Self-Determined Major Proposal: Computational Biology

Student Name: Kenneth Jung, Class of 2010

Address: 815 N. Broadway, Saratoga Springs, NY 12866

Faculty Advisors: Tom O'Connell, Patricia Hilleren(on sabbatical)

Core Course List:

1) BI105* (4) – Biological Sciences I - An introduction to the structures and processes common to all of life. The course explores topics in molecular biology, biochemistry, cell structure and function, transmission genetics, evolutionary theory, and population ecology. The laboratory portion of the course is inquiry-based and will introduce students to the methods and theory of modern biology. Three hours of lecture, three hours of laboratory per week. (Fulfills natural sciences requirement.) C. Freeman-Gallant and P. Hilleren

   As an introductory biology course, BI105 offers a sweeping summary of biology, exposing the various subfields computational biology can address.

2) BI106* (4) – Biological Sciences II - A comprehensive introduction to the diversity of life forms and life functions. The course explores topics in organismal biology with special emphasis on animals and plants, reproductive biology, physiology and developmental biology. *Prerequisite:* BI105. (Fulfills natural sciences requirement.) D. Domozych and J. Ness

   A continuation of the above course.

3) BI245* (4) – Genetics - A study of biological patterns of heredity explained by genes, their structure, function, and transmission from cell to cell and parent to

offspring, and the expression of genetic information. Topics include an in-depth study of mitosis, meiosis, Mendelian genetics and extension of Mendelian genetics, to complex traits and their analysis in individuals and populations. *Prerequisites:* BI106 or permission of instructor. B. Possidente

Gene sequencing is one of the most prominent topics in computational biology today (although more closely attributed to bioinformatics). This course provides a basic overview of genetics, and may promote further specialization.

4) BI247* (4) - <u>Cell Biology</u> - The course provides a cellular and organismal view of essential features of eukaryotic cell biology. Students will study cellular functions such as protein structure and function, cytoskeletal organization, cell migration, cellular metabolism, and cell signaling. In laboratory, students will gain experience with modern techniques for visualizing cell biological processes, with emphasis on differential interference contrast (DIC) optics, fluoresence, and confocal microscopy. *Prerequisites:* BI106, and CH105 or CH107H. J. Bonner

   Another quickly growing subfield of computational biology, modeling of both intracellular and intercellular interactions requires significant background knowledge. Gene expression, and the various methods of tracking changes, is also a main topic in this course, and a useful skill in molecular biology in general.

5) CH107H* (4) – <u>Intensive General Chemistry</u> - Study of the fundamental concepts of chemistry for motivated students who have a strong background in chemistry and intend to major in the natural sciences. Topics include atomic theory, chemical equilibria, acids and bases, electrochemistry, kinetics and bonding

theories. Emphasis is placed on active student participation and class discussion of course material. In addition, students are required to carry out an honors project (e.g., a written paper or poster) that involves library research on a topic in chemistry, proper citation of sources, and formal presentation to chemistry faculty and students. Laboratory experiments emphasize modern research techniques and instrumentation and prepare students for exploratory lab projects at the honors level that students complete during the last two weeks of the semester. Training in scientific writing will be an integral part of the laboratory experiences. Three hours of lecture-discussion and three hours of laboratory a week. *Prerequisites:* Consent of the department based on an online diagnostic exam administered during the summer, and QR1 results. Prepares students for CH207H and CH221. (Fulfills the natural science and QR2 requirements.) S. Frey, R. Nagarajan

The accelerated introductory course to chemistry, CH107H covers general inorganic chemistry, and is a requirement for the biochemistry major.

6) CH221* (5) – Organic Chemistry with Lab - The structures, physical properties, reactivity, and reaction mechanisms of aliphatic and aromatic hydrocarbons are investigated. The lab introduces the student to synthesis, purification, and chemical and spectroscopic methods of characterizing organic compounds. *Prerequisite:* CH106, 106H, or 107H. Three hours of lecture-discussion, and four hours of lab a week. R. Giguere, K. Cetto

   Organic chemistry goes hand-in-hand with molecular biology, providing a foundation on which higher biological concepts can be built. Synthesis of specific *organic compounds is also encompassed by computational biology, and this course provides a strong introduction.*

7) CH332 (5) – Physical Chemistry I with Lab - The fundamental principles and concepts of equilibrium thermodynamics including entropy, energy, temperature, heat, work, and chemical potential. Applications include chemical reactions, phase changes, environmental science, and biochemical systems. Lab experiments provide opportunities for quantitative experimental investigation of thermodynamic systems, including studies of heat exchange, chemical equilibrium, and phase equilibrium. Three hours of lecture-discussion and four hours of lab a week. *Prerequisites:* CH222 or 303, MA113, PY208. J. Halstead
Thermodynamics defines interactions that are frequently modeled or analyzed by computational biologists. Almost every published computational biology articles utilizes some concept from thermodynamics.

8) CH333 (5) – Physical Chemistry II with Lab - The fundamental principles of kinetic theory, reaction kinetics, statistical thermodynamics, chemical application of quantum mechanics, bonding, molecular spectroscopy, and structure. Lab and computer based experiments provide an opportunity for quantitative experimental investigation of phenomena such as reaction rates, transport properties, bonding, and spectroscopy. Three hours of lecture-discussion, four hours of lab a week. *Prerequisite:* CH330 or 332 or permission of the department. J. Halstead
A continuation of the previous course, this course covers quantum mechanics from a chemical perspective. Recommended by Prof. Standish over the offering

from the physics department for being more applicable to computational biology (less theoretical).

9) CH341 (5) – Biochemistry: Macromolecular Structure and Function with Lab - A study of the organic, physical, and biological chemistry of proteins, carbohydrates, lipids, nucleic acids, and enzymes. Structure-function relationships are explored at the molecular level using structural geometry and chemical reactivity concepts. The lab includes modern techniques for the purification, characterization, and identification of biomolecules. *Prerequisite:* CH222. Three hours of lecture-discussion and four hours of lab a week. M. Frey, R. Nagarajan
Biochemistry is the foundation of molecular biology, and is requirement by many graduate-level computational biology programs.

10) PY210* (4) – Foundations of Modern Physics - The significant historical discoveries leading to the development of atomic theory and quantum mechanics. Topics include discovery of the electron, blackbody radiation, the photoelectric and Compton effects, spectra, the Rutherford-Bohr atom, deBroglie waves, and Schrödinger's equation. Three hours of lecture, two hours of lab per week. *Prerequisite:* PY208. The Department
A brief introduction to quantum mechanics. A gateway course to all the 300-level physics offerings.

11) CS210* (4) – Design and Analysis of Algorithms - A study of techniques used to design algorithms that are efficient in terms of the time and memory required during execution. The course will also cover the techniques used to evaluate an algorithm's efficiency. Topics include advanced sorting techniques, advanced data

structures, dynamic programming, greedy algorithms, amortized analysis, and graph algorithms. *Prerequisites:* MC115 or MC215 and CS206, and MA111, or both MA108 and 109, or equivalent. The Department

The significant proportion of the computer science aspect of computational biology is dynamic programming. Other algorithms, as well as the overall methodology taught in this course, can be applied to problems in computational biology.

12) MC215* (4) – <u>Discrete Structures</u> - An introduction to mathematical reasoning in the context of studying discrete structures fundamental to both mathematics and computer science. Topics include elementary logic and sets, methods of proof including mathematical induction, algorithms and their analysis, functions and relations, elementary combinatorics, discrete probability, and graph theory. (Fulfills QR2 requirement.) *Prerequisites:* QR1, and CS106 or MA113, or permission of the instructor. The Department

    A prerequisite for CS210, as well as other higher level CS courses.

13) MC316 (3) – <u>Numerical Algorithms</u> - An introduction to using computation to obtain approximate solutions to mathematical problems. A variety of algorithms are studied, as are the limitations of using computational methods. Topics include algorithms for solving equations, systems, and differential equations; approximating functions and integrals; curve fitting; round-off errors, and convergence of algorithms. *Prerequisites:* CS106 and MA111, or both MA108 and 109, or permission of instructor. Offered on sufficient demand. The Department

Hugely applicable to modeling in computational biology, as the models usually include very difficult equations, thus requiring approximate solutions.

14) MA200* (4) – Linear Algebra - Vector spaces, matrices and linear transformations, determinants, solution of linear equations. *Prerequisite:* high school preparation including trigonometry or consent of department. Offered fall semester. (Fulfills QR2 requirement.) The Department

A prerequisite for higher level math and MC courses.

15) MA270* (4) – Differential Equations - An introduction to the theory and applications of differential equations. *Prerequisite:* MA113 and 200. Offered spring semester. The Department

Modeling in computational biology usually results in a large array of differential equations; this course offers methods for solving those problems.

16) MA202 (4) – Calculus III - Multivariable calculus. *Prerequisites:* MA113 and 200. Offered spring semester. The Department

Multivariate calculus is also highly utilized in modeling molecular systems; is a required program at similar undergraduate computational biology programs.

17) ID371 (4) – Interdisciplinary Independent Study - Independent work for juniors and seniors whose academic interests require an interdisciplinary approach beyond the academic structures available through established departmental courses. The student must have background appropriate to the proposed study, must have completed at least one other interdisciplinary course at Skidmore, must carefully define a plan of study, and must enlist the guidance of one or more faculty as appropriate. Proposals for ID371 and 372 are reviewed by the chair(s)

of the sponsoring faculty member's home department or interdisciplinary program; or in the case of students with an approved self-determined major, by the chair of the Self-Determined Majors Subcommittee.

Required by the self-determined major.

* - Have taken or am taking currently

** - Have placed out of

Possible Electives:

1) MA303 (4) – Advanced Calculus
2) MA204 (3) – Probability and Statistics
3) MA276 (3) – Selected Topics in Mathematics (Mathematical Modeling)
4) ID372 (3) – Interdisciplinary Independent Study
5) CH222 (5) – Organic Chemistry II
6) CH342 (3) – Biochemistry: Intermediary Metabolism
7) PY346 (4) – Electricity and Magnetism
8) MC306 (3) – Theory of Computation
9) MA324 (4) – Complex Analysis

Independent Project:

The proposal for the independent project addresses the subtopic of computational biology that most interests me: modeling. It will consist of constructing a mathematical model for a biological system on the molecular level. Although the exact system is undecided as of now, it will most likely incorporate electrostatics and binding thermodynamics and their influence on binding affinities of protein kinases, or similar molecules. Although the analysis can be purely theoretical, actual experimental data, whether collected by myself or obtained from literature, can be used to assess the validity of the model.

Projected Courses:

Spring 2009:

1) CH222 (5) – Organic Chemistry II
2) MC316 (3) – Numerical Algorithms
3) PY346 (4) – Electricity and Magnetism
4) MA204 (3) – Probability and Statistics
5) ID372 (3) – Interdisciplinary Independent Study

Fall 2010:

1) CH332 (5) – Physical Chemistry I with Lab
2) CH341 (5) – Biochemistry: Macromoleclar Structure and Function with Lab
3) MA303 (4) – Advanced Calculus
4) ID371 (4) – Interdisciplinary Independent Study
5) MA276 (3) – Selected Topics in Mathematics (Mathematical Modeling)

Spring 2010:

1) CH333 (5) – Physical Chemistry II with Lab
2) MC306 (3) – Theory of Computation
3) MA324 (3) – Complex Analysis

4) CH342 (3) – Biochemistry: Intermediary Metabolism

5) MA202 (4) – Calculus III

Rationale:

Computational biology is a new interdisciplinary field that is quickly gaining traction in the academic world. Computational biology is also known as computational chemistry, bioinformatics, computational genomics, mathematical modeling, or systems biology. As the various names imply, computational biology is a field where computational and mathematical strategies are used to solve or model complex problems in biology. Especially in this day and era, where science is able to amass large quantities of data, and where computers are only getting faster, computational biology is playing a larger and larger role in discovering new processes, automating older methods, and drawing connections between the various disciplines.

Most of the well-regarded science-oriented graduate schools already have a specific computational biology graduate program, including Yale, University of Washington, Harvard, and MIT. The computational biology major proposed is merely preparation for further studies at graduate-level institutions, and covers the many fields of science necessary for even the most basic work. From biology, computational biology requires a solid understanding of the many processes of life, on both macro and micro-scales. From chemistry, computational biology draws on everything from intra-atomic behavior to intermolecular interactions. From physics, computational biology requires knowledge of fluid, quantum, and thermodynamics. Mathematics is used significantly in all of the three major sciences, and computer science is used to harness the power of computing to solve unimaginably complex problems.

The proposed computational biology major recognizes the limitation of time, and is designed to maximize exposure to each field in hopes of stimulating interest in a specific subfield of computational biology. However, it also covers a solid foundation in each field, only slightly short of completing majors in each respective department. The lack of depth in biology is not an oversight; the program purposely avoids higher level biology classes in anticipation of the student taking the graduate-level counterparts. Instead, the program has a heavy chemistry, math, and computer science foundation, so that when the student finally does take graduate-level biology courses, he will learn from the perspective of a computational biologist.

All in all, the proposed major is a necessary stepping stone for any student wishing to pursue further studies and research in the specialized field of computational biology.

Special Aspects:

The proposed self-determined major differs from existing majors in the sense that it takes only the most relevant classes from each department in the interest of time. Simply pursuing a major in a single department, and even double-majoring, would leave the student with a lack of knowledge in the other sciences. This major intends to focus on breadth over depth, to help those that have not yet chosen a subfield from the vast selection within computational biology.