellular | 4.5 |
| | | Cellular | 6 |

*Total: 48h of coursework and 36 h of computer laboratory sessions over a 4-week period.

3.7.2. Computer Laboratory Sessions: Computer laboratory sessions complement the lectures described above. These sessions are scheduled to introduce students to the software applications related to the topics covered during coursework, and allow students to receive hands-on knowledge of such software. Computer laboratory sessions are taught by the core instructors who are often assisted by teaching assistants familiar with the applications.

3.7.3. Research Experience: Research is clearly the most important aspect of this program. It forms the basis of the BBSI and trains students to identify, address, and solve a biologically significant question in a logical and organized manner. Based on their interests, students choose their research mentors and are matched to a laboratory by the Program Coordinator. After discussion of their research project with their mentors, all students are required to submit a written proposal that outlines their research goals during the program, the hypothesis that will be proposed and the methodologies that will be used to test the hypothesis, and the interpretation and significance of potential results. While research activity may be limited during the coursework component of the program, students are engaged in full-time research thereafter until the conclusion of the program.

3.7.4. Seminars: Three types of seminars are organized for this program: (i) research, (ii) career, and (iii) an ethics forum. Research seminars expose students to state-of-the-art research opportunities and challenges, and provide them with a variety of perspectives by active researchers in the field. Examples of topics include: characterization of allosteric effects in biomolecular complexes, mathematical modeling of cell cycle regulation, signaling/regulatory networks of apoptosis, simulation of microphysiological processes such as calcium dynamics, microarrays or gene expression profile analyses, sensor

technology in medicine, etc. Career seminars expose students to career opportunities, helping them to make cognizant decisions at graduation. These seminars present options in academia as well as prospects in industry. Ethical issues are an important aspect of research, and a team-based ethics forum is organized every year for students in the program. This allows the students to evaluate and discuss the implications of ethical choices they may be presented with during their career.

3.8. Professional Development

3.8.1. Student Presentations: All students are required to present an article related to their research project after consultation with their Research Mentor. Every student also gives an oral presentation of their research to the group and participates in the poster session scheduled at the Annual Duquesne Research Symposium. At the end of the program, a complete written report is submitted to both the Research Mentor and to the Program Coordinator for inclusion in the annual report to the funding agencies.

3.8.2. Team-based activities: Students are assigned to groups for the preparation of the research articles and for the ethics forum. The objective of this is two-fold: (i) it provides a foundation for team-based activities, and (ii) it allows students from different academic backgrounds to complement lack of coursework in a certain area.

3.8.3. Career workshops: Two career workshops are organized with current graduate students and postdoctoral fellows during which program participants have the opportunity to discuss the basis of selection of graduate schools, the application process, and career options following graduation.

3.9. Dissemination

All course material, including lecture notes, assignments, and solutions to assignments, are made accessible on the internet via a dedicated website at www.ccbb.pitt.edu/bbsi. This also includes student presentations and written proposals and reports. This type of information is useful both for future applicants and other institutes that are (or plan to be) engaged in similar efforts.

4. OBSERVATIONS, CHALLENGES, AND DISCUSSION

Consistent with the interdisciplinary nature of the field, the BBSI @ Pitt has accepted student participants from very diverse academic majors since the program's first year in 2003 (Table 3). While applications to the program were received from universities and colleges throughout the nation, a majority of these applications were from institutions in the Northeast and the Midwest, most likely due to their proximity to Pittsburgh. Student applications were evaluated on academic performance (GPA), completed relevant coursework, a personal essay describing their reasons for attending such a program, and any prior research experience. As seen in Figure 1A, the participating student pool was well represented by students entering their junior and senior years, indicating that there is strong interest amongst undergraduate students at a very early stage of their education to identify a potential career path and subsequently determine if it is the correct career choice for them. More importantly, these students were eager to apply knowledge acquired in traditional disciplines such as chemistry, mathematics, engineering, and computer science toward a biological question using non-traditional methodologies. Furthermore, the students were represented nearly equally by both genders (Figure 1B)

suggesting a lack of gender-specific preference for the field of computational biology and bioinformatics. While we have been successful in recruiting underrepresented minority groups to our program (Figure 1C), our goal is to increase that fraction within the next few years.

Table 3: Distribution of academic majors of student participants at the BBSI @ Pitt, 2003-2005. The program accepted a total of 42 students (14 students each year).

| Academic Major | Number |
|------------------------|--------|
| Biological Sciences | 20 |
| Biochemistry | 3 |
| Bioinformatics | 6 |
| Biology | 7 |
| Biophysics | 3 |
| Computational Biology | 1 |
| Chemistry | 2 |
| Computer Science | 6 |
| Engineering | 10 |
| Bioengineering | 5 |
| Biomedical Engineering | 4 |
| Computer Engineering | 1 |
| Mathematics | 4 |
| Applied Mathematics | 3 |
| Mathematics | 1 |

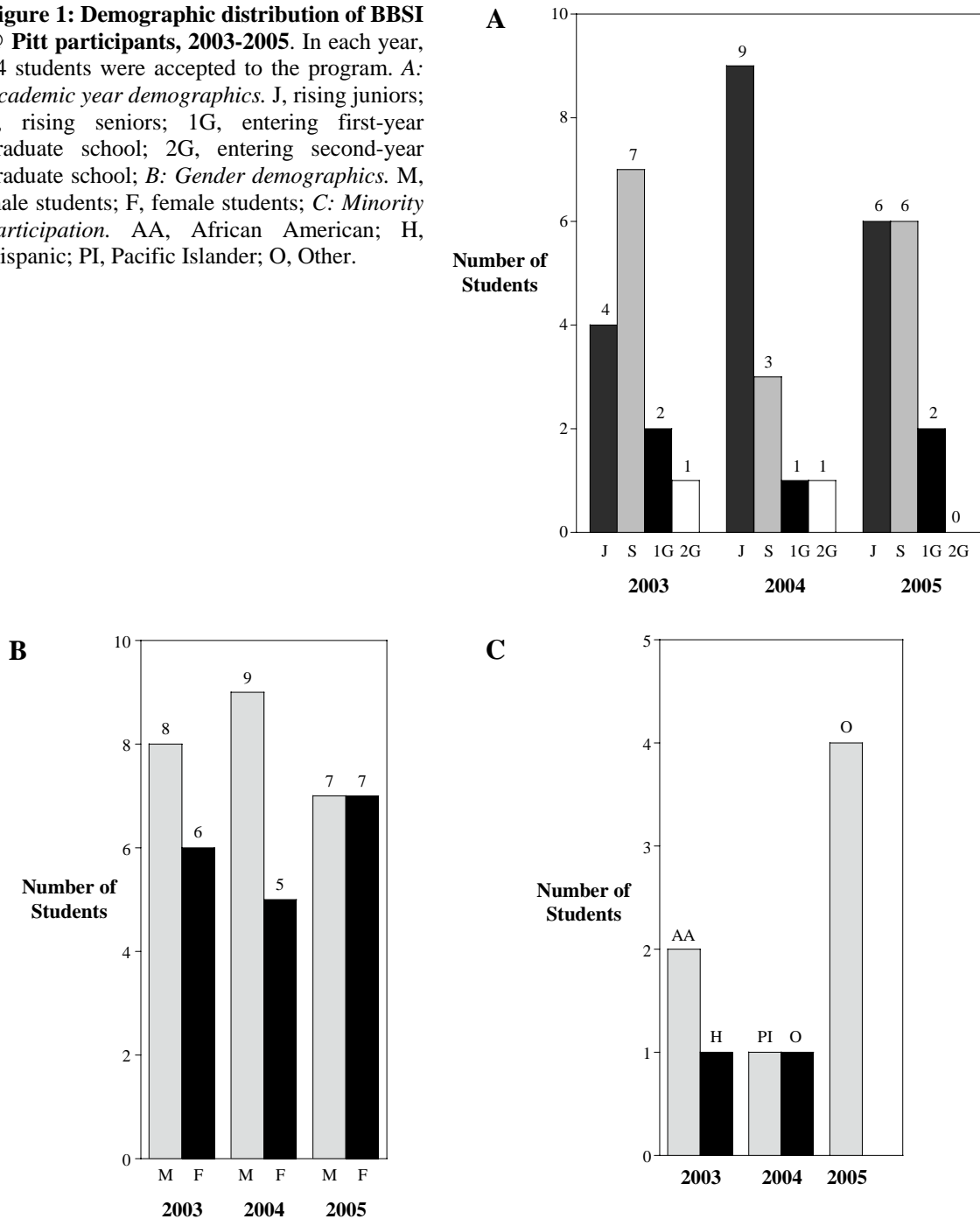
Challenges

During the two years that this interdisciplinary training program has been offered, we have identified four aspects/characteristics that distinguish it from traditional summer REU programs:

(1) Group heterogeneity: Due to the interdisciplinary nature of the field and the program, all program participants do not belong to the same academic major (Table 3). This results in a very diverse group of individuals with varied academic training and research interests. Thus, the BBSI is faced with the task of building a coherent learning

environment amongst students from vastly differing academic philosophies. This necessitates a teaching approach that is effective to biologists and engineers, and requires a higher degree of preparation of coursework material.

Figure 1: Demographic distribution of BBSI @ Pitt participants, 2003-2005. In each year, 14 students were accepted to the program. *A: Academic year demographics.* J, rising juniors; S, rising seniors; 1G, entering first-year graduate school; 2G, entering second-year graduate school; *B: Gender demographics.* M, male students; F, female students; *C: Minority participation.* AA, African American; H, Hispanic; PI, Pacific Islander; O, Other.



(2) *Academic curriculum:* The program includes comprehensive lectures and computer laboratory sessions, thereby broadening the student's view of post-genomic computational and mathematical research areas in molecular, cellular, and systems biology. A major challenge in developing an academic curriculum for such a diverse group is the ability of program participants to effectively understand the material at the desired level. We have found it necessary to include review sessions in mathematics and biochemistry to ensure that the lack of coursework in these basic fields does not limit understanding of advanced concepts in computational biology. Furthermore, the extent of coursework that can be covered is dependent to a great degree on the participant pool each year. Even though BBSI program participants are academically excellent, we recognize that group diversity may hinder the ability of the faculty to cover the desired coursework in the time period assigned (see next point).

(3) *Time constraint:* One of the major challenges faced by the core faculty is the time constraint in covering all planned didactic activities. Though the participating students are academically strong and eager to develop skills in computational biology theory and methodology, the breadth of this discipline precludes exhaustive coverage of every topic in the time available. Furthermore, varying interests within the student group may even make this undesirable. Thus, for the faculty preparing the lecture material for a particular topic, finding the right balance between necessary coursework and appropriate coursework can often take several years of teaching such a program.

(4) *Balance between coursework and research:* While coursework is undoubtedly necessary for a strong foundation in any field, especially for a relatively new discipline such as computational biology, the distribution of research and lecture time is critical.

This is even more significant for summer training programs due to the limited time during which an entire research project must be completed. Coursework must not only provide an appreciation of the potential of the field, but also provide the knowledge to effectively conduct research. Thus, it is key to find a compromise between fundamentals and application, thereby providing an appreciation of the potential of the field. Indeed, our student feedback has indicated that more research time is desirable because of the importance of making a substantial contribution towards a research project. Thus, emphasizing coursework over research, or vice versa, could lead to an unsatisfactory research experience.

In an effort to address the challenges and to improve the program, we have evaluated detailed online feedback surveys and personal discussions with students. While certain challenges, such as group heterogeneity and time constraints, are inherently part of such an interdisciplinary program, other challenges can be addressed effectively. As mentioned earlier, we have included additional review sessions in biochemistry at the suggestion of our students. We have greatly increased the amount of full-time research time available to students, though this could only be achieved at the expense of coursework time. Based on feedback from faculty, we have also included a computer laboratory session that reviews the basics of research computing for those students who may not be entirely familiar (or comfortable) with relevant tools such as a command line-based interface to the operating system.

Discussion

In recent years, a considerable amount of attention has been directed toward supporting interdisciplinary research, especially in biology programs (2, 11). A change in teaching principles has accompanied this interest, where active learning strategies have been shown to improve the learning process (12). We present here a novel summer research training program, the BBSI, which educates and prepares students for research in a relatively new field, and initiates the development of talented young scientists with the necessary skills to make an impact in the fields of computational biology and bioinformatics, as outlined in the NIH Roadmap Initiatives.

In a survey of students who completed at least some undergraduate research (4), 38.3% of the respondents indicated that they wished to pursue graduate education (PhD or MD/PhD) in a biology-related discipline, while an additional 14.2% indicated a desire to complete a PhD in the physical sciences. The goal of the BBSI program is to offer such students the opportunity to consider a career in computational biology, bioengineering, and bioinformatics by providing them with guidance and training at an appropriate time of their education. While a majority of students (56.5%) do not change their plans for graduate education following a research experience, such an experience does confirm or alter plans for 30.5% of the students surveyed (4). This indicates that novel summer research training programs, such as the BBSI, can indeed be effective in retaining and/or increasing the number of individuals wishing to pursue a career in this field.

The critical question is: “How effective has the BBSI @ Pitt been in providing a career path in computational biology?” While we have yet to establish and confirm the career paths of our student participants – primarily due to the infancy of the program, there are certain trends that we have been able to ascertain. (i) 36% of our students have indicated confidently that computational biology is the career choice for them. However, several of these students were entering their junior year at the time of this declaration; (ii) 28% of the students would likely or possibly continue in this field; (iii) Of the remaining 36%, 11% indicated that this was an unlikely career choice, while 25% stated with conviction that computational biology or bioinformatics was not a career option. This distribution notwithstanding, we believe that presenting an integrative view of molecular-to-system analytical models has been instrumental in introducing students from diverse academic backgrounds to the fields of bioengineering, computational biology and bioinformatics, and has allowed them to make an informed decision regarding their future career prospects.

5. ACKNOWLEDGEMENTS

The BBSI is a joint initiative of the NIH-NIBIB and NSF-EEC and was supported by the National Science Foundation under Grant No. EEC-0234002. All contributing authors of this manuscript acknowledge partial support by this award.

6. REFERENCES

1. Sung, N. S.; Gordon, J. I.; Rose, G. D.; Getzoff, E. D.; Kron, S. J.; Mumford, D.; Onuchic, J. N.; Scherer, N. F.; Sumners, D. L.; Kopell, N. J. Science education. Educating future scientists. *Science* **2003**, *301*, 1485.
2. Ares, M., Jr. Interdisciplinary research and the undergraduate biology student. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1170-1172.
3. Gross, L. J. Interdisciplinarity and the undergraduate biology curriculum: finding a balance. *Cell Biol. Educ.* **2004**, *3*, 85-87.
4. Lopatto, D. Survey of Undergraduate Research Experiences (SURE): first findings. *Cell Biol. Educ.* **2004**, *3*, 270-277.
5. Mogk, D. W. Undergraduate research experiences as preparation for graduate study in geology. *J. Geol. Educ.* **1993**, *41*, 126-128.
6. Tomovic, M. M. Undergraduate research—prerequisite for successful lifelong learning. *ASEE Annual Conference Proceedings* **1994**, *1*, 1469-1470.
7. Zydney, A. L.; Bennett, J. S.; Shahid, A.; Bauer, K. W. Impact of undergraduate research experience in engineering. *J. Eng. Educ.* **2002**, *91*, 151-157.

8. Hathaway, R. S.; Nagda, B. A.; Gregerman, S. R. The relationship of undergraduate research participation to graduate and professional education pursuit: an empirical study. *J. Coll. Stud. Dev.* **2005**, *43*, 614-631.
9. Bauer, K. W.; Bennett, J. S. Alumni perceptions used to assess undergraduate research experience. *J. Higher Educ.* **2003**, *74*, 210-230.
10. Nagda, B. A.; Gregerman, S. R.; Jonides, J.; von Hippel, W.; Lerner, J. S. Undergraduate student-faculty partnerships affect student retention. *Rev. Higher Educ.* **1998**, *22*, 55-72.
11. Cech, T. R.; Rubin, G. M. Nurturing interdisciplinary research. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1166-1169.
12. Handelsman, J.; Ebert-May, D.; Beichner, R.; Bruns, P.; Chang, A.; DeHaan, R.; Gentile, J.; Lauffer, S.; Stewart, J.; Tilghman, S. M.; Wood, W. B. Education. Scientific teaching. *Science* **2004**, *304*, 521-522.
13. Papoulis, A. Probability, random variables and stochastic processes. 2nd ed., 1965. McGraw-Hill, Inc.
14. Spiegel, M. R. Probability and stochastics: theory and problems of probability and statistics. Schaum's outline series in mathematics. 1975. McGraw-Hill, Inc.

15. Hill, T. L. Thermodynamics and statistical mechanics: an introduction to statistical thermodynamics. 1986. Dover, NY.
16. Bahar, I.; Erman, B.; Monnerie, L. Effect of Molecular-Structure on Local Chain Dynamics - Analytical Approaches and Computational Methods. *Adv. Polymer Sciences* **1994**, *116*, 145-206.
17. Ozkan, S. B.; Bahar, I.; Dill, K. A. Transition states and the meaning of Phi-values in protein folding kinetics. *Nat. Struct. Biol.* **2001**, *8*, 765-769.
18. Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **1977**, *81*, 2340-2361.
19. Gillespie, D. T. A rigorous derivation of the Chemical Master Equation. *Physica A* **1992**, *188*, 404-425.
20. McAdams, H. H.; Arkin, A. Simulation of prokaryotic genetic circuits. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 199-224.
21. Stiles, J. R.; Bartol, T. M. Monte Carlo methods for simulating realistic synaptic microphysiology using MCell. In *Computational Neuroscience: realistic modeling for experimentalists*, Ed. DeSchutter, E.; CRC Press, Washington, DC; 2001; pp87-127.

22. Allen, M. P.; Tildesley, D. J. Computer simulation of liquids. 1987. Oxford University Press, NY.
23. Leach, A. R. Molecular modeling principles and applications. 2nd ed., 2001. Prentice Hall, Pearson Ed. Ltd.
24. Oster, G. Brownian ratchets: Darwin's motors. *Nature* **2002**, *417*, 25.
25. Stiles, J. R.; Bartol, T. M.; Salpeter, E. E.; Salpeter, M. M. Monte Carlo simulation of neurotransmitter release using *MCell*, a general simulator of cellular physiological processes. In *Computational Neuroscience*, Ed. Bower, J. M.; Plenum, NY; 1998; pp279-284.
26. Tyson, J. J.; Chen, K.; Novak, B. Network dynamics and cell physiology. *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 908-916.
27. Fussenegger, M.; Bailey, J. E.; Varner, J. A mathematical model of caspase function in apoptosis. *Nat. Biotechnol.* **2000**, *18*, 768-774.
28. Aguda, B. D. A quantitative analysis of the kinetics of the G(2) DNA damage checkpoint system. *Proc. Natl. Acad. Sci. U. S. A* **1999**, *96*, 11352-11357.

Appendix I: BBSI Course Descriptions

Biochemistry Review: A knowledge of protein structure is essential to a deeper understanding of protein dynamics and molecular motions. This review describes amino acid structure, peptide bond formation, the four levels of hierarchy of protein structure (viz. primary, secondary, tertiary, quaternary), and Ramachandran plots. Examples are discussed to illustrate these concepts. The second of these reviews outlines basic principles governing cell signaling pathways, including secondary messenger formation, signal amplification, and eventual signal transduction. Finally, microarray technology, including cDNA microarrays and oligonucleotide arrays, are also discussed.

Linear Algebra: Review of basic concepts from linear algebra used in modeling and simulations. Description of matrix notation and matrix multiplication. Solving systems of linear equations using manipulations of matrices. Introduction of eigenvalues and eigenvectors, Taylor series and linearization of equations. Laboratory sessions: demonstration of programs such as MatLab and use of simple commands for solving equations and finding eigenvalues.

Differential Equations: Ordinary differential equations play a pivotal role in the kinds of modeling done later in the course. Starting with nonlinear differential equations, we discuss equilibria, in general. The behavior near equilibrium is determined by solving the linearized system of equations. This requires knowledge of the eigenvalues of an associated matrix. For differential equations that cannot be solved explicitly, we discuss

several simple numerical methods such as Euler and Runge-Kutta. Related software is demonstrated and used in the laboratory component.

Probability and Stochastics: Fundamental concepts of probability and stochastics are introduced, such as elementary probability spaces, permutations, combinations, random variable, expectation value, variance, probability density, stationary and nonstationary processes, stochastic differential equations, time averages and correlations, ergodicity and thermal noise. Brownian movements are presented as an important physical application of stochastic processes. Finally, Markov processes and transition probabilities are covered. In the laboratory session, simple problems from classical comprehensive textbooks (13, 14) are solved, and all of this material is related to subsequent lecture and lab work on microphysiological simulations based on Monte Carlo algorithms (Sessions II and III).

Data Visualization: We introduce issues such as data structure design and organization, data flow and manipulation, grid design and data point (string, scalar, vector, tensor, matrix, etc.) representation, using the open source package OpenDX (www.opendx.org). OpenDX includes a visual programming environment, software and hardware rendering engines, and extensive tutorials and documentation. In addition, this material serves as groundwork for more specialized visualization covered in Session II and III course work.

Molecular Visualization: Topics covered include importing and exporting coordinates, various rendering modes, visualizing and mapping various properties, and manipulation

of the structures. Software introduced includes InsightII, MOE (www.chemcomp.com), Cerius2, RasMol, WebLab Viewer, Molscrip, Midas, and Ribbons. In the laboratory sessions, students gain experience with common web-based search and retrieval systems.

Thermodynamics and Statistical Mechanics: First, the difference between mechanical and thermodynamic systems, and the relation between statistical mechanics and classical thermodynamics are presented. The canonical partition function is defined, discussed, and solved for an ideal gas. The concept of configurational partition function is introduced, followed by a discussion of the statistical nature of statistical mechanics. Finally, problems at the introductory level (15) are solved in the problem sessions to illustrate these theoretical concepts.

Kinetics: Kinetics includes both physical (structural change and/or diffusion) and chemical (reaction) kinetics. Elementary concepts of molecular mobilities, conformation and velocity distributions, changes (or fluxes) driven in the direction of decreasing potential (or concentration), and diffusion laws are introduced within the scope of physical kinetics. Chemical kinetics topics include the definition and applications of differential and integrated rate laws, the concepts of rate-controlling steps, steady-state approximation, transition states, and Michealis-Menten mechanism of enzyme reactions. From a different perspective, both the classical (simple mass-action) and modern (energy landscape) theories of kinetics applied to biochemical and biophysical changes are introduced. The master equation formalism is presented as a tool for exploring macromolecular motions (16), protein folding (17), and subcellular or cellular interaction

dynamics (18-20). In addition, stochastic interactions are covered together with Monte Carlo algorithms for 3D-reaction/diffusion problems (21).

Molecular Simulations: Molecular simulations require the definition of two essential features, geometry and energetics, of the investigated structure. Accordingly, models commonly used for representing molecular geometry, and the energy functions used for describing intra-and intermolecular interactions, are introduced. We then proceed to the description of the fundamental principles of molecular simulations, such as the numerical solution of Newton's equation of motion, or the evaluation of molecular forces as the negative gradient of potential. Simple algorithms for molecular dynamics (MD), Brownian dynamics (BD) and Monte Carlo (MC) simulations coupled with the Metropolis algorithm are presented. Students are referred to the comprehensive books by Allen & Tildesley (22) and Leach (23). The utility of adopting low resolution (or reduced) models as a means of proceeding to longer time or larger scale dynamics is emphasized and illustrated in the laboratory sessions. The students thus acquire a basic understanding of the methodology of MD, BD and MC simulations, and have access to related software and databases. Finally, the use of the mathematical methods presented in the first week (e.g. eigenvalue decomposition and correlation analysis) for analyzing MD or MC trajectories are illustrated in the problem sessions.

Subcellular Simulations: Models for various components in the cell are presented. We will consider models for molecular motors that drive much of the machinery for cell motility (24). As applications, some simple ion pumps are described as well as motors for

cytoskeletal components. Recent engineered gene networks are discussed in the context of building simple “computational devices”. Methods for characterizing the structure and function of protein channels (or pores), and the transport of simple ions (e.g., Na^+ , Cl^-) and biopolymers (DNA, proteins) through these pores are illustrated in the lectures and laboratory sessions. We discuss the techniques for studying non-equilibrium dynamics, which play an essential role in the operation of living systems (e.g., energy storage and conversion, signal transduction, etc.). Finally, the software MCell (21, 25), developed for quantitative simulations of synaptic physiology, is presented as an example of a simulator of cellular physiological processes. Students have access to the software and see its applications in the laboratory sessions.

Cellular Simulations: In this section, we describe, analyze, and simulate models describing cellular-level interactions. We examine the cell cycle models of Tyson and collaborators (26). We describe a model for apoptosis recently developed by Fusseneggar et al. (27) as well as our own extension of it to incorporate the effects of nitric oxide. Another topic is DNA damage and repair (28). Models for metabolic pathways (e.g. glycolysis and the associated oscillations) are also introduced. Recent models for the MinD/MinE system in bacteria are used to illustrate some partial-differential equations and pattern formation ideas. Both in lectures and lab sessions software packages such as VCell (www.nrcam.uchc.edu), MCell, E-cell (www.e-cell.org), and Gepasi (www.gepasi.org) are introduced.