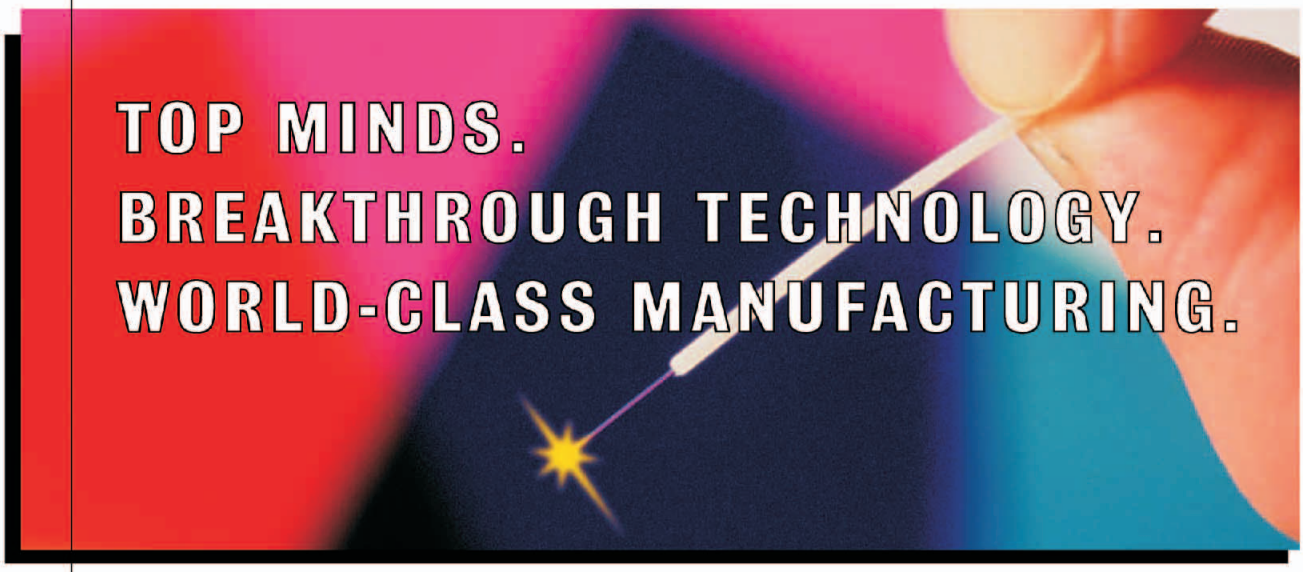
CURJ

APRIL 2001
VOL.1 NO.1

CALTECH UNDERGRADUATE RESEARCH JOURNAL

PREMIERE ISSUE

COMPUTER GESTURE RECOGNITION



THE OPTICAL FOUNDATION FOR CONNECTING THE PLANET

At Novalux, we've attracted some of the top minds in the laser and compound semiconductor industries. Like CTO Aram Mooradian, former lead of the Quantum Electronics Group at MIT's Lincoln Labs, CEO Malcolm Thompson, former CTO of Xerox PARC, and Hem Takiar, Director of Packaging at National Semiconductor. All that's missing is your unique contribution! You'll join teams perfecting our NECSEL platform—a tiny, surface-emitting laser with higher power coupled into fiber than any laser previously demonstrated. If you have a passion for working at high velocity and challenging the de facto methods of engineering and manufacturing, then consider these opportunities at our Sunnyvale location.

MANAGEMENT
- VP Engineering
- VP Sales and Marketing
- Director of Manufacturing Operations
- Product Managers

DEVELOPMENT ENGINEERING
- RF Engineers
- RF/Semiconductor Engineer
- Device Engineer
- Optical Engineer
- Sr Assembly Engineer
- Laser Engineer

MODELING
- Semiconductor Laser Modeling Engineer
- Optical Modeling Engineer
- Thermal Modeling Engineer
- Device Simulation Software Engineer
- Mechanical Design, Modeling and Simulation Engineer

WAFER FAB
- Sustaining Engineer
- Sustaining Process Tech

MECHATRONICS
- Automation Engineer
- Mechanical Engineer
- CAD/Design Drafter
- Tech

PACKAGING
- Sustaining Engineer
- Laser Packaging Engineer
- Mechanical Engineer
- Laser Packaging Tech

GROWTH
- Characterization Engineer
- MOCVD Engineer

RELIABILITY/TEST
- Failure Analysis Engineer
- Yield Analysis Engineer
- Statistician
- Tech

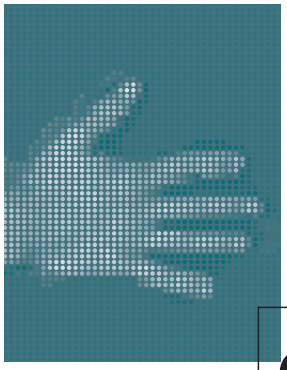
Thrive in a high-velocity career at a well-funded, pre-IPO startup. Send your resume for immediate consideration to **employment@novalux.com, mail to Novalux, Inc., Attention: Human Resources, 1170 Sonora Ct., Sunnyvale, CA 94086, or fax to (408) 730-3987.** Novalux is an Equal Opportunity Employer.

NOVALUX

LIGHT THAT MATTERS

www.novalux.com

CURJ
APRIL 2001



Cs$_1$

COMPUTER GESTURE RECOGNITION: USING THE CONSTELLATION METHOD by Dinkar Gupta

If machines could respond to our actions, they would truly be at our beck and call. A new algorithm brings computers closer to recognizing commands through our hand movements.

P.26

REVIEWS

RESEARCH

Ap *Applied Physics*
Be *Biomedical Engineering*
C *Chemistry*
Ce *Chemical Engineering*
Cs *Computer Science*
G *Genetics*
N *Neuroscience*

FINIS

SURF
PROGRAM

Caltech's Summer Undergraduate Research Fellowships program introduces undergraduate students to research under the guidance of seasoned mentors.

Students experience the process of research as a creative intellectual activity and gain a more realistic view of the opportunities and demands of a professional research career.

For more information go to
http://www.its.caltech.edu/~surf
or phone 626.395.2885

FROM THE EDITOR

Like Sauternes dessert wine, a successful journal must at once be enjoyable and refined. While preparing the premiere issue of the Caltech Undergraduate Research Journal, we were determined to create a composition that could be relaxing to read, and still offer the highest quality of undergraduate research.

More undergraduates than ever participate in laboratory work, and CURJ can help you find out what that work is about. In Reviews, get up to speed on cutting edge fields. In Research, read about discoveries that other undergraduates are making. In Finis, sample an artistic or literary perspective on science. Our Fall issue will also feature Letters for viewpoint pieces from undergraduates, prominent scientists, and industry leaders.

We've worked to ensure the journal is a satisfying vintage. Our selected authors and content team have developed a clear and informal writing style, for a rich flavour. Our production team adds color and artwork like a sweet bouquet. CURJ is available online and in print, in Spring and Fall. The online version features enhanced HTML text and PDF reprints of all editions.

CURJ is produced through the talent of undergraduates, staff, and faculty at the California Institute of Technology and the Art Center College of Design, in Pasadena, California. We publish writing from undergraduates anywhere in the world. We'd love to see your submissions. Visit our website to learn how to add your fine ingredients to our next release.

CURJ is distributed free of charge through generous support from Caltech Student-Faculty Programs, a partnership with the Caltech Alumni Association and the Art Center College of Design, and a strong showing from smart business and recruiting advertisers. Visit our website for advertising information and a list of our sponsors.

Pour yourself a glass of CURJ, online at www.curj.caltech.edu and in print now at Caltech, Art Center College of Design, MIT, University of California Los Angeles, and the Jet Propulsion Laboratory. Taste it, and let us know what you think. Write to us, and contribute. But for now, kick back, relax, and enjoy our April cuvée.

Ram Srinivasan
Editor-in-Chief

CURJ

CALTECH UNDERGRADUATE RESEARCH JOURNAL

www.curj.caltech.edu

TISSUE ENGINEERING

BY AL VALDIVIA

EVERY YEAR IN THE U.S. AND AROUND THE WORLD, MILLIONS OF SURGICAL procedures are performed that require the use of tissue substitutes. As a result, there is an ever-growing need for tissue and organ substitutes to replace or repair diseased or otherwise damaged body parts. Several conventional approaches towards solving this problem have met with some success but have severe limitations. The current approach to providing these substitutes depends on grafting existing tissue from either the recipient's own body or from another donor. Unfortunately, both these methods have problems that hamper their impact. The rise of tissue engineering as an alternative approach has the potential to reduce or even remove many of these limitations. Tissue engineering is a relatively new interdisciplinary field that allows for the development of viable tissue substitutes by combining the principles of biology and engineering. This approach allows for new therapies, which use tissue that can be seamlessly integrated into the existing tissue, thereby improving greatly on current methods. Tissue engineering has its own set of obstacles, but unlike the problems of some current approaches, most of them are solvable in theory.

TRANSPLANTS TODAY

The conventional means of providing tissue substitutes rely almost exclusively on using existing tissue for grafts. This approach is hampered by certain inescapable limitations regarding tissue rejection, availability of donors, and the risks involved in any surgery.

Autografting is the process of harvesting a tissue from one part of the body and transplanting it to another. Typically these types of grafts exhibit the best results because rejection isn't an issue. However, the process does require invasive surgery

"[Grafting] is hampered by certain inescapable limitations regarding tissue rejection, availability of donors, and the risks involved in any surgery."

that exposes the patient to significant risk of infection, pain, and blood loss. Of the approximately 300,000 coronary autograft bypasses performed each year, 20% result in infections and post-operative pain. Besides the potential risks associated with the transplant itself, autografting is less than ideal because the types of tissues available for reconstruction are often unsuitable and result in loss of some function at the donor site.

Allografting is the procedure of transplanting one person's tissue into the body of another person. Typically, the graft is accepted at first, but within 11-15 days is rejected and attacked by the body's immune system. This occurs because the immune system's T lymphocyte cells identify the graft as foreign tissue and begin to coordinate its elimination, much as they would invading bacteria. Lymphocytes are a special class of white blood cells, which are a key component of the body's immune system. T-cells in particular are responsible for cell-mediated immunity, meaning that they identify whether or not cells in the body are foreign. This immune response presents a major obstacle to the potential of tissue transplants and over the years, scientists and doctors have devised means to minimize the threat of rejection.

On their surface, all cells contain peptide-binding proteins called Major Histocompatibility Complex (MHC) molecules that are key components of how cells recognize each other. The MHC molecules present cell-specific peptides to the outside environment and T-Cells sense these peptides and bind to them. This binding triggers an immune response and ultimately results in the death of the tissue graft. This interaction makes MHC compatibility the clearest indicator of initial graft rejection or acceptance. If donors and recipients have comparable MHC molecules, the immune system perceives the transplanted tissue as its own and thus ignores it. Matching donors and recipients by MHC molecules has made it possible to greatly reduce the risk of initial rejection. Nevertheless, rejection over the long term is still common because of other minor histocompatibility antigens that trigger a smaller but significant immune response. Consequentially, unless the donor and recipient are identical twins, the patient must take immunosuppressive drugs indefinitely. Such a regimen is dangerous as these drugs leave the patient especially vulnerable to other infections. As a result of the unavoidable incompatibility, two different cell types can result not only in graft rejection, but also in significant drug therapy costs.

The practice of xenotransplantation was developed in response to the limited availability of donor organs, and has met with mixed success. Xenotransplantation is the term for an allograft, known as a xenograft, from a non-human species. Some thirty or so years ago, doctors experimented with transplanting livers, kidneys, hearts, and bone marrow from species such as chimpanzees and baboons. Unfortunately, none of these transplants were successful, as the patients involved succumbed to organ rejection or severe infections. Another danger introduced by using non-human tissue as a source for tissue substitutes is the danger of transgenic diseases. Often, the human immune system has never had to deal with certain diseases common among these animals and consequentially, patients' immune systems are caught unprepared. Over the past ten years or so, xenotransplantation has come to focus on baboons and pigs and has met with increased success. Scientists are currently attempting to genetically alter the organs of these animals to make them more compatible with the human immune system and develop therapies to eliminate various potentially fatal transgenic viruses. However, xenografts still have a great deal of development to undergo before they can be considered viable alternatives to allografts.

Autografts and allografts have been refined and perfected over the years to the point where successful transplants now rely on procedure rather then luck, but they still constitute non-ideal solutions. All these procedures call for surgery that inherently results in patient pain, places the body in danger of functional loss and creates a need for a lifetime of drug therapy.

"The difference between a colony of cells and a tissue is the level of organization present in the tissue, which allows it to accomplish its task."

These are only the dangers faced by those individuals who are fortunate enough to receive a transplant; every year thousands of individuals die while still on waiting lists that are far too long for the number of organs available. During the past decade there has been an exponential increase in the number of patients on the waiting list, as Figure 1 demonstrates.

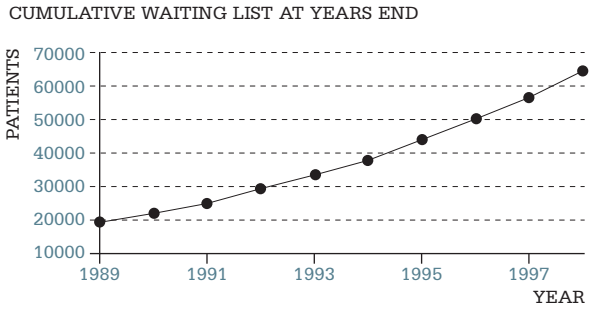


FIGURE 1. Number of Patients on U.S. Organ Lists at years end
Source: United Network for Organ Sharing, (www.unos.org) Annual Report for 1999

Although transplantation techniques and practices have improved dramatically over the past ten years, there is still a lot of room for improvement. Figure 2 displays the percentage of patients who manage to survive certain transplants sorted by survival periods of varying length.

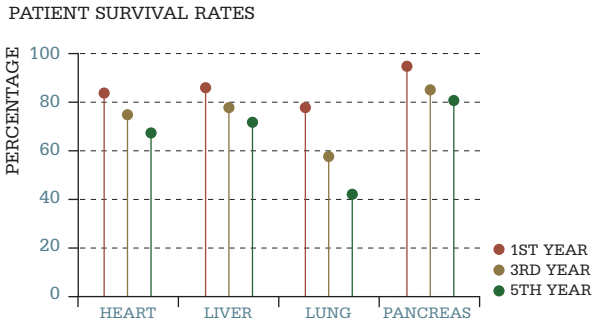


FIGURE 2. Percentage of patients who survive certain transplants, presented over three time intervals
Source: United Network for Organ Sharing, (www.unos.org) Annual Report for 1999

TISSUE ENGINEERING: A NEW HOPE?

Tissue engineering is a relatively new interdisciplinary field that holds promise for synthesizing skin, cartilage, and eventually entire organs. This is accomplished by combining biological and engineering principles to create new tissue substitutes. It has long been possible to culture individual cells in artificial mediums, but this differs greatly from the conditions required to fabricate a tissue. The difference between a colony of cells and a tissue is the level of organization present in the tissue, which allows it to accomplish its task. Cells form tissue structures only after receiving a variety of cues (chemical, structural and electrical) directing them to do so. Recent experiments are attempting to artificially stimulate the differentiation of cells into viable tissue constructs. The experiments that attempt to mimic the cell's environment and cues have shown the greatest success.

In the body, the majority of cues emanate from the extracellular matrix (ECM) that surrounds all cells. The ECM is made up of a framework of gel and connective tissue that fills the spaces in the body between organs and tissues. Another important set of components of this cellular growth pathway is various specialized proteins, called growth factors. Growth factors are secreted by specialized cells and regulate the variety of cues received by cells. These growth factors are essential components of how the body controls its cell growth. The basic approach to tissue engineering, pioneered by Joseph Vicanti and Robert Langer, is to recreate this cellular environment as best as possible. Their research has resulted in the

development of scaffold-guided tissue regeneration, which is currently used in one form or another for virtually all tissue engineering. This process involves seeding a porous (biodegradable and made of cellulose or other similar substances) scaffold with donor cells and then bathing the region in the appropriate growth factors (see Figure 3 below). The donated cells adhere to the scaffold and begin to differentiate into a shape resembling that of the desired tissue. In this scheme, the scaffold serves as the template for the emerging tissue structure in place of the ECM. Over time, the scaffold biodegrades and leaves behind a functioning tissue structure in the desired shape. This approach typically results in patients experiencing significantly reduced levels of pain (both during surgery and afterwards) and scarring.

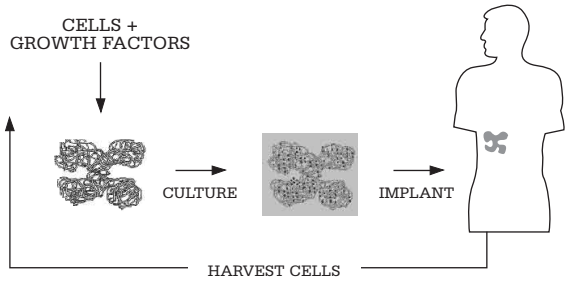


FIGURE 3. Overview of Scaffold-Guided Tissue Regeneration
Source: http://www.pittsburg-tissue.net/about_te/techniques.html

REPLACEMENT PARTS, MADE TO ORDER

Most of the early groundwork on tissue engineering was laid during the 1970s and 1980s. This work was unnoticed by the general public and was conducted without government funding. As a result, most of the researchers involved were forced to move towards industry for funding and this fact is reflected in the rapid development of tissue engineering companies. In recent years, the Food and Drug Administration (FDA) and the National Institute of Health (NIH) have paid increasing amounts of attention to this new discipline because it has begun to generate actual products and to show real potential. Tissue engineering has the potential to greatly alter the manner and cost of health care in the US and abroad. In 1993, it was estimated that $400 billion dollars, approximately half of the annual national health-related cost, is spent on individuals afflicted by organ or tissue loss. It has also been reported that along with the individuals who die on the organ transplant waiting lists, approximately 100,000 people each year die who, because of their present medical state or medical history, were not able to even qualify for the list. The sheer economic and humanitarian potential of tissue engineering has helped fuel its advance in recent years.

Current tissue engineering research is progressing from relatively simple tissue, such as skin, towards tissue of an increasingly complex structure, such as fully functional organs. Tissue such as skin and cartilage is geometrically simple to construct and requires a relatively small blood delivery support system. Consequentially, trials that have focused on these tissues have experienced the greatest amount of success thus far. Organogenesis Inc. has been able to transform a few seed cells of skin into vast quantities of cells ready to be grafted into patients. This product, Apligraf, uses foreskin cells from an infant which have not fully developed their identifying proteins. The issue of tissue rejection could potentially be avoided by using these cells as seeds. Organogenesis received FDA approval to market Apligraf in the treatment of ulcers in 1998. Considering that there are approximately 600,000 Americans suffering difficult-to-heal ulcers every year, this represents a significant breakthrough. Since then, another company, Genzyme Tissue Repair, has received FDA

> "Genzyme Tissue Repair has a means to repair knee-cartilage injuries by harvesting some donor cells, growing them and then re-introducing them into the injured site."

approval to market a means to repair knee-cartilage injuries by harvesting some donor cells, growing them and then re-introducing them into the injured site. Depending on the patient and the extent of the injury, the knee-cartilage can be fully healed within 12 to 18 months. The early success with growing cartilage has encouraged others to use it in new areas such as treating urological disorders. Some of this research is being done by Reprogenesis, which is studying whether cartilage cells could be harvested, grown and then re-introduced into patients' urinary tracts as a means to replace missing muscle tone, which is characteristic of incontinence.

The success of skin and cartilage trials has prompted other researchers to work on larger tissue structures. Researchers at the Carolinas Medical Center are attempting to use tissue from a woman's leg to stimulate the growth of tissue for the purposes of replacing breasts lost to mastectomies or lumpectomies. Viewing this project as engineering a larger tissue is deceptive, because the new tissue would contain none of the functionality of a true breast, but it is a step in the right direction and could eliminate the need for implants or prostheses. It is through these small steps that researchers are progressing towards the goal of whole organ synthesis. Researchers at Cedars-Sinai Medical Center have been able to synthesize liver-like tissue that is capable of displaying the partial functionality of the whole organ. Other researchers at Harvard Medical School's Children's Hospital have been able to generate kidney-like constructs that possess some of the filtering capabilities of fully functional kidneys. Additionally, researchers at the University of Toronto are working on a long-term project to synthesize a human heart within 10 to 20 years. Considering that in 1999, cardiovascular diseases cost the U.S. around $326.6 billion dollars in direct and indirect costs, this is a project that could have huge ramifications for the health care industry in the US and around the world.

Another developing field, which has the potential to make a huge impact on the future of tissue engineering, is that of stem cells. Most cells in the body are terminally differentiated, which means that they are fully developed into their unique type and are locked into that structure. Stem cells, on the other hand, are a special brand of cells that exhibit the unique property of being able to reproduce and have their progeny develop into any of the body's specialized cells. The body uses stem cells to replace cells that have lost the ability to divide on their own and must be replaced. The outer lining of your skin and inner lining of the small intestine are examples of tissue that is replaced through the growth and development of stem cells.

These cells could have a tremendous impact on the field of tissue engineering by allowing for the generation of an endless supply of specific cell types. These cells types could lead to vast improvements in organ growth technologies. Several companies around the U.S. are currently trying to develop medical uses of stem cells. SyStemix is attempting to improve bone marrow transplants by using blood-producing stem cells. Using these cells it is possible to make a recipient's immune system tolerant of any of the donor's other cells. Such an advancement could open up the door to a wide range of possibilities for combating transplant rejection. Osiris Therapeutics of Baltimore is in the testing stages of a stem-cell preparation technique that has the potential for regenerat-

ing cartilage and other types of damaged tissue. Recent advances have shown that the stem cell approach has a great deal of potential and could help tissue engineering overcome some of its current hurdles.

SETBACKS AND THE FUTURE

There are various obstacles that must be overcome before tissue engineering can reach its full potential. The most significant of these difficulties is the generation of a blood supply delivery network to feed the increasingly complex synthesized structure. Tissues that exceed a few millimeters in thickness require the introduction of additional blood vessels. Research aimed at developing a means to starve tumors of their blood supply resulted in the isolation of angiogenesis-stimulating molecules, molecules which stimulate the construction of new blood vessels. The use of these molecules works well for relatively small structures, but the angiogenesis molecules would take too long to reach the core of a complex structure such as an organ and in the meantime the organ would die.

Another difficulty inherent in increasing tissue size comes along with the necessary increase in scaffold size. It is difficult to distribute donor cells evenly on a large scaffold, particularly toward the center of the structure, where the diffusion of cells is impeded. A potential solution would be to construct the scaffold with the cells pre-attached, but current construction techniques occur in environments too hostile for living cells to survive. Since organs are complex and consist of different types of cells, scaffold structures must contain various donor cell and corresponding growth factor densities. Such scaffolds cannot be constructed with current building techniques; new, more flexible biomaterials must be developed in order to build them. If tissue engineers wish to create universal donor cells, then the old problem of rejection reappears unless it is somehow possible to strip the cells of identifying proteins or use undifferentiated stem cells. Unfortunately, stem cells themselves are not free of complications. Due to the nature of their tasks, stem cells must divide on a timescale that greatly exceeds that of the typical cell. This pace leads to premature aging and poses a potential limit to the use of stem cells. Work is currently being carried out on embryonic stem cells which are able to side step this problem. However, due to ethical concerns, this research is moving along at a very sluggish pace. Before the engineering of complex tissue structures can become a reality, these problems must be resolved.

Al Valdivia is a third year undergraduate in Chemical Engineering at the California Institute of Technology. *The author wishes to thank Mark Davis, Warren and Katharine Schlinger Professor of Chemical Engineering at Caltech.*

FURTHER READING

B. Alberts *et al. Molecular Biology of the Cell* (Garland Publishing, New York,), . - .

C. Arnst, J. Carey. Biotech Bodies. *Business Week* **July** , - ( ).

R. D. Guttmann. The Graft Survival Curve: Ideology & Rhetoric. *Transplantation Proceedings* , - ( ).

C. Janeway, P. Travers. *Immunobiology: The Immune System In Health and Disease* (Current Biology Limited, San Francisco,).

J. Klein. *Immunology* (Blackwell Scientific Publications, Boston,), . - .

H. Lodish *et al. Molecular Cell Biology* (W. H. Freeman & Company, New York,), . - .

D. J. Mooney, A. G. Mikos. Growing New Organs. *Scientific American* ( ), - ( ).

L. C. Paul. Chronic Rejection of Organ Allografts: Magnitude of the Problem. *Transplantation Proceedings* , - ( ).

R. Pichlmayr *et al.* Organ Transplantation: What Are the Limits? *Transplantation Proceedings* , - ( ).

G. Stix. Growing a New Field: Tissue Engineering Comes Into Its Own. *Scientific American* ( ), - ( ).

CONSCIOUSNESS, MEET NEUROSCIENCE

BY FLORIAN MERKLE

FOR MILLENNIA, PHILOSOPHY AND WESTERN RELIGION HAVE TENACIOUSLY maintained that the body and mind are two separate entities that together form a conscious human being. This view is perhaps best presented in the writings of Descartes. In his *Meditations*, he argues that the fusion of mind and body takes place in the pineal gland, located in the center of the brain. Nowadays, however, scientific evidence strongly suggests that consciousness is a property that is *derived* from structures such as the human brain. Mind and matter are not split: the latter is intimately connected to the former. This realization raises some interesting questions. *How* does the brain produce consciousness? How can we determine if a system is conscious? If the human brain is conscious, can a machine be conscious? To begin to answer these difficult questions, it would be a good idea to find some common ground to stand on. We should agree on a handful of premises.

FIGURE 1. René Descartes (1595-1650), perhaps the most influential of the modern philosophers. It is now difficult to escape from his conception of the duality of mind and matter, which permeates almost all modern thought in psychology, philosophy, and neuroscience.
Source: Gregory, 1997

PREMISES

We should agree that consciousness actually exists and that the human brain creates human consciousness. Consciousness has been described again and again in a myriad of different accounts. It is noteworthy that most accounts of consciousness state that subjectivity and attention are important to consciousness, so these phenomena should be taken into account when determining if something is conscious. There is also an important distinction, which will be described later, between consciousness proper and the contents of consciousness. Finally, it must be recognized that although consciousness is ubiquitous in our lives, it is extremely difficult to study and we know very little about it. As you read this document, you are probably conscious of the little symbols on the page that represent words, the meaning of the words, and the fact that this sentence is written in English. If I call your attention to it, you are also conscious of your breathing and the feel of the chair you are sitting in. These conscious experiences are taken for granted; they are our most immediate source of information, yet they are poorly understood. Most studies of consciousness are therefore very exploratory and theoretical, as they are in any young field of science. We are living in exciting times: research in upcoming decades may revolutionize the way that we think about thought itself.

DEFINING THE QUESTION

What is consciousness anyway? We all have a pretty good idea of what we mean when we say "consciousness", but all attempts to definitively describe it have failed. The word has so many different uses and definitions that some believe that it has lost all meaning. However, from the multitude of accounts of consciousness, there does seem to be a common element: that of subjectivity or "phenomenal experience". When you touch a stove, your hand will reflexively withdraw and it is only later that you have the phenomenal (conscious)

"Although consciousness is ubiquitous in our lives, it is extremely difficult to study and we know very little about it."

"The main contribution of AI studies was the realization that the brain and the serial computer resemble each other only minimally."

experience of pain. Although a machine might be able to withdraw from a source of excessive heat, it is not likely that it would *feel* the pain and say "ow" of its own accord.

Another common element to many definitions of consciousness is that of attention. Before I asked you to attend to your breathing, chances are that you were not conscious of how fast or deeply you were breathing. The state of being aware of a stimulus is distinct from the state where a stimulus is present but is not in the awareness of the observer. There are many studies that support this, including fMRI (functional magnetic resonance imaging) and visual attention studies. Attentional studies show that when subjects attend to an object, different areas of the brain are activated than when they do not attend to it, even though the object is "seen" in both cases. Reaction times are also much slower when subjects are asked to respond to a change in a random non-attended object than if they are asked to respond to a change in an attended object.

There is an important distinction between consciousness proper and the contents of consciousness. The term *contents of consciousness* refers to consciousness of sensory information whereas *consciousness proper* is an awareness that is not directly dependent on sensory input. Visual awareness, which is closely linked to visual attention, is an example of a content of consciousness. However, consciousness proper is usually meant when consciousness is discussed. Whether pure consciousness can exist in the absence of all sensory input is not certain. There is debate about whether a blind man is less conscious—in the sense of consciousness proper—than a man who can see. This point of contention may be circumvented when addressing the questions we have asked, but it is crucial to note that the word "consciousness" carries different meanings in different contexts. Unless otherwise noted, we will assume that "consciousness" refers to consciousness proper.

PLANS OF ATTACK

The questions we have posed have been attacked from many angles, leading to several schools of thought. The first, philosophy, is also the oldest. Philosophers have themselves posed some interesting questions and attempted to answer them in large part though introspection and thought experiments. However, these methods generally do not lead to definitive answers, and the results give anything but a consensus. Since the only undeniably conscious system is the human brain, this is the system that should be studied. In this regard, philosophers are clearly on the right track. However, in order to get answers that can be supported with evidence and compared objectively, the brain must be studied scientifically. To do this successfully presupposes knowledge of the anatomy and physiology of the brain and neurons, the brain's substituent elements. Scientific approaches to the problem of consciousness include: artificial intelligence based on the computer—brain analogy, anatomical and physiological study, and abstract modeling that functionally simulates the way groups of neurons process information.

ARTIFICIAL INTELLIGENCE

Unfortunately, the main contribution of artificial intelligence (AI) studies was the realization that the brain and the serial computer resemble each other only minimally. The belief that

the brain works like a serial computer has entrenched itself as one of the great misconceptions of popular science. In reality, the brain is highly parallel, highly connected, abstract, relatively slow, and computationally probabilistic. A computer, on the other hand, is serial, weakly connected, millions of times faster than the brain, and completely logical. The hardware and software analogy to the brain that was drawn in the early days of AI is simply inapplicable. Memory storage is another difference between the brain and the computer that cannot be disregarded. Memory is stored in discrete units in computers, but in the brain, memory and processing capability is globally distributed.

Artificial intelligence has proven unable to robustly display the behaviors associated with consciousness. Most "behaviors" that AI systems display are not emergent properties of computation; rather, they have been explicitly programmed to feign cognition. Due to the relative lack of communication between processing units and the serial nature of processing in computers, it is unlikely that such a system will become conscious in our sense of the word. For this reason, we will devote no more time to AI. Anatomical studies of the brain, however, have produced many revelations.

FUNCTIONAL AND ANATOMICAL STUDIES

It is very difficult to understand the nature of consciousness and its source without understanding the nature of the brain. The monkey brain and the human brain are structurally and functionally very similar. Hence, intimate knowledge of the primate brain has greatly contributed to the study of consciousness. The main advantage of studying the brain directly is the fact that there are many different methods of obtaining tangible and reproducible information. When data can be acquired by different methods and compared, the knowledge gained from the comparison is more accurate and comprehensive than when the information stems from only a single source of data.

The disadvantage of this approach is that the brain is incredibly complex. It is difficult to reduce things to a level where they can be understood. The brain is heavily interconnected, from the level of molecules all the way up to the level of brain systems. To study one area of the brain means to study not just the area itself, but also (and most often unwillingly) its connections to other areas, the chemicals that happen to be present in the area at the time, the changes in activity that the particular method of measurement causes, and so on. To better understand the problem of complexity and the problem of determining the mechanism of consciousness, we need to know some basic facts about the brain.

BRAIN BASICS

There are tens to hundreds of different types of neurons in the human brain (see Figure 2). The word "brain" generally refers to a portion of the central nervous system that sits in the cranium and is continuous with the spinal cord. In humans, it is made up of about 10 billion neurons, the brain's fundamental computational and structural elements. Neurons are excitable cells, meaning that they are able to discharge electrically. Though they usually discharge in response to a stimulus, they can also fire autonomously. The term "firing" refers to a coordinated electrical discharge. The frequency, patterns, spatial location, and timing of these discharges carry the information of the brain's computations.

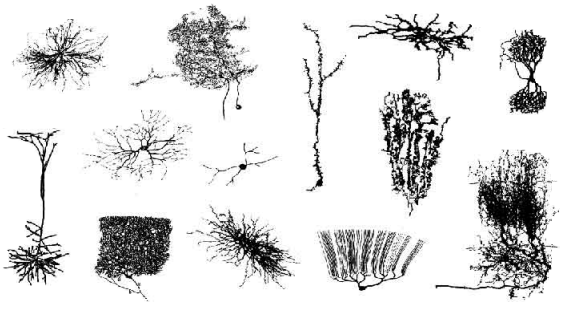


FIGURE 2. An assortment of neuronal types. Some are highly branched and complex, while others are quite a bit simpler. In addition to being structurally different, neurons are functionally different in their responses to stimuli. Each type of neuron is specialized to perform a particular task.
Source: Koch, 1999

"To study one area of the brain
means to study
not just the area itself,
but also (and most often unwillingly)
its connections to other areas,
the chemicals that happen to be present
in the area at the time
and the changes in activity
that the particular method
of measurement causes."

Practically all neurons can be split into three main structural elements: the dendrite, the soma and the axon (see Figure 3). The soma is the cell body from which the axon and dendrites originate. Dendrites are branchlike projections that transmit electrical impulses from other cells to the soma. The axon is the long structure that carries electrical impulses away from the soma and can range from fractions of a millimeter to several meters in length. The axon, particularly when it must transmit electrical signals over long distances, is covered in a fatty material called myelin. The axon is wrapped in myelin to ensure that the signal will be transmitted along the entire length of the axon without the loss of signal intensity. The presence of myelin makes the tissue fatty and white in areas where there are many axons. This is the source of the white color in white matter, while gray matter gets is color from cell bodies and unmyelinated tissue. Gray matter, which is dense in neurons, is found in areas where a lot of computation happens, such as the cortex.

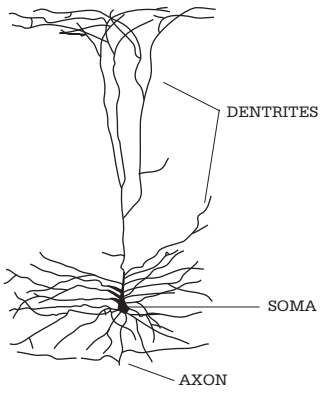


FIGURE 3. A typical cortical neuron. The axon, of which only a small portion is shown, connects the neuron to neurons in other layers of cortex, while the lateral projections off of the axon allow it to communicate with neighboring cells. The dendritic branching is quite complex, in contrast to the rather minimal axonal branching.
Source: Zigmond, 1999

Neurons communicate with each other with chemical messengers called neurotransmitters. An electrical signal prompts one neuron to release neurotransmitters from the end of its axon over a very small gap to a second neuron that has receptors for the neurotransmitters. The system of neurotransmitters, the receptors, and the gap is called a synapse. The transmitting and receiving neurons are called the presynaptic and postsynaptic neurons, respectively. The presynaptic cell stores neurotransmitters at the ends of its axon, while the receptors of the postsynaptic neuron are found mainly in its dendrites. The release of neurotransmitters can cause an electrical signal to be generated in the postsynaptic cell. Therefore, the mechanism of information transfer between neurons is electrical to chemical to electrical.

The brain can be divided into the cerebrum and cerebellum, the brainstem, the midbrain and the basal ganglia. The cortex is on the surface of the folded, outer layer of the cerebellum. It is divided into four lobes: the frontal lobe, the parietal lobe, the temporal lobe, and the occipital lobe. The basal ganglia include the amygdala, the globus pallidus, the putamen, and the caudate nucleus. They are heavily connected to many structures including the thalamus, which acts much like a "relay station" for the brain.

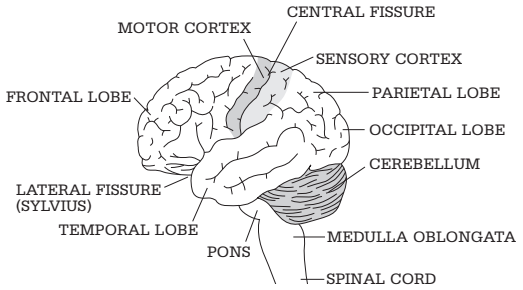


FIGURE 4. The lobes and major features of the brain, as seen from the side.

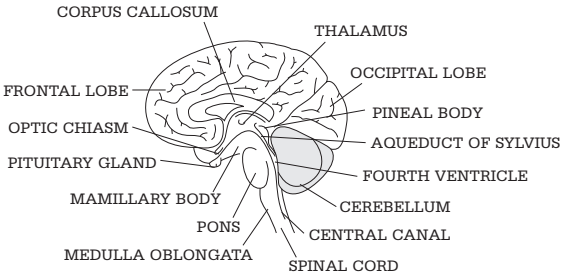


FIGURE 5. Important features of the brain, as seen when the brain is cut vertically down the midline and then viewed from the side. Note the central location of both the pineal body and the thalamus. The basal ganglia (not shown) also occupy this general area.

INSIGHTS FROM THE DIRECT STUDY OF THE BRAIN

Anatomical and physiological studies have uncovered several central facts. First, there is an anatomic correlate of function, meaning that functional properties may be ascribed to structural properties. This correlation is not absolute, since there are generally no clear borders between where one brain area ends and another begins. More often than not, a function such as vision is the result of the interaction of many areas. This can make analysis of particular areas particularly difficult. For example, when only one area is damaged and one sees a deficit in a particular behavior, then it is not correct to say that the damaged area is responsible for producing that behavior. It is probably true that an important connection was damaged, with interesting but usually unclear repercussions. There is, however, enough localization of function in the brain that it is possible to predict the deficit that will be seen if a particular area is damaged.

Second, processing is often hierarchical in nature. For instance, hierarchical organization has been observed in the visual system in the form of layers and as associated areas. The removal of any of these areas has fascinating effects on the overall perception. Lower areas, or areas that are early in the visual pathway, are involved in detecting the crude characteristics of visual stimuli. Higher areas seem to be involved in internal representation of visual stimuli, and have been associated with conscious perception.

Third, the connections and functions associated with structures are not necessarily logical or intuitive. Rather, they were created and shaped by the forces of evolution. Nature is a tinkerer. She takes working systems, modifies them slightly to take over new functions, and finds creative ways to achieve the desired results. The result is a brain that doesn't have the qualities that an engineered brain might have, which makes study of it even more difficult.

Fourth, inhibition plays a very important role in the proper functioning of networks of

cells. While excitatory cells will respond to a particular stimulus by integrating the information and transmitting the message, inhibitory cells decrease the activity of the neurons they affect. Excitation and inhibition depend on the neurotransmitters that are used to relay information from the presynaptic to the postsynaptic neuron. Excitatory neurotransmitters increase the probability that a cell will fire within a certain time window, while inhibitory neurotransmitters decrease this probability. The interaction between excitation and inhibition spatially and temporally constrains neuronal information relevant to the stimulus, and thereby increases the efficiency and resolution of processing. Somewhere between 20% and 50% of all neurons in the human brain are believed to be inhibitory, a percentage that is often underestimated or ignored.

Finally, brain regions are highly connected in many directions. Feedback and parallel connections play a vital role in the regulation of computation. Again, there is an outstanding example in the visual system. On the way from the eye to the primary visual area (V1), the optic nerve passes through the lateral geniculate nucleus (LGN) of the thalamus. For every nerve fiber sent from the LGN to V1, ten are sent back, suggesting that what we see is strongly influenced by what we have just perceived.

DISTRIBUTION AND LOCALIZATION

There are several schools of thought among neuroscientists regarding the localization of consciousness in the brain. The first hinges upon the belief that consciousness is a distributed, emergent property, a certain coherence of activity that transcends traditional methods of analysis and understanding. Arguments for distributed consciousness are based on observations about other properties of the brain. The structure of connections and synapses between neurons, in addition to neuronal firing patterns, generates the behavior of the network. In this sense, information from past experiences and memory are globally encoded.

"Consciousness might also be globally distributed, residing nowhere in particular but resulting from the orchestrated simultaneous functioning of many different systems."

When neurons are slowly removed from this global distribution of information, the system still functions more or less normally. This effect is called graceful degradation. This is in stark contrast to serial computers, which undergo so-called catastrophic degradation since a purely serial system cannot compensate for the loss of a component. These observations suggest that consciousness might also be globally distributed, residing nowhere in particular but resulting from the orchestrated simultaneous functioning of many different systems. The difficulty with this theory, however, is that it cannot truly be tested unless we understand all the intricacies of the brain. Furthermore, this view does little to delimit which systems are conscious and which are not. Still, given the fact that we know so little about the subject of consciousness and there is no large body of compelling evidence to the contrary, it is a position that many neuroscientists take.

"If indeed there is a neuroanatomical correlate of consciousness, the ILN is a good candidate."

THE NEUROANATOMICAL CORRELATE OF CONSCIOUSNESS

A second, smaller, fraction of neuroscientists thinks that there is an anatomically distinct machinery, or neuroanatomical correlate of consciousness, that evolved to generate consciousness. This belief has some very interesting consequences. Advocates of this theory state that consciousness is the synthesis of information from many different sensory modalities, none of which are independently required for consciousness to exist. Sensory information is incorporated into consciousness only after it has gone through significant processing. An elegant experiment that illustrates this was done with blindsight patients. Due to a stroke or trauma that injured the primary visual area of half of the brain, blindsight patients claim to be blind to visual information from the side opposite the injury. These patients were presented an array of lights in their neglected hemifields. They were asked to point to the light that was lit, and though they claimed that they could see nothing at all, they could point to the correct light with surprising accuracy. The information was there, but somehow it could not be integrated into the consciousness of the patient.

If indeed there is a neuroanatomical correlate of consciousness, where might it be found? It is probably represented bilaterally, since most neural structures show lateral symmetry. Studies with split-brain patients also support this idea. A split brain patient has had his corpus callosum, the large bundle of fibers that allow the two hemispheres of the brain to interact, severed surgically as treatment for seizures. When tested, the patient behaves as if he had two separate conscious brains. For example, Roger Sperry and his colleagues showed that a monkey's hemispheres could be taught to elicit conflicting responses to the same situation, "as if the animal [had] two separate brains".

Surprisingly, and in contradiction to many scientists who believe that consciousness is distributed, consciousness it not necessarily found in cerebral cortex. There have been patients who independently have bilaterally lost one or more entire lobes but retained consciousness. These lesions can be huge, covering up to 40% of all cortical area. For the same reason, the hippocampus, a structure important for memory, probably does not produce consciousness either. Some researchers believe that the reticular activating system (RAS), creates consciousness, but this is probably on the wrong track. The RAS arouses the brain and thus is more likely to activate consciousness than to produce it. While cortical lesions may be huge and have little effect on consciousness, small, specific lesions may make the person unresponsive.

Very small lesions to the thalamus, especially the intralaminar nucleus of the thalamus (ILN), can eradicate consciousness. The ILN is a bilaterally, centrally located structure that receives input from many areas, including association, visual, parietal and prefrontal cortex, the optic tract, some of the basal ganglia, and the hippocampus. It projects back to motor cortex, the reticular nucleus of the thalamus, and some of the basal ganglia. All of these connections would be important to a structure that was mediating consciousness. If indeed there is a neuroanatomical correlate of consciousness, the ILN is a good candidate.

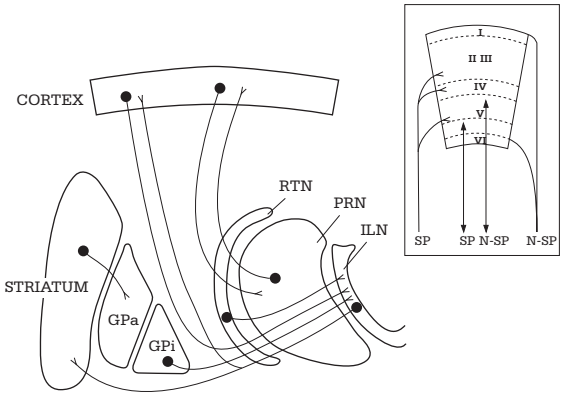


FIGURE 6. A highly schematic outline of the connectivity of the dorsal thalamus and the associated reticular nucleus of the ventral thalamus. The filled-in circle represents a connection that begins in a layer and the V-shape represents a connection ending in a layer. GPe = external division of the globus pallidus, GPi = internal division of the globus pallidus, RTN = reticular nucleus of the thalamus, PRN = principal nucleus of the thalamus, ILN = intralaminar nuclei of the thalamus, SP/N-SP = specific/non-specific afferents in the cerebral cortex. Roman numerals indicate cortical layers.
Source: Bogen, 1995

Theories that claim that there is a particular location in the brain that creates consciousness are very difficult to verify experimentally and are still speculative. The same constraints and difficulties that apply to localizing any area of brain apply to localizing the area that is the neuroanatomical correlate of consciousness. The boundaries between brain areas are not distinct and fade into each other. Connected areas have the capability of performing the function to some degree if the primary mechanism is damaged. For these reasons, as well as the fact that consciousness itself is so poorly understood and therefore not readily comparable to other brain functions, the number of supporters of this theory is quite small.

THE SIMILAR MECHANISMS APPROACH

Another approach with relatively few advocates states that all aspects of consciousness employ similar mechanisms. This approach applies to processing the contents of consciousness, such as vision or pain, and assumes that the mechanisms that lead to consciousness of each of these modalities are similar. Proponents of this theory believe that consciousness is created in the neocortex and some areas of paleocortex. The rationale for this claim is that when these areas are damaged, certain contents of consciousness are lost. Since the claim is that similar mechanisms are utilized to generate different forms of consciousness, emphasis has been placed on the genesis of consciousness in the primate visual system, which is well studied and documented. Of particular interest is selective visual attention.

Selective visual attention is a serial process that acts like a spotlight, bringing only small areas into our awareness at a time. Anyone who has ever tried to find Waldo in a Where's Waldo® book has experienced it. The role of attention is probably to bring things to visual awareness, making us conscious of the attended phenomena, such as Waldo. The precise mechanism of attention and how it mediates awareness is not clearly understood, but it may be the result of the semi-synchronous activity of relevant areas. A deeper understanding of this mechanism might allow us to apply the basic principles to other sensory processing areas and so understand how we are conscious of each of them. This understanding, in turn, might illuminate the causes of consciousness proper. Just how it is produced is still a mystery, but there are a variety of approaches to directly studying the human brain that may clarify the issue. But even if we understand the mechanics, the conundrum of determining when something is conscious still remains.

HOW CAN WE TELL?

A haunting problem for those who seek to identify consciousness lies in Descartes' famous aphorism "cogito ergo sum" (I think, therefore I am). Though we are able to judge that we ourselves are conscious, we have no way of knowing whether any other system is conscious. Since we can only observe in others the behaviors that are associated with consciousness, we cannot know what the system is experiencing. As Gray puts it: "One's own conscious experience is a datum that stands in need of explanation; the conscious experi-

"It is simply impossible to prove that others are conscious and that the apparent consciousness of others is not the elaborate deception of an 'evil genius' who put you in the world with zombies..."

ences of theirs, however, can function only as a hypothesis by which to explain their behavior." This problem is sometimes referred to as the "zombie problem": it is possible that everyone but ourselves is a zombie that claims to be conscious and behaves as such, but has no conscious experience of anything. It is a problem that is debated in the field of philosophy, but cannot be attacked by science. It is simply impossible to prove that others are conscious and that the apparent consciousness of others is not the elaborate deception of an "evil genius" who put you in the world with zombies who only appear to be conscious. It has therefore been considered to be an unsolvable problem.

However, there are characteristics of our own observable consciousness that we can apply to the apparent consciousness of others. Armed with the faith that there is no evil genius, one can assume that if other beings rigorously meet the criteria, then they too are conscious. The question of which criteria are important is hotly debated, but there are certain criteria that are frequently included. These are alertness and attentiveness, ability to cooperate and interact with others, responsiveness, apparent sense of self and subjectivity, and characteristic EEG rhythms. These criteria are guidelines rather than absolutes. So far, clinical utilization of these criteria has provided a fairly good indicator of consciousness in humans. When trying to apply them to animals, however, there are clear problems. It is hard to tell if a cat has a sense of self. To apply these criteria to machines is fraught with still more problems. This is an important point that must be considered when trying to construct an artificial consciousness: it may be impossible to tell when the goal has been achieved.

MODELS

Almost all neurobiologists would agree that consciousness is rich in information. The brain can handle very large amounts of information about an attended phenomenon in a short amount of time, but can also switch rapidly between attended phenomena. We know of no machines that have this property, but if we could construct them, we might gain insights into consciousness that, by current methods and understanding, cannot be directly obtained from studying only the brain. Studying the brain is a messy business, because there are

so many variables that can interfere with what you want to study. To simplify the problem, models called artificial neural networks (ANNs), which mimic the parallel processing of a biological group of interconnected neurons, are being increasingly utilized. The advantage of using these models is that they provide a controlled, simplified environment in which to experiment and observe behavior in a way not possible by studying the brain directly. Of course, there is a fine balance that must be achieved between oversimplifying the problem and including so many details that the model is too complicated to analyze efficiently. In consideration of this problem, there are two main classes of ANNs: the connectionist model, which is governed mainly by computational constraints and tends to apply a more theoretical approach, and "biological networks", which are constrained mainly by biological considerations including the anatomical and physiological properties of different cell types.

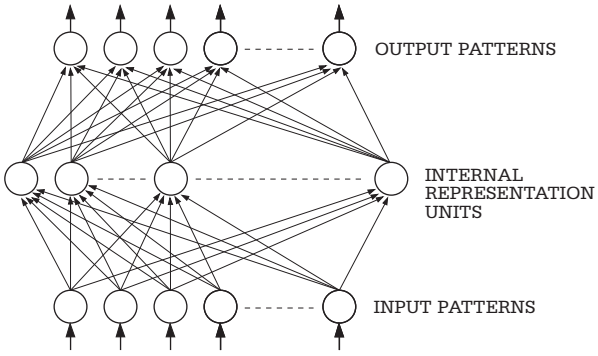


FIGURE 7. An example of a simple multilayer feed-forward network. Internal representation units are sometimes referred to as "hidden units". After training with a learning algorithm or other process, the network can learn to give the desired output pattern for a given input pattern.
Source: Crick, 1994

At the basis of all ANNs there are three elements: the node, the connection, and the modification of connections. The behavior of the node, though simplified, is computationally similar to that of a neuron. Typically, the nodes are arranged in layers, resulting in hierarchical processing. One layer is connected to the next either unidirectionally or bidirectionally. The connections themselves are given varying strengths or weights depending on the saliency of the connection to the desired output. The weights are determined by one of many learning algorithms. When two nodes exhibit similar behavior, the connection is strengthened; when they do not exhibit similar behavior, the connection is weakened. This is referred to as Hebbian learning and is similar to biological processes of synapse modification by activity. Hebbian learning is the basis for many learning algorithms.

There are many different types of ANN, but all share some very interesting, brainlike characteristics. Like a brain, neural networks can cope with loss of processing units with no drastic effect on the output; in other words, they degrade gracefully. An ANN can have several different "memories" for different tasks, in that a particular pattern of connectivity can be superimposed over another with minimal change in the output of either. Hence, an ANN can be trained to perform several different tasks. These memories are also similar to memories in the brain. If the memories are similar or too numerous, the network can get confused and give the wrong output. We have all experienced this phenomenon ourselves at one time or another, perhaps while taking an exam that required lots of memorization.

Unfortunately, many of the impressive feats of ANNs are achieved though the use of unbiological structures or learning methods,

Join Caltrans' Team

Take California's Transportation system into the 21st Century!

Caltrans is involved in an aggressive program to seismically retrofit all bridges in California, including the San Francisco - Bay Area toll bridges.

Caltrans needs:

- Civil Engineers
- Transportation Planners
- Land Surveyors

Visit our website for more hiring & exam information.

www.dot.ca.gov

Caltrans

XonTech, Inc.

People, Science and Technology

www.xontech.com

XonTech, one of the nation's premier Research & Development firms, specializes in the processing and analysis of data, interpretation of results, algorithm development, performance assessment, modeling and test planning. Over the years, we have carefully attracted the best scientific and technical specialists to work with leading-edge technology. Do you have what it takes to join our elite team? We have many exciting positions at our offices throughout the country for:

❖SOFTWARE ENGINEERS❖

❖PHYSICISTS❖

❖MATHEMATICIANS❖

❖ELECTRICAL ENGINEERS❖

We seek bright, inquiring minds with PhD's, Master's or Bachelor's degrees in Physics, Mathematics, Computer Science or Electrical Engineering. Experience or education in Signal Processing is desired.

U.S. Citizenship is required for all positions. Selected candidates will be subject to a government security investigation and must meet eligibility requirements for access to classified information.

XonTech offers competitive salaries and comprehensive benefits in a stimulating R&D environment. For immediate consideration please submit your resume to: **XonTech, Inc., Employment Dept.-IT, 6862 Hayvenhurst Ave., Van Nuys, CA 91406; FAX: (818) 787-8584; or E-mail to: human_resources@xontech.com.** EOE M/F/D/V

Northern & Southern California • Alabama • Colorado • Florida
Ohio • Massachusetts • Washington, D.C.

Engineering

AN INVENTIVE MIND knows no limits

Litton Guidance & Control Systems, *is a world leader in the design, development, and operation support of advanced inertial navigation electronics systems for ground, space, airborne, and seaborne applications.*

Opportunities available in the following disciplines:

Electrical & Computer Engineering

Material Science

Physics

Math

Mechanical Engineering

Some positions may require a security clearance.

If you are graduating with a BS/MS/PhD, please fax or send resume to: ***LITTON GUIDANCE & CONTROL SYSTEMS, HUMAN RESOURCES-DEPT GRAD, 5500 Canoga Avenue, Woodland Hills, CA 91367*** *or* ***E-mail: millerd@littongcs.com***, *or* ***Fax (818) 715-2488***. *For more information on Litton Guidance & Control Systems and these extraordinary opportunities, visit our website at:* ***www.littongcs.com***. *EOE*

Litton
Guidance & Control Systems

Structural Design Engineers

DES Architects + Engineers, a leading mid-Peninsula A/E firm, seeks Structural Design Engineers with CE or SE license, minimum 5 years experience in wood, steel, concrete and masonry. Projects include schools, hospitals, corporate campuses, etc.

Competitive salary and complete benefit package.

Send resume with intro letter and salary requirements to: DES, PO Box 3599, Redwood City, CA 94064. Fax to 650.364.2618, Attn: Niki Hithe or email to jobs@des-ae.com.

DES
ARCHITECTS
ENGINEERS

www.des-ae.com

so it is difficult to justify comparing them to neural networks in the brain. Structural connections in ANNs are often contrived and never encountered in nature. ANNs are also minuscule when compared to areas of cortex that contain many millions of neurons. However, although it has not been achieved, it is conceivable that a machine based on the principles of ANNs that exhibits the properties of consciousness could be built. Of course, the construction of such a machine would rely on our knowledge of the mechanism of consciousness in the human brain. Thus, the real utility of ANNs lies in their relative simplicity and the ease and speed with which they allow new theories to be tested. This will assist in unraveling the intricate functioning of the human brain and thus indirectly contribute to our understanding of the origin of consciousness.

CONCLUSION

Models and neuroscientific data complement each other and they will continue to do so as more knowledge is accumulated. Biological studies have revealed many of the defining characteristics of neural computation and organization. From our observations, we can draw several conclusions. First and foremost, consciousness is a slippery subject and will remain difficult to define and assess. It is still unknown whether consciousness has a specific anatomical correlate or whether it is distributed. It seems reasonable that attentional mechanisms might be closely connected to consciousness, but this too is debatable. Careful study of anatomy and physiology of the primate brain will continue to be the best way to study consciousness at many levels of complexity. Neural networks are more controlled and provide a clean field on which to test new theories. As they become larger and more "biological", they will become more useful tools to tackle the question of consciousness. It is conceivable that machines that exhibit observable features of consciousness could be constructed, but it will be difficult to tell when the goal has been achieved.

Florian Merkle is a third year undergraduate in Biology at the California Institute of Technology. *The author wishes to thanks Christof Koch, Lois & Victor Troendle Professor of Cognitive & Behavioral Biology and Professor of Computation and Neural Systems at Caltech.*

FURTHER READING

I. Aleksander. *Impossible Minds: My Neurons, My Consciousness* (Imperial College Press, London,).

J. E. Bogen. On the Neurophysiology of Consciousness: . An Overview. *Consciousness and Cognition* , - ( ).

J. E. Bogen. Memory: A Neurosurgeon's Perspective. In *Surgery of the Third Ventricle*. M.L.J. Apuzzo, Ed. (Williams and Wilkins, Baltimore, MD,).

S. R. Cajal. *Histology of the Nervous System of Man and Vertebrates*. N. Swanson, L. W. Swanson, Trans. (Oxford University Press, New York,).

F. Crick. *The Astonishing Hypothesis: The Scientific Search for the Soul* (Touchstone, New York,).

F. Crick, C. Koch. Towards a Neurobiological Theory of Consciousness. *Seminars in the Neurosciences* , - ( ).

F. Crick, C. Koch. Consciousness and Neuroscience. *Cerebral Cortex* , - ( ).

R. Descartes. Meditations on the First Philosophy in Which the Existence of God and the Distinction Between the Body and Soul are Demonstrated. In *Classics of Western Philosophy*. S. Cahn, Ed. (Hackett Publishing Company, Indianapolis, IA,), . - .

H. Eichenbaum. Conscious Awareness, Memory and the Hippocampus. *Nature Neuroscience* , - ( ).

J. A. Freeman, D. M. Skapura. *Neural Networks: Algorithms, Applications, and Programming Techniques* (Addison-Wesley Publishing Company, New York,).

J. A. Gray. The Contents of Consciousness: A Neuropsychological Conjecture. *Behavioral and Brain Sciences* , - ( ).

R. Gregory. *Eye and Brain: The Psychology of Seeing* (Princeton University Press, Princeton, NJ,).

P. N. Johnson-Laird. *The Computer and the Mind* (Harvard University Press, Cambridge, MA,).

C. Koch. *Biophysics of Computation: Information Processing in Single Neurons* (Oxford University Press, New York,).

T. Metzinger. Introduction: The Problem of Consciousness. In *Conscious Experience*. T. Metzinger, Ed. (Imprint Academic, Exeter, UK), . - .

W. Penfield. Studies of the Cerebral Cortex of Man: A Review and an Interpretation. In *Consciousness Symposium: Studies of the Cerebral Cortex of Man* ( ).

D. E. Rumelhart, J. L. McClelland, the PDP Research Group, Eds. *Parallel Distributed Processing* (Bradford Books, MIT Press, Cambridge, MA,).

T. J. Sejnowski, C. R. Rosenberg. Parallel Networks that Learn to Pronounce English Text. *Complex Systems* , - ( ).

T. J. Sejnowski, P. S. Churchland. Brain and Cognition. In *Foundations of Cognitive Science*, M. I. Posner, Ed. (MIT Press, Cambridge, MA,) . - .

C. Williams. *Free Will and Determinism: A Dialogue* (Hackett Publishing Company, Indianapolis, IA,).

M. J. Zigmond *et al*. *Fundamental Neuroscience* (Academic Press, San Diego, CA,).

COMPUTER GESTURE RECOGNITION:
USING THE CONSTELLATION METHOD

BY DINKAR GUPTA

IMAGINE A COMPUTER THAT SHUTS OFF IF THE USER simply waves goodbye, a TV that changes channels when the viewer does a "thumbs up" or "thumbs down", a room whose lights you can turn on and off with a flick of your fingers. In short, consider a world where simple, everyday tasks would not require grappling with levers, squinting over control panels or hunting for the remote control. Instead, all you would need is a simple hand gesture to turn up the stereo just when your favorite track starts.

The attempt to have machines recognize dynamic hand gestures is hardly unreasonable. After all, humans use hand gestures to communicate and get things done all the time. In fact, human gesticulation is extremely sophisticated. Simply consider clenching your fist and raising one finger. Every finger has a different connotation—the thumb implies "well done" or "keep it up", the index finger implies the number one, and so on.

Developing such functionality for computers and microprocessor-based machines opens up a completely new vista in the field of Human-Computer Interaction. The machines require both a way of 'observing' the gestures and a method for interpreting them. Most automated systems 'see' using video cameras that mimic our eyes. The reason that gesture recognition schemes have not even come close to matching the human performance is the enormous power of an organ that has always intrigued neurologists, philosophers and the layman—the good old human brain.

"Consider a world where simple, everyday tasks would not require grappling with levers, squinting over control panels or hunting for the remote control."

TO SEE OR NOT TO SEE? HOW, IS THE QUESTION

The traditional approach therefore has not focused on understanding and replicating human recognition systems. Instead, people have tried to develop application-specific systems. For example, Becker and Pentland in 1996 developed a system that can recognize 18 T'ai Chi gestures. They use a stochastic process, where the outcome is governed by probabilistic rules known as the Hidden Markov Model. For more information on the T'ai Chi system, see Becker and Pentland, 1996.

In 1994, William Freeman and Michal Roth from Mitsubishi Electric Research Labs developed a static gesture recognition system using orientation histogram matching (OHM). The approach involves calculating a gradient for all pixels as a measure of how the luminosity varies around every pixel. The gradients are used to generate an orientation histogram vector (OHV) for a gesture. Freeman and Roth train gesture models to their system by finding the average OHV from multiple samples of a gesture. When run on a test image, the system finds its OHV and computes its Euclidean distance from each of its stored model OHVs. The test gesture is declared to be of the same type as the model OHV that gives the smallest distance.

In our lab, George Panotopoulos is interested in using dynamic gestures as biometrics. A biometric is a feature that can be used to identify people—a thumbprint, the iris and so on. Earlier in the year, Naru Sundar, an undergrad at Caltech, had modified Freeman and Roth's algorithm to use on a database of gesture samples gathered by Panotopoulos. The aim was to train the system with samples well enough that it would be able to distinguish the gesturer from the gesture made. He concluded that the algorithm could not perform recognition accurately enough to use gestures as a biometric.

My project took some inspiration from Sundar's, and was aimed at determining if we could develop a system that would correctly recognize and classify different types of gestures in Panotopoulos' database. Thus, the problem was not "who is the gesturer?" but rather "what is the gesture?" The system was to learn a model for a gesture type from manually provided samples. However, it would run its probabilistic recognition algorithm without any human supervision. The algorithm powering the system was to be the constellation model.

The constellation model had been developed by Markus Weber, a graduate student at Dr. Pietro Perona's (Caltech EE) lab, for distinguishing between different types of objects, e.g. faces vs. cars in 2-D images. When trained on multiple images of a certain object, for example on human faces, Weber's algorithm learned a constellation model for the object type—a model based on the arrangement of important features, such as the eyes and nose for human faces, and how they co-vary spatially as shown in Figure 1. For any test image, the algorithm would extract the features from the image and compute a probabilistic score using all its learnt models. The highest scoring model was declared a match.

"The system was to learn a model for a gesture type from manually provided samples. However, it would run its probabilistic recognition algorithm without any human supervision."

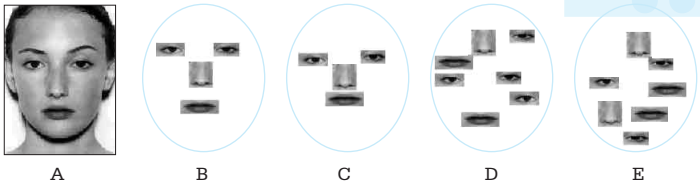


FIGURE 1. (A) A typical face. (B) A good constellation model for a face. (C) Another good constellation model of a face. (D) A bad constellation model for a face. (E) Another bad constellation for a face.

The constellation model has two major advantages. The models it generates are translationally invariant. Whether a face appears in the upper right or lower left corner of an image, it will be correctly identified, just as we can correctly identify the constellation of Orion irrespective of where it lies in the night sky. Secondly, it is quite resilient to background clutter. Objects in the background do not affect its performance significantly, which means that it can recognize an object type against many different backgrounds.

WHAT YOU SEE IS WHAT YOU GET

The system we developed can be described briefly by the block diagram in Figure 2. Effectively, each gesture sample datum is processed (see Figure 3) and its features are extracted. The features can then be used to learn a new model (in conjunction with many more samples of the same kind) or can be classified as one of the models the system has learnt.

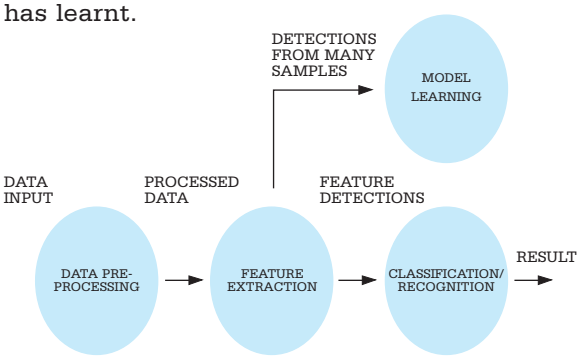
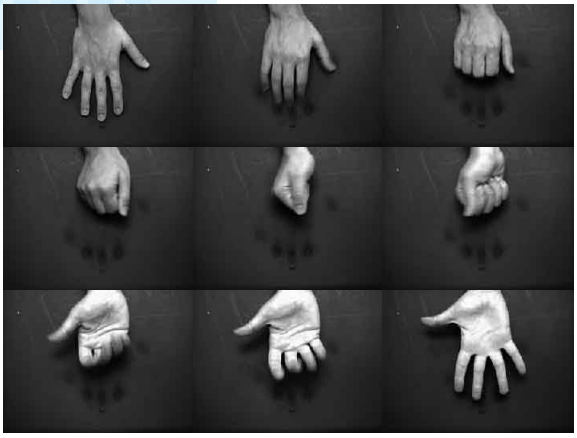
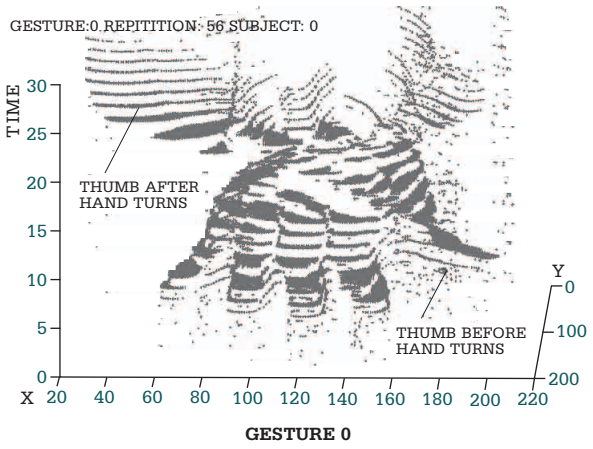


FIGURE 2. Gesture Recognition system block diagram

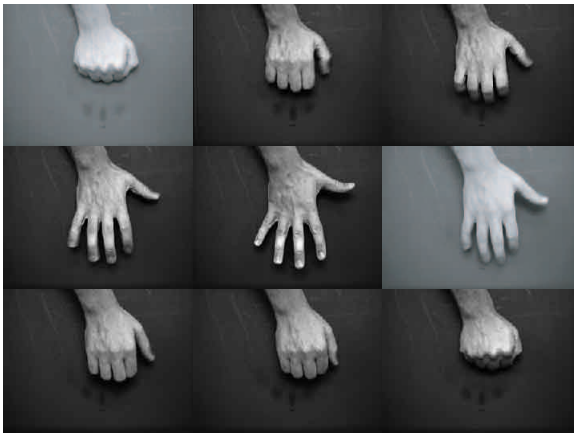
The project involved using the constellation model in very different conditions from Weber's experiments. Firstly, Weber had used the constellation model to distinguish between significantly different objects like cars and faces. All our gestures involved tracking a moving hand. As the hand motion plots in Figure 3 demonstrate, the 'objects' differed only in subtle ways. The similarity across gestures was further enhanced when we processed the gesture data. Regions where hand movement was recorded between consecutive frames were colored 'blue' and everything else was 'white'. As a result, a lot of the surface information, like the positions of the knuckles, was discarded. The aim was to concentrate on the motion of the silhouette of the hand, which we considered important in gesture tracking, to speed up the computation.



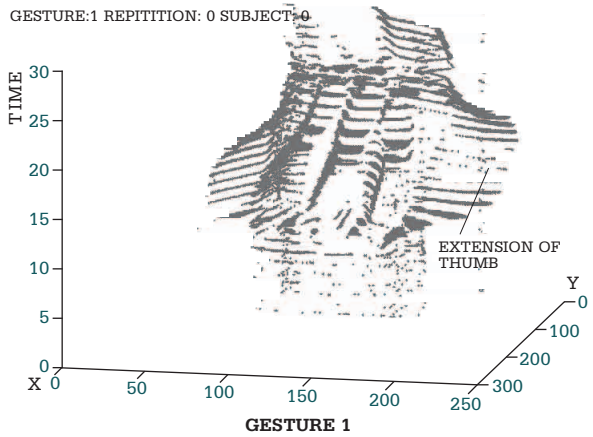
A SELECT FRAMES FROM GESTURE 0



B 3D MOTION OF GESTURE 0 SHOWN IN A



C SELECT FRAMES FROM GESTURE 1



D 3D MOTION PLOT OF GESTURE 1 SHOWN IN C

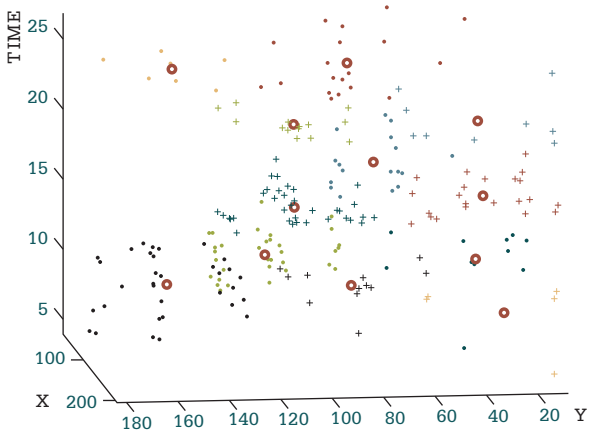
FIGURE 3. Visualization of the data set and hand motion. Figures (A) and (B) contain the frames and the 3D motion plot of a sample of Gesture 0 respectively. Figures (C) and (D) contain the frames and 3D motion plot of a sample of Gesture 1 respectively.

Secondly, every sample of gesture data involved many successive 2-D image frames of a hand in motion. As a result, our data was inherently 3-D, not 2-D. In fact, Figures 3(B) and (D) have been generated by stacking these frames one on top of each other. Extracting features from these samples required extending Forstner's 2-D Interest Operators, which Weber used to detect corners and circular object in images, to run in 3-D. Thus, the program could locate vertices of polyhedra and spherically symmetric features in 3-D and retain their coordinates.

During the training phase, the system extracted and separated vertex and spherically symmetric feature coordinates from multiple training samples (supplied manually). The algorithm, however, did not know what part of the hand exactly generated a given detection. To locate salient features for models, we had to verify if the detections clustered i.e. if they fell into certain groups or clumps. Each cluster would then correspond to detections of a distinct feature. The mean value of every cluster was recorded as the feature's coordinate in the constellation model. The model also stored the feature covariances—a measure of how the different features vary together. If no clustering occurred, we would have to provide better training samples. A successful sample clustering of feature coordinates is presented in Figure 4.

Figure 5 shows an extracted model superimposed on a motion plot of the actual data. As expected, many features lie close to the fingertips and the joints between fingers. Some features seem to appear in the middle possibly because the particular gesture type involved rotating the hand 180°.

A CLUSTERING FOR CORNERS FOUND IN SUBJECT 0 GESTURE 0 REPITITIONS 1,22,45,76



B CLUSTERING FOR CIRCULAR FEATURES FOUND IN SUBJECT 0 GESTURE 0 REPITITIONS 1,22,45,76

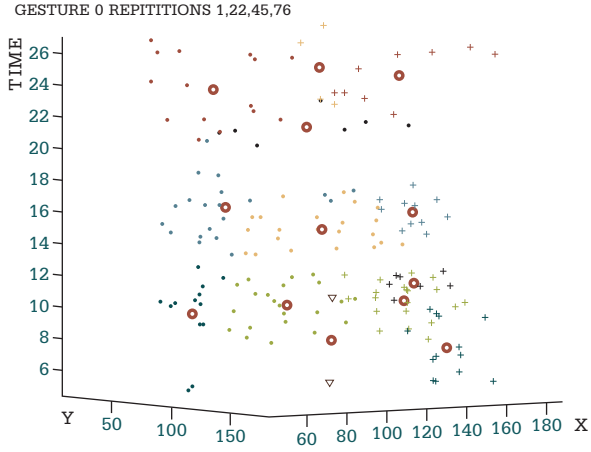


FIGURE 4. (A) Vertex feature detection clustering over extractions from 4 samples of the same gesture. Here the detections plotted in different colors belong to different clusters. The means are plotted in red circles. (B) Spherically symmetric feature detection clustering over extractions from 4 samples of the same gesture. Here the detections plotted in different colors belong to different clusters. The means are plotted in red circles.

CLUSTERING FOR CORNERS FOUND IN SUBJECT 0 GESTURE 0 REPITITIONS 1,22,45,76

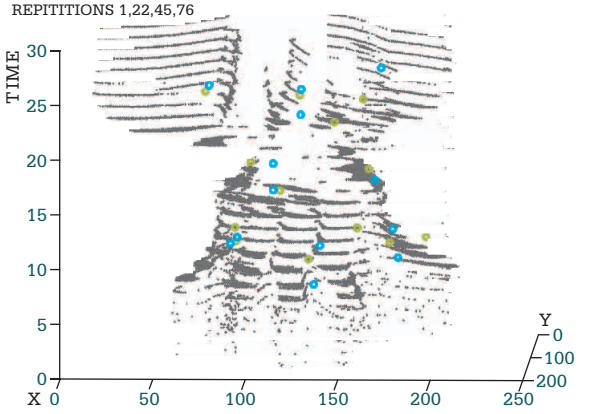


FIGURE 5. Constellation model superimposed on the gesture data. Here the green circles represent the vertex features and the cyan circles represent the spherically symmetric features.

"Every step of the system relies on certain parameters,

Once the model extraction was complete, the system was ready to classify any test input. The features from the test input were extracted and assigned a probabilistic score using every learnt gesture model. The model with the highest score was declared a match—a winner take all situation. Unlike Weber's system that looked for relatively complex features, our system looked for the most banal features—vertices and spherical symmetry—and stored their coordinates. As a result, we had significantly more features than in Weber's experiments. Thus computing the best probabilistic score per model exhaustively would have required at least half a million matrix manipulations and calculations. This was simply unacceptable. We therefore naively matched each detection to the closest feature in the model and computed a score. In principle, this compromise to speed up the recognition process could significantly cancel out the advantage of the constellation model, and our results seem to confirm this suspicion.

We extracted models under nine different parameter conditions and tested the systems with 320 samples (160 samples of two different types of gestures). At best, the system managed to correctly classify 192, or 60.0%, of the samples, under relatively stringent conditions. At worst, it only classified 104, or 32.5%, of them correctly. This is quite unsatisfactory. To put the figure in perspective, a simple coin flip algorithm deciding between two types of gestures should have achieved 50% accuracy in the long run. Thus, at best the system does only 10% better than a penny would!

LIFTING THE VEIL

The system is certainly a long way from home. As explained, the system outperforms the most naive of probabilistic 'algorithms'—coin flip-

some of which significantly affect the results."

ping—by very little. Currently it is not running very fast either, taking around two minutes to classify one gesture type. This is too slow for a real-time system. But is it totally sunk?

We do not believe so. The current system is the first implementation. We now have a complete chain that can process data captured from a video camera, learn new gesture models and then classify them based on a very simplified version of the constellation model. However, nothing in the system has been optimized sufficiently, either because of the numerous degrees of freedom, or due to lack of computational resources. Also, as explained above, the score computation is too naive. Although it seems that with our approach an exhaustive score computation would not be feasible, a further look into the heuristics for better score computation would certainly help.

Every step of the system relies on certain parameters, some of which significantly affect the results—for example, the number of parts the actual data is split into during the search for features directly affects the number of features located. Though there probably is an optimum number of such parts for the image to be split into, we currently know very little about this. The model extraction is carried out under different parameter values. Currently, we use the same parameter values for all types of gestures. Across gesture types however, the gestures are often inherently different and would more likely respond better to different parameter values. Again, though, we have no a priori information as to what these better values would be. Last but not least, the experiments were run in Matlab. Though convenient, this certainly slows the execution manyfold, thus putting another cap on the computational resources of our system. It is therefore still too early to declare whether the above system is truly worth something or doomed to failure.

In a much more encompassing sense, this reflects the current state of the field of computer vision. It may turn out that what we expect to achieve is physically not possible, at least with the current levels of computational technology. However, endowing machines with perception capability as sophisticated as gesture recognition can have immense technological returns. If a system that equals the human brain at gesture recognition were ever implemented, it would provide an unprecedented breakthrough in the field of robotics. In the real world, applications abound. Gesture recognition systems can be used for much more intuitive, rapid interaction with machines. Some versions, especially electronic glove-based ones, are already being used in virtual environments, such as 3-D modeling of molecules. Gesture recognizers can provide a great interface for the vocally challenged by responding efficiently to sign language. Furthermore, we might even be able to create a system that responds to voice and gesture cues, making it even more effective at interacting with humans. Gesture recognition systems therefore seem an integral part of the human-machine interfaces to come. C

Dinkar Gupta is a third year undergraduate in Electrical and Computer Engineering at the California Institute of Technology. *This work was completed with Demetri Psaltis, Thomas G. Myers Professor of Electrical Engineering at Caltech, and funded by the 2000 Caltech Summer Undergraduate Research Fellowship. The author wishes to thank Demetri Psaltis, George Panotopoulos and Pietro Perona.*

FURTHER READING

M. Weber. Thesis. California Institute of Technology ( ).

R. L. Haralick, L. Shapiro. *Computer and Robot Vision II* (Addison-Wesley Publishing Company, New York,).

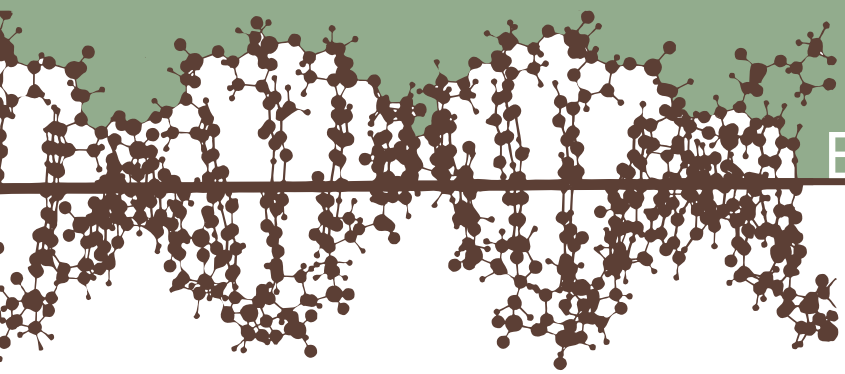
C. M. Bishop. *Neural Networks for Pattern Recognition* (Oxford University Press, London,).

V. I. Pavlovic, R. Sharma, T. S. Huang. Visual Interpretation of Hand Gestures for Human Computer Interaction. *IEEE Transactions on Pattern Analysis and Machine Intelligence* , - ( ).

W. T. Freeman, M. Roth. *Orientation Histograms for Hand Gesture Recognition* (Mitsubishi Electric Research Laboratories, Cambridge, MA,).

P. Perona. 2D Forstner operator implementation, http://www.vision.caltech.edu/ee - /projects/projects.html (link: foerstner.m)

D. A. Becker, A. Pentland. Using a Virtual Environment to Teach Cancer Patients T'ai Chi, Relaxation and Self-Imagery. In *Proc. Int'l Conf. Automatic Face and Gesture Recognition* (Killington, VT,).



EVOLVING ENZYMES FASTER

BY YINGBING WANG

ENZYMES ARE PROTEINS THAT CATALYZE CHEMICAL REACTIONS in living organisms. If exploited, they can be of great benefit to industry and medicine. Unfortunately, these catalysts evolved in nature according to the survival needs of organisms and do not always exhibit the features or efficiency required for industrial application. The research team under Frances Arnold at Caltech applies the methods of directed evolution to alter and improve protein function. One of directed evolution's success stories is the development, by Arnold and Jeffrey Moore, also of Caltech, of an enzyme that is more effective than the natural enzyme at cleaving an ester bond of an antibiotic precursor. This reaction has to take place before the antibiotic can be used. By accelerating the reaction that cleaves the ester bond, the antibiotic can be produced more efficiently. More recently, Hyun Joo and coworkers in the Arnold lab modified an enzyme (cytochrome P450) that requires an expensive (1 gram for $700) cofactor to catalyze reactions. They overcame this problem by evolving an enzyme that could use hydrogen peroxide instead of the cofactor. The modified enzyme functions very effectively with the cheaper substitute.

In directed enzyme evolution, DNA carrying the gene for a specific enzyme is randomly mutated by chemical treatment. Multiple DNA pieces join together to form a vector that can be taken up by the bacteria. This results in a library of mutant genes. The mutated genes are transformed into bacteria and used to produce the individual new proteins. The bacteria are screened for the property of interest. For example, thermostable enzymes can be found by incubating the enzyme-producing bacteria at high temperature. The bacteria with the most active enzymes are selected and the vector containing the mutant gene isolated. With the improved gene, the process begins again. By alternating mutation and selection, enzymes evolve with higher and higher activity. To increase the efficiency of mutant library production, I developed a ligase-free method of DNA cloning that overcame some of the flaws in the original method. I tested the method by trying to improve the thermostability of ∝-amylase enzyme from the hyperthermophilic archaeon *Pyrococcus furiosus.*

"Now in the twenty-first century, we speed up evolution and breed perfect enzymes."

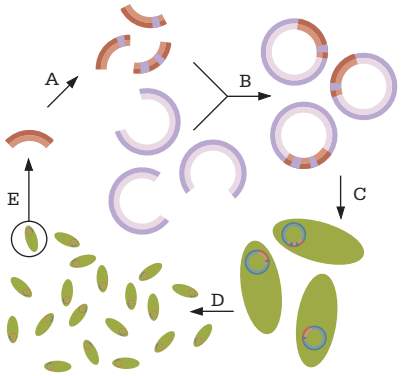


FIGURE 1. Directed enzyme evolution. (A) The gene of the enzyme of interest (red) is replicated and mutated during PCR (mutations in purple). (B) The genes are inserted into plasmid vectors (blue). (C) The vectors are transformed into bacterial cells where the DNA is used to synthesize the individual ∝-amylase variants. (E) The activity of the secreted enzymes is tested by measuring the decomposition of starch. The bacteria producing the most active enzymes are isolated and the process is iterated until the desired enzyme is produced.

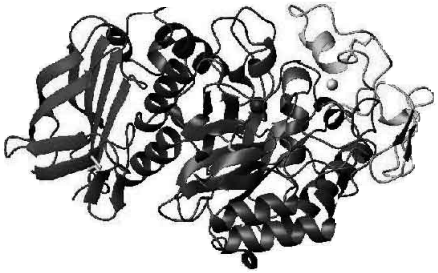


FIGURE 2. The 3D-structure of an ∝-amylase enzyme. It cleaves starch into smaller components, a reaction used industrially for brewing and syrup production. At temperatures higher than those currently used industrially, syrup can be produced with reduced risk of contamination. With thermostable ∝-amylase, the reactions can occur at higher temperature.
Source: http://www.mol.biol.ethz.ch/glockshuber/picturegallery/picturegallery.html

HOW HIGH CAN IT GO?

∝-Amylase cleaves the ∝-1,4-glucosidic bonds of starch, producing ∝-1,4-oligoglucosaccharides and maltose. This reaction is essential in syrup production, brewing, and textiles, making ∝-amylase one of the most commercially valuable enzymes. Starch is water soluble only above 100˚C, which requires that an ∝-amylase used in starch processing be stable and active at that temperature. With more thermostable enzymes, higher temperatures can be used, thereby reducing the risk of contamination.

The traditional DNA cloning technique uses restriction enzymes. Restriction enzymes cut at specific DNA sequences called restriction sites to generate cohesive ends on two DNA fragments, an insert containing the gene of interest and a cut circular piece of DNA called a plasmid vector. DNA ligase links the fragments with covalent bonds and the plasmids are taken up by *E. coli* cells in a process called transformation. Ligase has proven to be a bottleneck in generating large libraries because suitable restriction sites are not always available to combine a specific insert and vector. Another problem with ligase is that DNA fragments flanking the cohesive overhangs attach to each other indiscriminately; inserts bind to other inserts and vector ends to other vector ends. This obstacle reduces transformation efficiency and results in fewer bacteria to select from. The ligase-free DNA cloning technique developed during this project links DNA fragments by creating customized cohesive ends, avoiding the need for ligase.

This technique was recently pioneered by Hung Tseng at the University of Michigan.

The second aim of the project was to explore the thermostability limits of ∝-amylase, one of the most thermostable enzymes known. ∝-Amylase boasts an optimal temperature of 100˚C, a half-life (time to lose half of its activity) of two hours at 120˚C, and is active up to 130˚C. Less thermostable enzymes unfold and denature at much lower temperatures, losing their conformation and consequently their functions. We wanted to evolve ∝-amylase to be active at even higher temperatures.

LIGASE-FREE CLONING

The ligase-free DNA cloning method uses complementary sequences at the junctions of the plasmid vector and the insert to precisely combine the two DNA fragments. First, the plasmid is cut with restriction enzymes. Then, the ∝-amylase gene insert is amplified via Polymerase Chain Reaction (PCR). PCR rapidly copies DNA, each successive round doubling the number of pieces. Specially designed DNA that complements the cut vector ends is added to the ends of each insert copy. These additions provide a long overlapping sequence at the ends of the vector and the insert. Addition of Exonuclease III enzyme (ExoIII) removes nucleotides on double-stranded DNA in the 3' to 5' direction, creating 5' phosphate-end overhangs. The overhangs of the vector and insert fragments bind together because their ends are complementary. Due to the ExoIII digestion, gaps remain where single-stranded DNA does not have a complementary piece to bind to. These gaps are closed by T4 DNA Polymerase. After the process is complete, four nicks remain at the junctions where the nucleotides on each side are not covalently bonded to each other (See Figure 3). The resulting recombinant plasmid solution is used to transform *E. coli* cells.

In order to generate a mutant library, $MnCl_2$ is added to the PCR reaction to induce random errors in the DNA copying. The mutant recombinant plasmids are transformed into

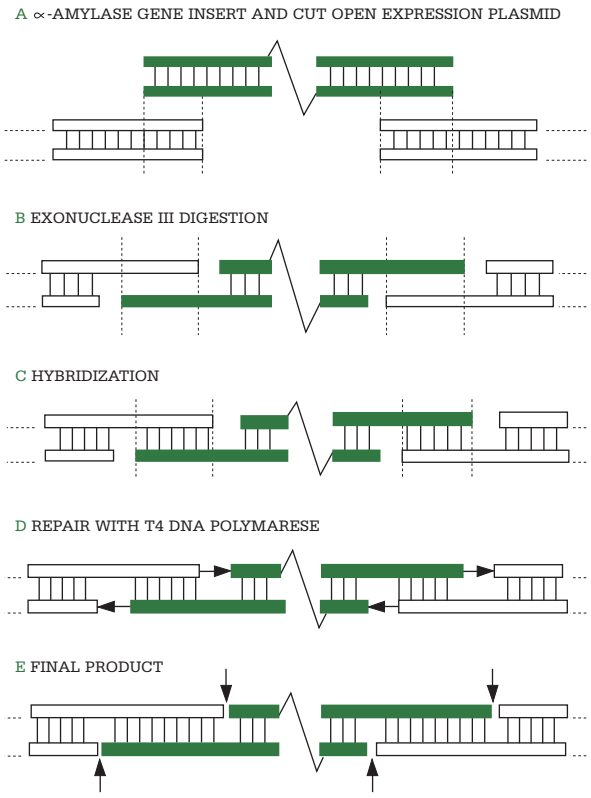
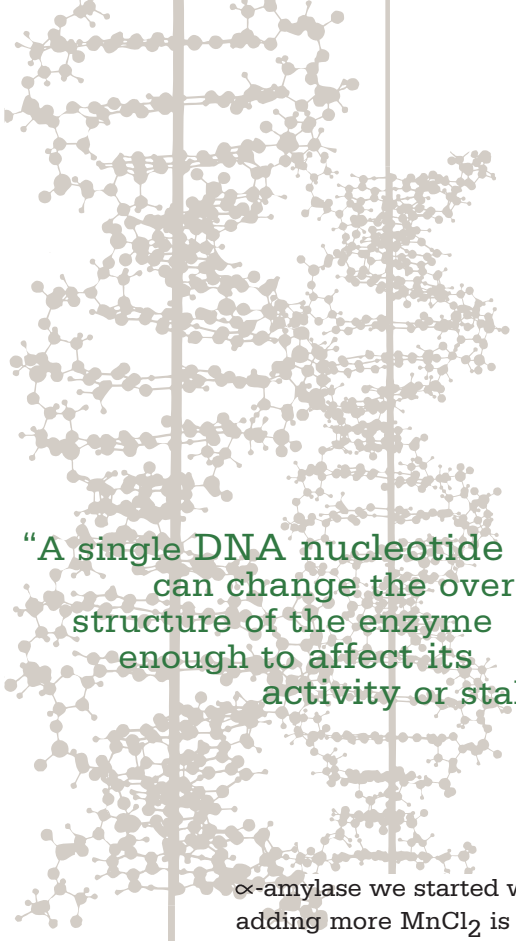


FIGURE 3. Ligase-free DNA cloning procedure. (A) Original cut circular plasmid vector ends (white) and insert containing the gene of interest (green). Dotted lines mark complementary sequences—DNA that is the same on both double helices and thereby has a natural affinity for the other. (B) After brief digestion with Exo III to strip away nucleotides on the 3' ends, short 5' overhangs remain. (C) The vector and insert fragments bind because their remaining nucleotides are complementary. (D) Gaps left after Exo III digestion are closed in the 5' to 3' direction by T4 DNA polymerase. (E) The hybrid regions have been extended and leave nicks (indicated by arrows). The plasmid now contains the gene of interest and is ready to be introduced into the bacterial cells.

expression cells. These cells synthesize the ∝-amylase, and the enzyme is secreted into the surrounding medium. Photometric analysis reveals the amount of amylase.

SEARCHING FOR SUPERACTIVITY

Adding more $MnCl_2$ increases the number of mutations and hence the odds that a particular mutation will change the shape of the enzyme. A single DNA nucleotide error can change the overall structure of the enzyme enough to affect its activity or stability. What we want are mutant clones with activity higher than the



"A single DNA nucleotide error can change the overall structure of the enzyme enough to affect its activity or stability."

∝-amylase we started with. The tradeoff with adding more $MnCl_2$ is that as the number of mutations increases, so will the likelihood that enzymes will not be functional after the changes due to detrimental mutations. Although some of the clones in the error-prone PCR will be superactive, most will have normal activity or no activity at all since the majority of the mutations are likely to be neutral or detrimental. By selection of the rare increased-activity enzymes, the evolution of an enzyme can be directed toward producing an enzyme with the desired characteristics. The project is as yet unfinished, although the ligase-free cloning method appears to be a success. It generates the desired mutant library, but the activity of the resulting enzymes has not yet been carefully analyzed.

The ligase-free cloning technique exhibits many advantages over the traditional method. By generating customized complementary sequences, this method yields more frequent accurate joining between the vector and the insert. With this easy procedure, there is no need to pine over the availability of a restriction site. In addition, because the overhang sequences are custom-designed, symmetry is absent, thus avoiding self-complementarity of the DNA fragments. This reduces the number of unwanted insert-insert and vector-vector joinings, increasing the number of clones that successfully get copies of the ∝-amylase gene.

FUTURE ENZYME ENGINEERING

The search for ∝-amylase that is active at high temperatures continues. Now that we have established an efficient method for generating mutant libraries, we try to identify superactive enzymes. This method can be applied to any type of DNA cloning. It is easily accessible and has the potential for widespread use.

Once superactive enzymes are found, their thermostability is determined by measuring activity after different durations of incubation. The level of activity in each case is compared to the initial activity and the thermal half-life of the enzyme is calculated. DNA that confers improved thermostability is sequenced and the mutations identified. Modified amino acid residues in the proteins are thus pinpointed so that contributions to the thermostability of the enzyme can be evaluated.

Current pursuits in the directed evolution effort include probing the limits of enzyme stability and function in non-natural environments and determining how enzymes evolve novel properties. The soul of directed evolution lies in the ability to synthesize enzymes for adaptation and analysis. As we improve our technology, we will be able to produce enzymes more and more rapidly to suit our needs.

Yingbing Wang is a second year undergraduate in Chemical Engineering at the California Institute of Technology. *This work was completed with Frances Arnold, Dick and Barbara Dickinson Professor of Chemical Engineering and Biochemistry at Caltech, and funded by the Robert E. Anderson Endowment for the 2000 Caltech Summer Undergraduate Research Fellowship. The author wishes to thank Frances Arnold and Holger Berk.*

FURTHER READING

G. Q. Dong, C. Vielle, A. Savchenko, J. G. Zeikus. Cloning, Sequencing, and Expression of the Gene Encoding Extracellular ∝-Amylase from *Pyrococcus Furiosus* and Biochemical Characterization of the Recombinant Enzyme. *Appl. Environ. Microb.* , - ( ).

H. Tseng. DNA Cloning without restriction enzymes and ligase. *Biotechniques* , - ( ).

H. Joo, Z. Lin, F. H. Arnold. Laboratory Evolution of Peroxide-Mediated Cytochrome P450 Hydroxylation. *Nature* , - ( ).

J. Moore, F. H. Arnold. Directed Evolution of a Para-Nitrobenzyl Esterase for Aqueous Organic Solvents. *Nature Biotechnology* , - ( ).

THE BASIS OF ACIDITY

BY PAUL J. CHOI

ORGANIC ACIDS DRAMATICALLY EXHIBIT THE INTERPLAY BETWEEN structure and function. Swivel a bond twenty degrees or push two atoms an Angstrom closer together, and the acidity may increase by a factor of a million. Now change the solvent, and the acidity may increase by another factor of a million. Clearly, anyone wanting to engineer molecules with acidic properties would need to grapple with the problems and grab the opportunities that arise because acids are extremely sensitive to their local environment. If we were, however, able to successfully and efficiently design molecules with target properties, there would be incredible consequences. The periodic table would become a collection of building materials used to make our nanoscale wishes come true. Chemical engineering would take on a new meaning.

Studying organic acids is valuable for several reasons. Attempts to design molecules often focus on catalysts, substances that increase rates of reactions. Catalysts often rely on subtle acid-base equilibria. A biological catalyst may even have organic acids at its active site, the critical region that binds to reactants. Organic acids also provide the chance to explore electrostatic effects that influence the function of complex molecules such as proteins. A charge on the molecule could destabilize a critical part of the molecule, help bind it to another charged ion, or cause a change in conformation, its overall shape. In fact, some of the motivation for this project originally stemmed from detailed work by others on the conformation of succinic acid, a simple dicarboxylic acid with two acidic carboxyl groups.

TWO IS BETTER THAN ONE

The carboxyl group is the functional group of organic acids. The acidity of a single such group is affected to a limited extent by its hydrocarbon base, the long chain of carbons and hydrogens attached to it. Adding a second carboxyl group so that the two groups can interact leads to many more interesting scenarios. Therefore, studying dicarboxylic acids is useful for examining the interaction of carboxyl groups in detail. Varying the structure of the acid will change the possible relative positions of the two groups. Varying the surrounding solvent will help pick out electrostatic effects. For example, the solvent may form a cage around the charges on the molecule.

When a carboxyl group ionizes, it loses a proton, a positively charged hydrogen atom, to another species, and becomes negatively charged. The ionization constants, K_1 and K_2, describe how easily an acid with two exchangeable protons loses its first and second protons to a more basic molecule. For example, if the ionization is easy, K is large, and if the ionization is difficult, K is small. Ionization constants are useful because they allow us to predict the extent to which reactions involving proton transfers proceed.

In 1953, L. Hunter of the University College of Leicester compared the successive ionization constants of fumaric and maleic acids in water. He attributed the increased K_1 and decreased K_2 of maleic acid compared to fumaric acid to intramolecular hydrogen bonding, an attractive force between hydrogen and a nearby atom. Since then, many have studied the interaction of carboxyl groups and their effect on the ratio K_1/K_2.

"If we were, however, able to successfully and efficiently design molecules with target properties, there would be incredible consequences."

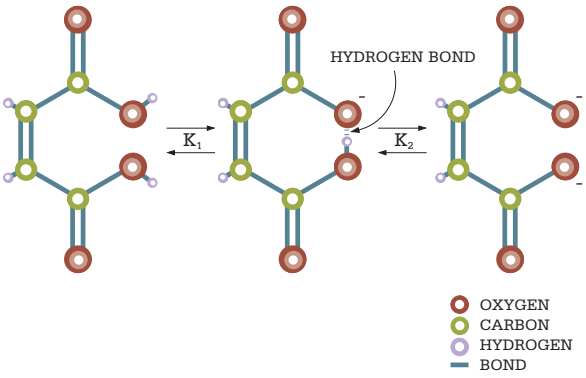


FIGURE 1. Maleic acid ionizes in two steps, losing a proton from a -COOH carboxyl group each time. The ionization constants describe the ease with which each step occurs. After the first ionization, intramolecular hydrogen bonding may be important, changing the values of K_1 and K_2.

CHEMICAL INTERACTIONS

Assuming that the carboxyl groups are identical and do not interact, the ratio K_1/K_2 from statistics alone should be four, since there are two ways for the unionized acid to lose a proton, and there are two ways for the dianion, the ion with two negative charges, to gain a proton. The ratio K_1/K_2 may increase because of intramolecular interactions and electrostatic repulsions. We could think of the unionized, monoanionic, and dianionic acids as corresponding to three evenly spaced rungs on a ladder. From this simplified viewpoint, the ionization constants K_1 and K_2 would be analogous to the ease with which a person could climb up the ladder. Intramolecular hydrogen bonding lowers the middle rung, so the first step is easy, but the second step is difficult. When the dicarboxylic acid is ionized for the first time, there is a negative charge on one carboxyl group. The hydrogen on the other carboxyl group could interact with the negatively charged oxygen, forming a hydrogen bond. The formation of this bond stabilizes the monoanion, because the unfavorable negative charge is spread over a larger region, thereby lessening its effect on any one part of the molecule. As a result, the first ionization is easier and the second ionization is more difficult, increasing the ratio K_1/K_2. Electrostatic repulsions raise the top rung of the ladder, making that last step very difficult. When the dicarboxylic acid loses a second proton, there are two negative charges on the molecule. These charges repel each other and destabilize the dianion. Because the feasibility of a reaction is largely determined by the relative stability of the product, the second ionization is more difficult, lifting the top rung of the ladder and increasing the ratio K_1/K_2. We want to know whether hydrogen bonding or electrostatic effects are more important in moving around the rungs of the ladder. For acids in water, L. Eberson concluded that electrostatic effects would increase K_1/K_2 by at most a thousand times.

It is not immediately obvious how to separate the magnitudes of the intramolecular hydrogen bonding and electrostatic effects. While at Harvard University in 1956, F. Westheimer and O. Benfey proposed a method for quantitatively describing the importance of hydrogen bonding within dicarboxylic acids. They examined dicarboxylic acids and their corresponding monoesters, where the acidity of one carboxyl group was disabled. The monoester's anion should have no hydrogen bonding, so comparing K_1 and K_E, the ionization constant of the monoester's acidic carboxyl group, would give an estimate of the effect of hydrogen bonding on the dicarboxylic acid. Westheimer and Benfey found that hydrogen bonding inside the molecule was insignificant in water, except in a few special cases.

Environment plays a critical role on the relative strengths of the ionization constants. The dielectric constant and hydrogen bonding ability are two characteristics that could affect the ratio K_1/K_2. The dielectric constant describes the ability of the solvent to separate charges. Molecules with multiple charges are more stable in a solvent with a very high dielectric constant, because the charges only interact weakly with each other. In addition, hydrogen bonding between the acid and solvent could compete with internal hydrogen bonding within an acid molecule. Hydrogen bonding ability may be the most important consideration. It has a greater effect on the region immediately surrounding the carboxyl group, because at the atomic scale, solvent molecules cannot always fit between the two carboxyl groups and so separate the charges.

THE GOOD, THE BAD, AND THE BULKY

Dimethyl sulfoxide (DMSO) has unique properties as a solvent. Dimethyl sulfoxide's dielectric constant is lower than water's dielectric constant but higher than that of a pure hydrocarbon. Dimethyl sulfoxide's solvating properties are different from water as well. The oxygen on DMSO can stabilize positive charges in solution, but DMSO lacks an effective way to surround and stabilize anions. The sulfur atom in DMSO is somewhat positively

"The NMR method quickly determines the magnitude of carboxyl interactions with reasonable accuracy."

charged, but any net positive charge is spread out over the large sulfur atom. Also, two methyl groups surround the sulfur atom and make it difficult for an ion to approach.

We chose acids by varying the carbon backbones upon which the carboxyl groups are attached. A carbon backbone composed of only single bonds can rotate about each bond and achieve a variety of positions, much like a sequence of ball-and-socket joints from the human body. Malonic acid has a short but flexible backbone. Diethylmalonic acid is similar to malonic acid but has additional bulky groups at its center that limit the flexibility of the backbone or make a small bubble of low dielectric constant around the acid. A carbon backbone with double bonds is rigid and does not rotate appreciably, similar to a sequence of welded steel rods. Fumaric and maleic acids both have identical, rigid backbones, but with their carboxyl groups apart and together, respectively. Phthalic acid also has a rigid backbone, but a ring of carbons affects the carboxyl group.

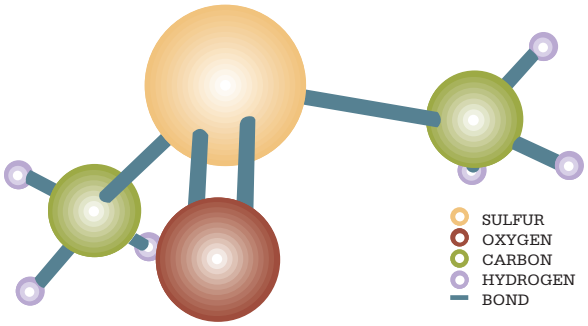


FIGURE 2. Dimethyl sulfoxide is an aprotic, polar solvent. The oxygen atom can interact with positive charges, but the sulfur atom does not interact with negative charges effectively.

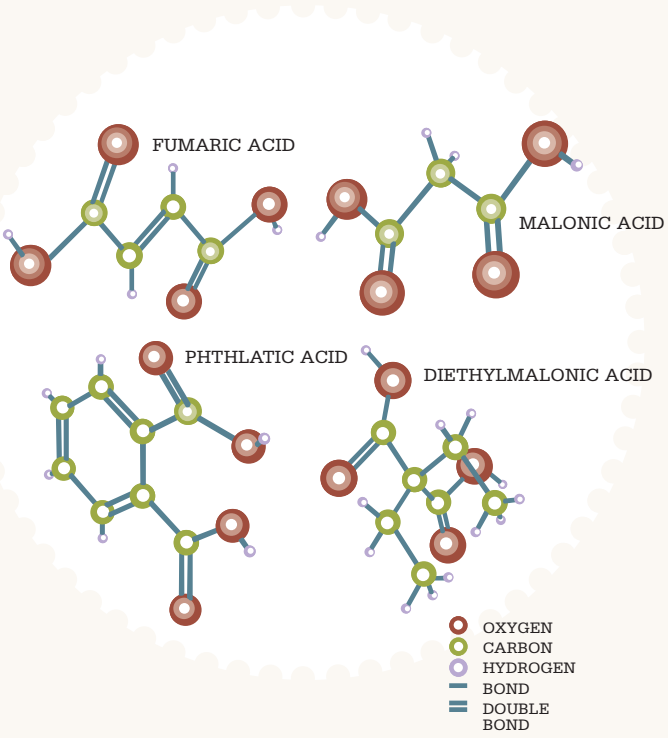


FIGURE 3. Acids with varying properties were examined. In the structures, single bonds can rotate freely, but double bonds cannot. Phthalic, malonic, diethylmalonic, and fumaric acids are shown.

MEASURE FOR MEASURE

Measuring the ionization constants in nonaqueous solvents has proven to be difficult in the past. We used nuclear magnetic resonance (NMR) spectroscopy to determine the ionization constants. Proton NMR detects the local electronic environment, called the chemical shift, of the hydrogen's lone proton nucleus. Since organic acids have many protons, NMR can easily characterize them. We determined the chemical shifts of the unionized, monoanionic, and dianionic acids from prepared standards. From these chemical shifts, we determined the relative amounts of the three species present in a solution of unknown composition using the solution's NMR spectrum.

To measure the ionization constants, we used test acids, which were monoprotic acids with known ionization constants. For example, acetic acid has an ionization constant of 10^{-12} in DMSO. If we add acetic acid to a solution of the maleic monoanion and the final solution contains mostly unionized acetic acid and ionized maleic monoanion, no reaction has occurred, and K_1 for maleic acid must be much greater than 10^{-12}. The NMR method therefore quickly determines the magnitude of carboxyl interactions with reasonable accuracy.

FACTORS OF A BILLION

Comparing the values of K_1/K_2 in water and DMSO demonstrates that reactivity depends heavily on both the type of solvent and the structure of the acid. For fumaric acid, the ratio K_1/K_2 increased by less than a factor of ten going from water to DMSO. Presumably, this effect stems from the lower dielectric constant of the solvent, which causes the two negative charges on the fumaric dianion to interact unfavorably. For the other acids, the ratio K_1/K_2 increased by a factor of a billion going from water to DMSO. The acids appear to be hydrogen bonding when possible, even if the flexible acids must sacrifice their freedom by forcing their two carboxyl groups close together. Only fumaric acid, whose structure forces its carboxyl groups to be far apart, cannot hydrogen bond. These large changes for a variety of acids strongly suggest that intramolecular hydrogen bonding in DMSO is important, though its effect is relatively weak in water. The inability

"A thorough understanding of the interaction between functional groups would be an essential part of our molecule-designing toolkit."

of DMSO to donate a proton for hydrogen bonding with an ionized carboxyl group's negatively-charged oxygen is most likely responsible for the large difference in K_1/K_2 values between water and DMSO.

WHERE TO NOW?

Even among the acids that showed evidence of intramolecular hydrogen bonding, the increase in K_1/K_2 going from water to DMSO varied. Predicting the actual ionization constants is the next step in understanding how the carboxyl groups are interacting. The ionization constants vary over several powers of ten, when a single power of ten can make the difference between an efficient and inefficient reaction. Thus, the specific interactions of the carboxyl groups need to be understood in detail, and this will be the basis of future work. Additional acids, whose ionization constants are unknown, will be studied using the NMR method. Along the way, the NMR method may be further refined to allow a more efficient determination of K_1/K_2 values. We hope to examine dicarboxylic acids that have been predicted to have particularly strong hydrogen bonds in order to understand what in their structure causes such favorable interactions.

After examining many dicarboxylic acids with interesting shapes and acidities, we will come to understand acids well enough to predict, in detail, the interactions between acidic side groups. Once we can do this, we will be in a better position to design molecules that have sensitive requirements for both structure and function, since carboxyl interactions not only affect the acidity of a molecule, but also affect the conformation. The overall conformation of a molecule determines how well carboxyl groups interact, and in turn, the carboxyl interactions further determine what particular shape the molecule favors. Thus, structure and function are truly interdependent. Knowing when a mechanism such as intramolecular hydrogen bonding is important will be necessary to understand the behavior of simpler molecules as well as more complicated ones like enzymes. Additional functional groups, such as amine groups, can also interact with carboxyl groups and need to be studied.

A thorough understanding of the interaction between functional groups would be an essential part of our molecule-designing toolkit. Experimental and theoretical advances continue to provide us with more of the tools we need every year. In the future, we will be able to draw any chemical structure and predict its shape and reactivity to a high degree of accuracy. Then, we would truly be able to engineer molecules to suit our needs. C

Paul Choi is a second year undergraduate in Chemistry at the California Institute of Technology. *This work was completed with John D. Roberts, Institute Professor of Chemistry at Caltech, and funded by the Peter A. Lindstrom Endowment for the 2000 Caltech Summer Undergraduate Research Fellowship. The author wishes to thank John D. Roberts and Krag Petterson.*

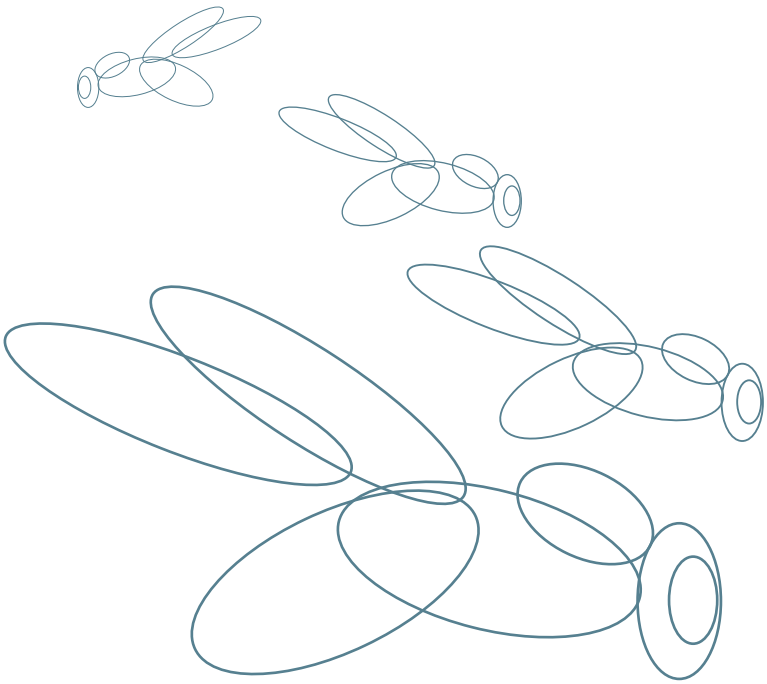
FURTHER READING

M. K. Chantooni, I. M. Kolthoff. Transfer Activity Coefficients of Acids and Ions. *J. Phys. Chem.* , - ( ).

C. L. Perrin, M. A. Fabian. Multicomponent NMR titration for simultaneous measurement of relative pKas. *Anal. Chem.* , - ( ).

F. G. Bordwell. Equilibrium Acidities in Dimethyl Sulfoxide Solution. *Acc. Chem. Res.* , - ( ).

L. L. McCoy. The Geometry of Intramolecular Hydrogen Bonding in 1, 2-Dicarboxylic Acids. *J. Am. Chem. Soc.* , - ( ).



THE AGING ENIGMA : CONSULTING THE FLY

BY ANNEMARIE SELAYA

A MULTITUDE OF GENETIC AND ENVIRONMENTAL FACTORS ARE ABLE to modify the longevity of an organism. The identification of these factors and associated molecular mechanisms is crucial to address the problem of aging and why it is associated with "normal" deteriorations (loss of agility, eyesight, etc.) and higher incidence of pathological conditions (Alzheimer, Parkinson, Huntington diseases, cancer, etc.). By selective breeding, Michael Rose of the University of California, Irvine has shown it is possible to isolate flies with extended longevity demonstrating the existence of a genetic component of aging. However, these long-lived flies carry multiple mutations making further genetic analysis impossible. Although multiple genes contribute to longevity, single gene mutants have been isolated in worms, fruit flies (*Drosophila melanogaster*) and mice. This critical observation predicts that classical genetics with an organism as simple as a fruit fly can be used to dissect a problem as complex as aging.

"Classical genetics with an organism as simple as a fruit fly can be used to dissect a problem as complex as aging."

Although this link has been established between aging and genetic elements, the exact interplay has yet to be determined. It remains unclear whether a specific genetic program controls aging, whether aging controls the genes, or both. Within the scope of that question lies another: if aging is a by-product of genetic elements, is aging equivalent to a disease to which the genes must slowly succumb, or do genes induce aging in an orderly program, making them our own worst enemies?

Direct investigation of how genetic and molecular mechanisms contribute to the aging process hinges on the establishment of an accurate description, termed phenotype by geneticists. Prior use of mortality as an endpoint indicator of aging is blind to the many events that contribute to that eventuality. It is crucial to identify what specific changes occur as we age. The majority of people can easily describe the external changes associated with aging, such as white hair and wrinkles. However, white hair or wrinkles cannot be used to pinpoint a specific age. The advent of molecular and genetic technology provides a powerful "scalpel" to dissect internal changes that could be used as a specific phenotype for every phase of the life span.

The enhancer trap technique has proven to be one such technology that can be used to indirectly visualize gene expression patterns. Using this approach, Blanka Rogina from the University of Connecticut School of Medicine demonstrated how gene expression in the adult fly antenna changed over the life span of the animal. Gene expression patterns were classified based on general trend of expression, i.e. genes that started expressing from an initial high level and decreased later in adult life, genes that started expressing from an initial low level and increased later in adult life, and classes of genes which showed constant expression throughout adult life. He also showed that gene expression changes can be used as biomarkers of aging, i.e. alteration of longevity by mutations or environment does not affect the pattern of change, but only the rate at which they occur. This suggests that a well-controlled and orderly gene expression progression takes place during aging. Thus, gene regulation probably remains active and dynamic during adult life. However, those studies focused solely on the fly antenna. Studying only that specific portion of the fly may have masked important genes that are active in the rest of the fly.

Employing the same enhancer-trap technique, 500 genes were examined in our lab for when, where and how much they were expressed. Several of these genes, such as one we called *jumpy*, exhibit age specific expression changes.

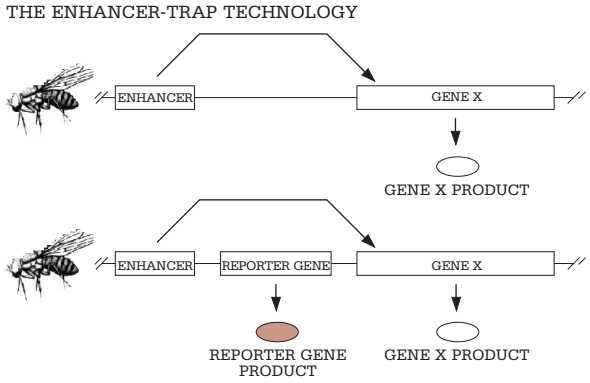


FIGURE 1. In a normal fly (top fly), gene X expression is controlled by a regulatory element called the enhancer. In an enhancer-trap fly (bottom fly), an engineered piece of DNA including a reporter gene is inserted near the gene X enhancer. The reporter is thus also controlled by the enhancer and mimics the expression pattern of gene X. We can measure the amount of the reporter gene product to monitor the amount of the gene X product.

Using *jumpy*, we have started to establish the causal link between gene expression and aging. This relies upon the characterization of the molecular changes reflected by variation in gene expression.

A FLY CALLED *JUMPY*

The *jumpy* enhancer-trap has a complex temporal pattern of expression, increasing up to midlife and subsequently decreasing. Examination of the expression localization within the fly shows it is restricted to muscle tissue. These observations indicate that muscles are subjected to specific molecular changes during aging.

JUMPY TEMPORAL PATTERN OF EXPRESSION

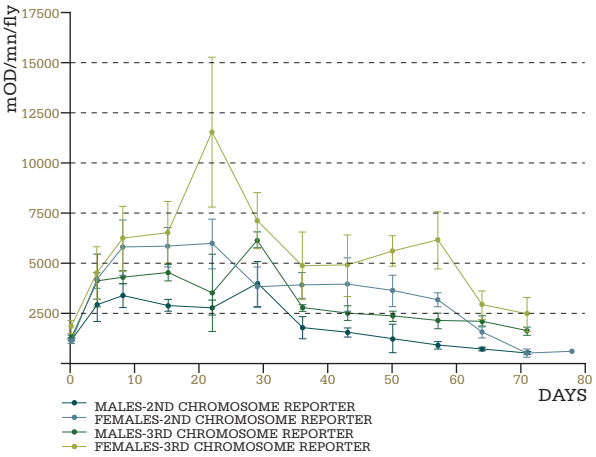


FIGURE 2. The *jumpy* gene is transcribed at varying levels throughout a fly's life.

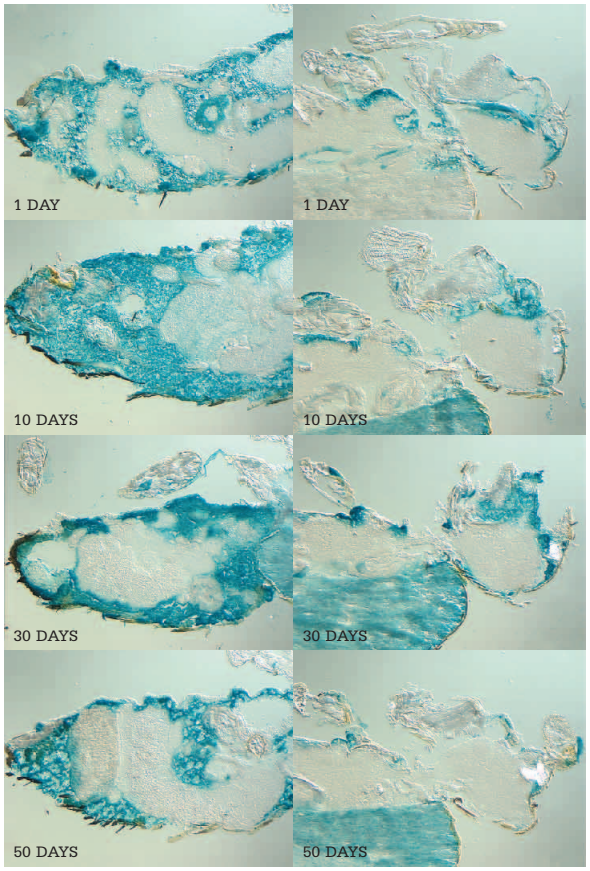


FIGURE 3. Dark blue areas on these whole fly sections reflect tissue in which the *jumpy* gene is expressed. Left panels show abdominal regions while head and thorax regions are on the right panels. For each panel, the age of the sectioned fly is indicated. *Jumpy* enhancer-trap expression is restricted to muscle tissues.

JUMPY ENCODES A TYROSINE PHOSPHATASE ENZYME

Analysis of the sequence surrounding the *jumpy* enhancer-trap revealed its location inside the LD39930 gene (previously identified by the Berkeley Drosophila Genome Project). This gene encodes for a putative tyrosine phosphatase, an enzyme that removes a phosphate group from a substrate by hydrolysis. Since we found that *jumpy* expression is restricted to muscle tissue, this enzyme must act on a substrate within muscles.

JUMPING TO HUMANS

A GenBank search identified the human version to this phosphatase (GenBank accession number: AY007098, p-value=2e-71). The fly and human *jumpy* proteins appear to be related to the myotubular myopathy (MTM) family, a group of phosphatases. Jocelyn Laporte and Ling Jia Hu, CNRS Strasbourg, France, and Petra Kioschis, Heidelberg, Germany, have shown mutations in the human MTM1 gene are associated with myotubular myopathy, a severe muscle disorder characterized by muscle weakness and severe hypotonia. Since the *jumpy* and MTM proteins only share the tyrosine phosphatase active site consensus sequence (I/V)HCXAGXXR(S/T)G and a FSG motif, *jumpy* defines a new family of muscle phosphatases. This finding raises the question of the biological function of the *jumpy* family. Removing the gene and examining the consequences enables us to resolve this issue. Obviously, such experimental proceeding is not suitable with humans.

JUMPING BACK TO FLIES

The *jumpy* enhancer-trap contains a reporter gene insertion in the *jumpy* gene that may renders it nonfunctional. By a simple genetic cross, it is easy to mutate the fly *jumpy* gene and analyze the effects.

"[*Jumpy*] males and females have a shorter life span compared to normal flies."

We tested flies carrying a homozygous *jumpy* insertion for phenotypes relevant to aging, i.e. lifespan, stress resistance and metabolic rate. Lifespan analysis shows that *jumpy* males and females have a shorter lifespan compared to normal flies.

Jumpy flies also show a reduced resistance to stress induced by free radical generator (paraquat) or starvation. However, the metabolic rate, as measured by CO_2 output, does not show any statistically significant differences with the wild type control until the flies get close to death. This result indicates that the decrease in lifespan cannot be attributed to higher metabolic rate.

In the youthful flies, age 6 days, *jumpy* has a randomly occurring *jumpy* behavior. However by banging the fly vial, they consistently show a hyperactive reaction and difficulties in stopping buzzing of their wings after landing. At 50 days, *jumpy* show dramatic differences in activity level. The rate of movement is strikingly slower as compared to white flies, and their bodies and legs appear to shake. Notably, when *jumpy* flies are dismantled from the surface of the glass container, they fall upon their backs and exhibit difficulty in righting themselves, often struggling frantically for many seconds. At that same age, the white flies instantly regain upright posture when dismantled. Interestingly, it appears the various defects appear progressively stronger, longer and more frequently with age. *Jumpy* males appear to be less mobile, moving a shorter distance in five minutes as compared to the white male controls. Although the absolute distance moved changes at various ages, with a decrease thereafter as the flies age, *jumpy* males consistently move less.

Taken together, these observations are reminiscent of muscle disorder symptoms, suggesting that, as the MTM family, the *jumpy* family is required for muscle function.

JUMPY MUSCLE DISORDER

To test this hypothesis, we investigated the origin of defects by examining *jumpy* muscles. The sagittal sections of the thoracic region show no significant difference in organization and structure of muscle fiber in *jumpy* and control white flies. However at the electron microscope level, individual *jumpy* sarcomeres have a darker M line, located in the A zone, than white sarcomeres.

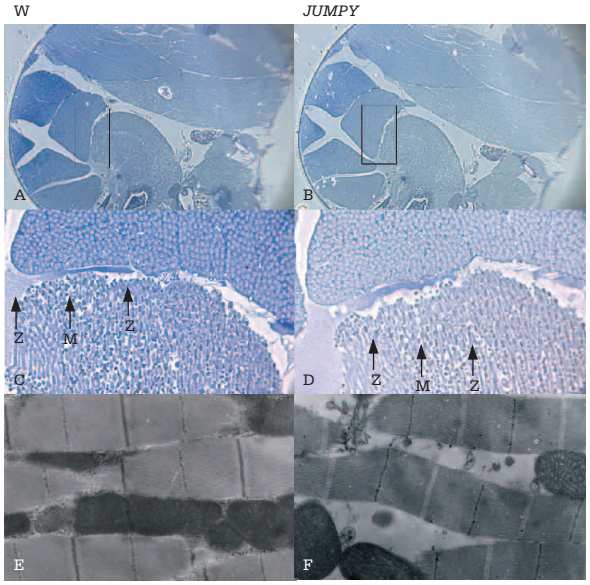


FIGURE 4. Thoracic muscle sections revealed by light (A, B, C, D) and electron (E, F) micrography. (A), (C), and (E) show control fly muscle tissue and (B), (D), and (F) show *jumpy* muscle. (C) and (D) are enlargements of muscle fiber cross-sections in (A) and (B). No significant difference in muscle fiber structure is revealed in these sections. (E) and (F) show longitudinal sections of muscle tissue. The M arrows point to the M-line in individual sarcomeres. The M-line is lighter in *jumpy* than control. The Z arrows point to the Z-line, the endpoint to each muscle sarcomere.

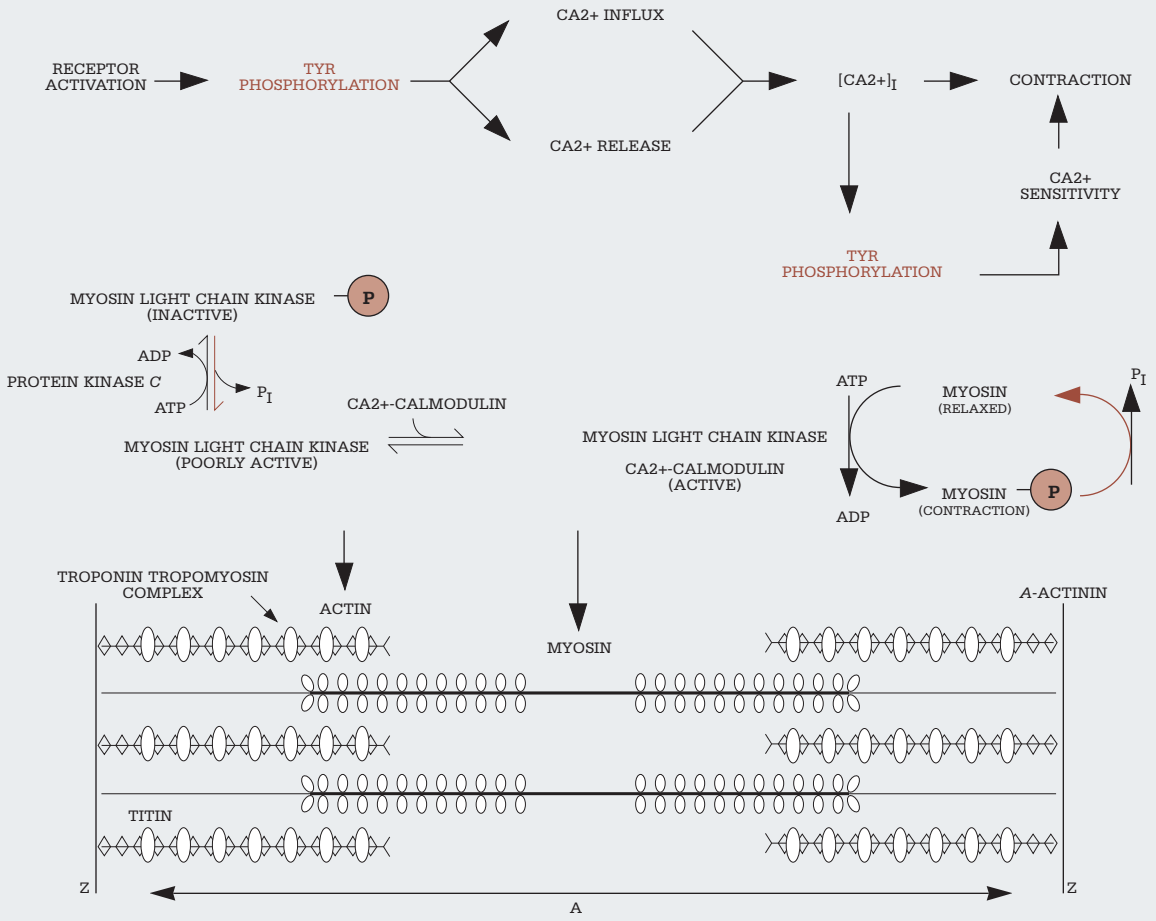


FIGURE 5. Sarcomere structure and models of muscle contraction. Sarcomere structure is illustrated in the bottom panel, showing the A-Zone (myosin filaments) and Z-line (marks the end of actin filaments and individual sarcomeres). Interactions between the myosin and actin lead to muscle contraction. The top panel outlines possible regulatory mechanisms occurring in actin filaments. Protein tyrosine phosphorylation participates in mechanisms occurring in actin filaments. Protein tyrosine phosphorylation participates in mechanisms changing cellular calcium concentration. This increase can induce further tyrosine phsophorylation, enhancing calcium sensitivity in actin-myosin reactions. The middle panel models the role of myosin light chain kinase (MLK) in muscle contraction. Dephosphorylation of MLK allows binding with calicum and calmodulin, activating the kinase. The activated MLK-Ca-Calmodulin complex phosphorylates the myosin filaments, leading to contraction.

The defects affecting behavior and locomotion in *jumpy* are associated with structural abnormalities in the muscle contractile units. Muscular contraction, as described by the sliding filament theory, involves interactions between the myosin filaments and actin, resulting in the shortening of the whole sarcomere. When all sarcomeres in a muscle fiber shorten, muscle contraction occurs. Muscle pathologies (myopathies) characterized in humans involve either shrinkage or degeneration of muscle fibers. *Jumpy* does not exhibit notable defects in the muscle fibers, but rather has differences in the sarcomere, where basic regulation of skeletal muscle contraction occurs (See Figure 5).

Several studies implicate protein tyrosine phosphorylation/dephosphorylation is an important mechanism in regulation of smooth muscle contraction. Specifically, tyrosine kinase-induced phosphorylation has been shown to be involved in the control of stored Ca2+ in rat aortic smooth muscle. Calcium, well known to play an integral role in muscle contraction, signals the regulatory proteins on actin (troponin and tropomyosin) to reveal the myosin binding sites, and hence allows for actin-myosin interactions.

Other regulatory mechanisms in muscle

"This gene's function in muscle contraction may play a critical role in the aging process."

involve phosphorylation as well. Myosin light chain kinase (MLK) is known to contribute to smooth muscle contraction when activated by binding to a calcium-calmodulin complex. The active MLK subsequently phosphorylates the myosin light chain, allowing for contraction. A phosphatase is then required to dephosphorylate the myosin light chain, relaxing the muscle.

Both models illustrate the complex regulatory mechanisms in muscle, and their reliance on an intricate series of phosphorylations and dephosphorylations. Although they were proposed for smooth muscle, Perry *et al*. admit that the phosphorylation of the myosin light chain must modulate the contractile response, which is important in fast muscle (such as flight muscle in flies). Although the specific function of the *jumpy* tyrosine phosphatase has not been established, it could fill the many niches of muscle regulation.

A FLY MODEL OF MUSCLE DISORDERS?

Sequential homology relates the *jumpy* tyrosine phosphatase to a well conserved protein family thought to be responsible for myotubular myopathy, a human muscle disorder characterized by muscle weakness and severe hypotonia. However, the *jumpy* protein is much more closely related to its human homolog, identified and described for the first time in this study. The *jumpy* human homolog has an unknown function and is found at chromosomal location 3p25. Interestingly, the gene encoding caveolin-3 has also been mapped to 3p25. Mutations in caveolin-3 are believed to cause a form of muscular dystrophy in humans. Our study prompts for human studies to investigate the presence of mutations in the human *jumpy* gene associated with muscular dystrophy or other muscular disorders of unknown etiology.

MUSCLE AND AGING

We propose that *jumpy* decrease of expression at older age might be responsible for locomotion and movement difficulties. It follows that this gene's function in muscle contraction may play a critical role in the aging process. The phenotype expressed by the mutation of this gene may represent a hastening of a normal step in aging, leading to a decrease in lifespan. C

Annemarie Selaya is a third year undergraduate in Ecology, Behavior and Evolution at the University of California, Los Angeles. *This work was completed with Seymour Benzer, James G. Boswell Professor of Neuroscience at the California Institute of Technology, and funded by the Howard Hughes Medical Institute and the Minority Undergraduate Research Fellowship. The author wishes to thank Seymour Benzer, Laurent Seroude, Rosalind Young, and Amparo Gomez.*

FURTHER READING

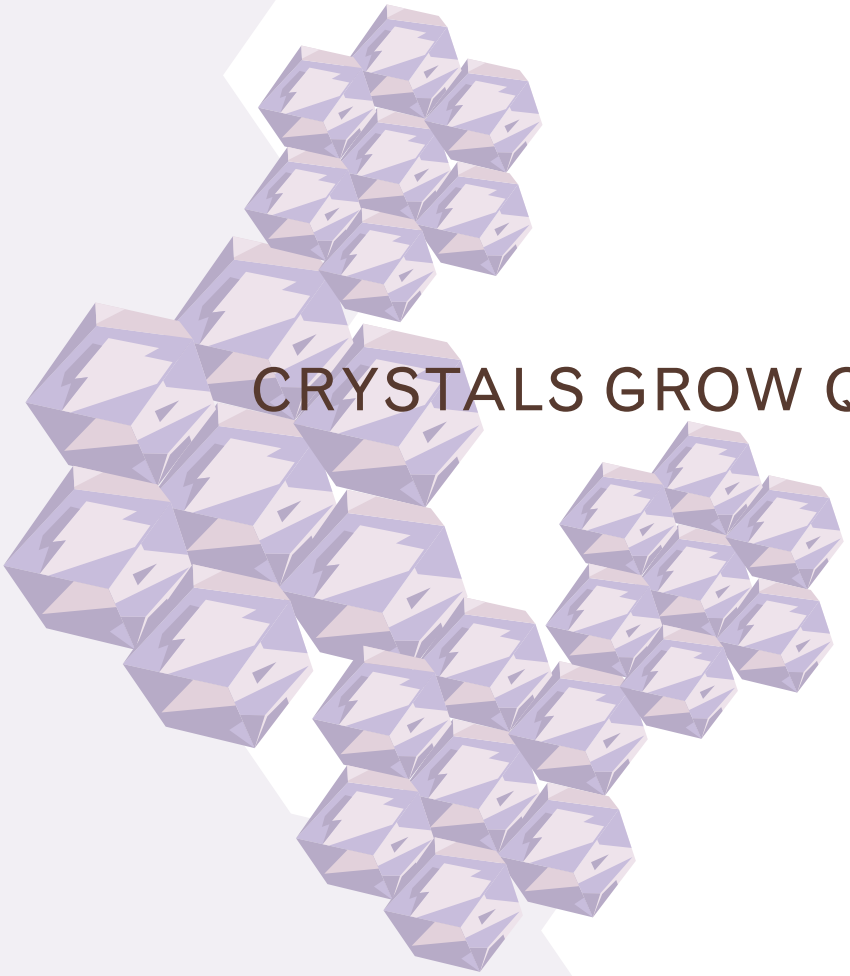
E. H. Fischer, H. Charbonneau, N. K. Tonks. Protein Tyrosine Phosphatases: a Diverse Family of Intracellular and Transmembrane Enzymes. *Science* , - ( ).

S. L. Helfand *et al*. Temporal Patterns of Gene Expression in the Antenna of the Adult *Drosophila Melanogaster*. *Genetics* , - ( ).

H. Huxley. The Mechanism of Muscular Contraction. *Scientific American* ( ), - ( ).

D. Jones, J. Round. *Skeletal Muscle in Health and Diseases* (Manchester University Press, New York,).

J. Laporte *et al*. A Gene Mutated in X-linked Myotubular Myopathy Defines a New Putative Tyrosine Phosphatase Family Conserved in Yeast. *Nature Genetics* , - ( ).



CRYSTALS GROW QUICKER IN E-FIELDS

BY MOLLY SWANSON

CRYSTALS HAVE LONG BEEN PRIZED for both their beauty and their utility, but the interest scientists have for them has little to do with their aesthetic value. The mechanics of crystal growth present an incredibly complicated system that has perplexed scientists for centuries. Much theoretical work has been done on the mechanisms involved in crystal growth, but several aspects are still unexplained. The equations describing crystal growth are very complicated, and have yet to be fully solved, while there are effects seen in experiments that the theories can't explain. Additionally, the behavior of highly faceted crystals is not well understood, because the highly complex shapes render the theoretical calculations prohibitively difficult to solve. This is unfortunate, as crystals have many industrial applications, including solid-state lasers, flat screen computer displays, and a wide variety of other devices, so information about crystal growth could potentially bear great dividends. But beyond these tangible benefits studies of crystal growth could bring, there is also the thrill of

uncovering a deeper understanding of these remarkable systems—crystals form some of the most beautiful structures in nature, such as snowflakes and diamonds, and doing these crystal growth experiments can help explain how and why these fascinating systems form.

One of the most interesting directions current research has taken is to examine the effect electric fields have on crystal growth. Studies have shown that interesting things happen when a field is introduced to an ice crystal system—the crystals grow faster and form different shapes, so there is definitely some interesting, new physical effect present. This also provides a way to test some of the crystal growth theories—if the theoretical calculations can be modified to include an electric field, theories can be tested under some different parameters. Also, knowledge about how they grow in an electric field could be used to control crystal growth—if growth direction and rate could be changed with the flip of a switch, applications that require great precision would become much more efficient.

Experiments on crystal growth are done in a sealed tank, with the crystals grown on a wire that is kept at a given voltage, creating an electric field around the growing crystal. In experiments done with ice, the electric field was expected to increase the rate of crystal growth by increasing the rate of diffusion of the polar water molecules. The results showed that when a voltage was applied, the ice crystals formed dendrites whose growth rate did indeed increase as the voltage was increased. They also showed a runaway growth of thin ice needles at a threshold voltage. However, it was later pointed out that the electric field also increased the threshold at which the water condensed (called the equilibrium vapor pressure), which effectively would have canceled out the effect from the polarizability of the water molecules. Some other mechanism, therefore, must have been responsible for the increased growth observed. One possible theory is that a smaller electrostatic effect reduces the equilibrium vapor pressure. The original theory did not take this into account because it is substantially smaller than the polarization effects, but since these cancel out, this electrostatic effect could become important. Theoretically, such an effect would operate like this: adding a molecule to a crystal means that it must get slightly larger, and if the crystal has charge on it, the charges will tend to repel each other and encourage this change. It is by this process that the applied voltage enhances the crystal growth. A natural experiment to test whether this understanding is correct is to examine the effect of an electric field on iodine crystals—being non-polar, there would be no increase in growth due to a polarizability-dependent increase in diffusion rate, so any enhancement would be due to this repelling of like charge.

GROWING THE CRYSTALS

These experiments were done in a 500 mL glass "fleaker," shown in Figure 1. It was suspended upside-down and a copper tube with an insulated wire inside came up through a hole in a rubber stopper at the bottom. The hole was sealed with vacuum grease so the tube could be rotated easily, and the section of tube inside the fleaker was wrapped with heat-shrink tubing for insulation and protection from corrosion. About 2 cm of the wire was stripped and extended beyond the end of the tube, providing a place for crystals to form. The other end of the wire was connected to a high voltage power supply to apply the voltage to the growing crystals. In order to get crystals to form on the wire, the copper tube was chilled with ice water. Both the top and the bottom of the fleaker were wrapped with aluminum foil and heater wire, and a temperature controller was hooked up to the top, which tells the heater to maintain a given setpoint.

To set up the experiment, a tablespoon of iodine was put in the chamber and the bottom heater was turned on, with foil wrapped around the middle of the fleaker to keep

"The greatest crystal growth seemed to be obtained by slowly stepping the voltage up."

crystals off the sides. The iodine crystallized on the top overnight. To actually perform the experiments, ice water was put around the tube and the top heater was turned to 80°C. The bottom heater was turned off during the experiments, so the bottom of the chamber was at room temperature. The iodine diffused down from the top and started crystallizing on the wire very quickly, and continued to grow good crystals for about two hours. After a trial, the bottom could be heated again to recrystallize the iodine on the top. Large crystals also formed on the sides of the chamber during the experiments, but this did not seem to affect the growth much unless the crystals from the wire were growing close to the sides. The crystals on the sides were knocked off periodically so they did not interfere with the growth or obscure the view of the crystals on the wire. There was no method for measuring the concentration of iodine in the air above the equilibrium vapor pressure, which is called the supersaturation, directly, but using computer models of the temperature and vapor pressure distributions with the program QuickField, the supersaturation in the crystal growing region was estimated to be around 6-10 percent.

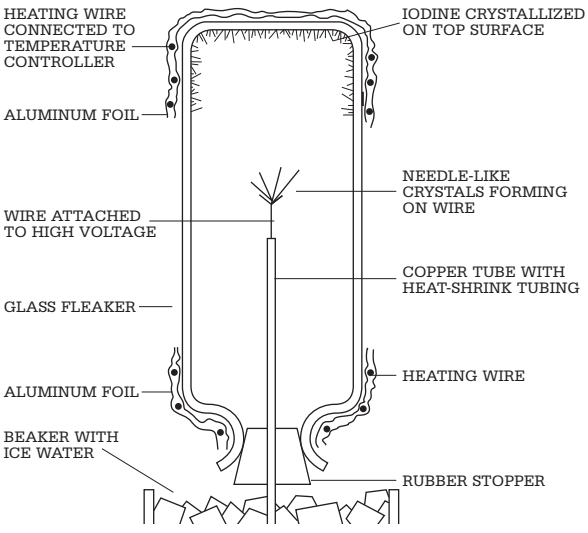
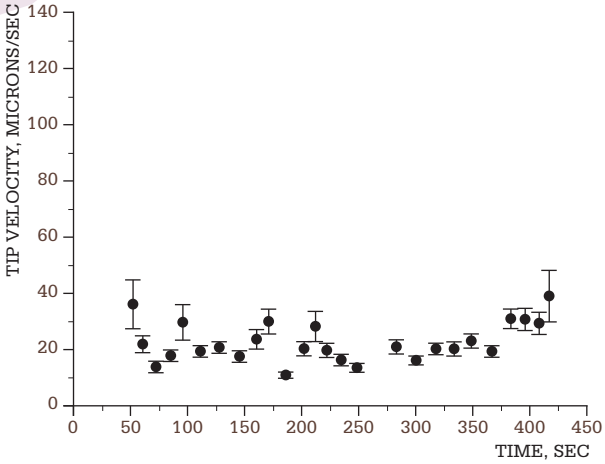


FIGURE 1. A diagram of the experimental setup.

Data was taken by videotaping the crystals using a microscope and a CCD camera. One crystal was isolated to reduce the effects of competition, turned so it was perpendicular to the camera, and followed through a series of voltage changes. During each two hour trial, 8 to 11 crystals were measured. The greatest crystal growth seemed to be obtained by slowly stepping the voltage up, but some data was also taken by stepping it down and some was taken at a constant voltage. The polarity of the voltage was occasionally switched from positive to negative to see if that had any effect. The tape was analyzed using a VCR with a timer, and tick marks were made on the screen with a marker at the tip of the crystal. The time each tick mark was made and the distance from the last tick mark were recorded. The distance was measured with a ruler on the screen and then calibrated. This information was then used to create profiles of the tip velocity versus time for each crystal, noting the times where the voltage was changed. These profiles are shown in Figures 2(A) and (B). For each crystal, the average and standard deviation of all of the velocities at each different voltage were also calculated. Average velocity vs. voltage combining data from six 2-hour trials was plotted, and some of the data was thrown out—points corresponding to crystals noted as competing, crystals growing into the sides of the chamber or other locations that made measurement difficult, forks that stopped growing in favor of the other fork, averages with less than three data points, and averages with a standard deviation of 50% or greater were deleted from the data, since they were not representative. This plot of the average velocities is shown in Figure 3.

A SAMPLE VELOCITY PROFILE OF SINGLE IODINE CRYSTAL—NO VOLTAGE



B SAMPLE VELOCITY PROFILE OF SINGLE IODINE CRYSTAL—VOLTAGE STEPPED UP

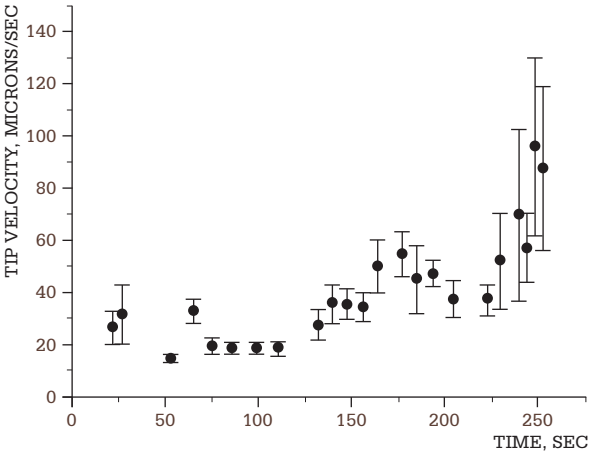


FIGURE 2. Two sample velocity profiles of individual crystals. (A) shows a profile with the voltage turned off, and (B) shows a profile where the voltage was stepped up slowly. Note the increase in velocity as the voltage is turned up.

AVERAGE TIP VELOCITY OF IODINE CRYSTALS AS A FUNCTION OF VOLTAGE

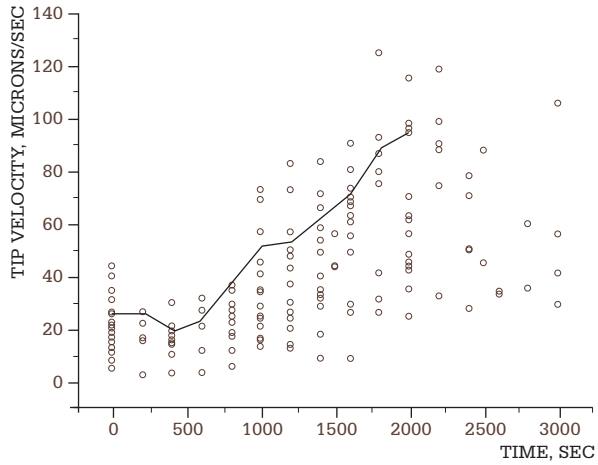


FIGURE 3. A summary of average velocity vs. voltage data. The 80th percentile line represents a hypothetical case in which competition between crystals and supersaturation fluctuations is eliminated.

FIGURE 4. A collection of crystals growing on the wire with the voltage turned off. Note the wide variety of shapes.

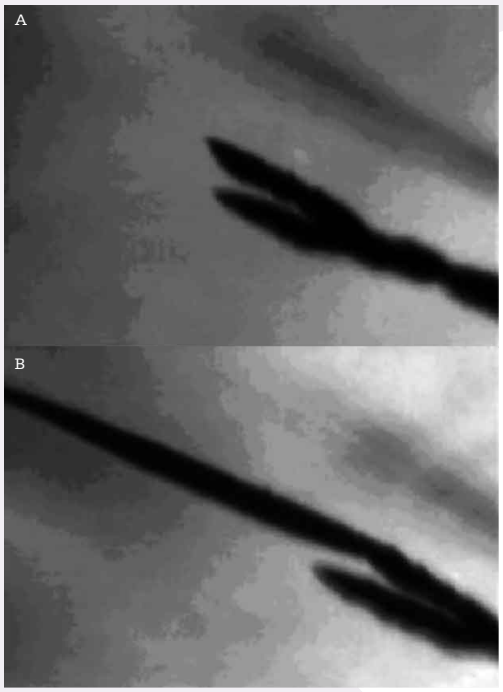


FIGURE 5. (A) shows a single crystal with the voltage turned off. (B) shows the same crystal several seconds after the voltage was turned on to 1500V. Note the thin, smooth, needle-like shape in (B).

"Since the results for iodine and ice are so similar, it seems clear that an electrostatically induced reduction of the equilibrium vapor pressure is responsible for the same effect in both systems."

HOW THEY GREW

The iodine crystals clearly show a voltage effect. Below 800V, not much happens, and the crystals grow about the same rate as without voltage (about 10-20 microns/sec). The appearance of the crystals in this range is varied: some dendrites, some blocky crystals, and some short, irregularly shaped needle-like crystals can be seen, as in Figure 4. Around 800-1000V, long, smooth needles start growing from the crystals that are already there, and the growth rate starts increasing, reaching about 100 microns/sec at 2000V, as indicated in Figure 3. An example of this growth can be seen in Figures 5(A) and (B).

Also, changing the polarity of the voltage seems to have no effect, which supports the theory of an electrostatic effect since it should be dependent only on the magnitude of the voltage and not the sign. The results are fairly messy since there are some uncontrollable factors, but clearly, the voltage does have some influence on the non-polar iodine molecules. In fact, the basic results are nearly identical to those for ice crystals: an increase in velocity beginning at about 1000V and continuing to 2000V, also with no polarity effect. This indicates that the effects in the two different materials are probably caused by the same, polarizability-independent mechanism, as per the theory.

THE NEXT STEP

It is difficult to take data cleanly with an experiment such as this because there are uncontrolled and unmeasurable variables. Every effort was made to reduce their effect, but the results still show a lot of scatter. The most important uncontrolled factor is competition between crystals. Often, four or five long needle crystals would form at the same time, growing in different directions, and the camera could only follow one of them to take data. Sometimes one would start growing even faster

while the others stopped, and sometimes they would grow relatively evenly. To reduce the influence of this factor on the data, the fleaker was gently tapped to knock off all but a few good crystals, and data was taken on what appeared to be the fastest-growing crystal in each batch. On several occasions, however, the crystal being analyzed stopped growing in favor of other crystals. The most obvious of these instances were noted during the data-taking and later removed, as noted above, but it was often hard to tell when this was happening.

The other major unmeasurable variable was the supersaturation. All of the data was taken within a moderately small range of supersaturation, estimated to be about 6-10% from computer modeling, but small local variations inside the chamber could not be monitored. Since variations depend on the vapor pressure distribution and air flow patterns in the chamber, the crystals could grow through regions of different supersaturation. Crude estimates could be made during the trials based on the density of the pink iodine vapor inside the chamber, and some crystals were clearly growing into low-density regions (usually near the sides of the chamber). Despite these efforts, competition and supersaturation variations still have a noticeable effect on the data, which is why many of the data points at higher voltages have a lower velocity than expected. The 80th percentile line on Figure 3, with 80% of the points below and 20% above, represents an estimate of what would happen in an ideal case where these factors are eliminated.

THE REAL VOLTAGE EFFECT

The primary conclusion to be drawn from this experiment is that an electric field can influence the growth of iodine crystals even though iodine is not polarizable. When the electric field is applied, distinct thin needle-like crystals form, and their growth rates can increase as much as tenfold. These effects are independent of both the polarizability of the material being used and the sign of the voltage, which indicates an electrostatic effect. Since the results for iodine and ice are so similar, it seems clear that an electrostatically induced reduction of the equilibrium vapor pressure is responsible for the same effect in both systems. Directions for further research include designing a cleaner setup on which the supersaturation can be monitored more closely and performing similar experiments on a variety of different materials. A series of investigations using other substances would be interesting because it could further verify the independence from polarizability and may reveal other properties that alter the voltage effects. Some preliminary tests have already been done with carbon tetrabromide, which does not seem to show any voltage effect. Perhaps this is because the crystals are softer and more rounded, as opposed to iodine, which is highly faceted. This might make it more difficult for sharp needles to form—the degree of faceting might be an important variable to investigate. Still, the potential is clear—this method could someday be used to precisely control crystal growth for industrial applications or to do chemistry in the strong fields at the needle tips. It could also point the way to a more highly developed and accurate theory of the mechanism behind crystal growth in electric fields.

Molly Swanson is a third year undergraduate in Physics at the California Institute of Technology. *This work was completed with Ken Libbrecht, Professor and Executive Officer of Physics at Caltech, and funded by the Mr. and Mrs. Carl Larson Endowment for the 2000 Caltech Summer Undergraduate Research Fellowship. The author wishes to thank Ken Libbrecht, Carolyn Merkel, and the Caltech Physics Shop.*

FURTHER READING

T. Kobayashi, T. Kuroda. *Morphology of Crystals: Part B.* I. Sungawa, Ed. (Terra Scientific, Tokyo,), chap. .

G. Crowther, C. P. R. Saunders. On the Aggregation and Fragmentation of Freely-Falling Ice Crystals in an Electric Field. *J. Meteorol. Soc. Japan* , - ( ).

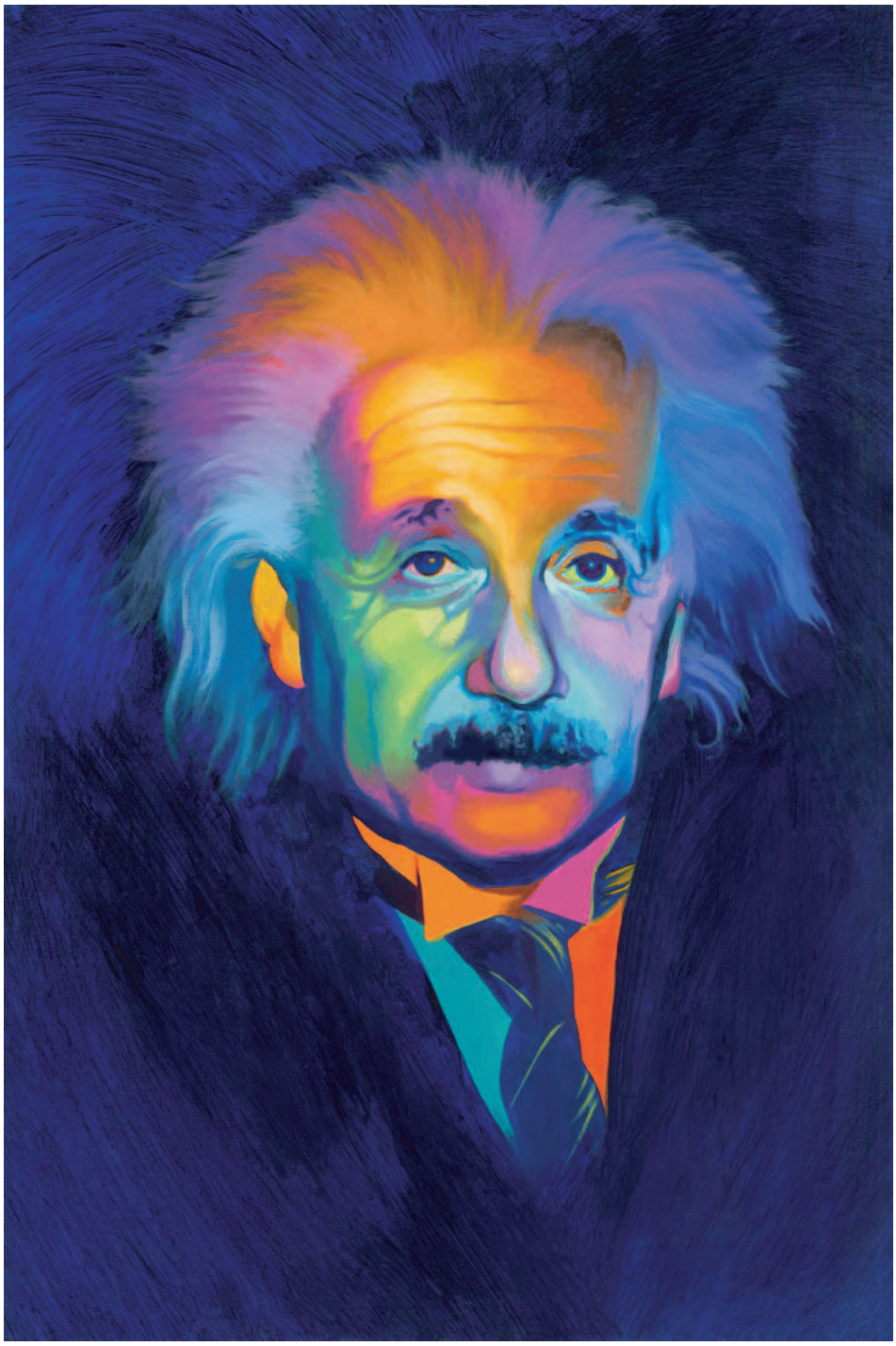
K. G. Libbrecht, V. M. Tanusheva. Cloud Chambers and Crystal Growth: Effects of Electrically Enhanced Diffusion on Dendrite Formation from Neutral Molecules. *Phys. Rev. E.* , - ( ).

K. G. Libbrecht, V. M. Tanusheva. Electrically Induced Morphological Instabilities in Free Dendrite Growth. *Phys. Rev. Lett.* , - ( ).

E. A. Brener, H. Muller-Krumbhaar. Comment on "Electrically Induced Morphological Instabilities in Free Dendrite Growth." *Phys. Rev. Lett.* , ( ).

K. G. Libbrecht, V. M. Tanusheva. Reply to comment on "Electrically Induced Morphological Instabilities in Free Dendrite Growth." *Phys. Rev. Lett.* , ( ).

GRAY MATTER BY BRYAN NORTON



Gray Matter
acrylic on board, " x "
Artist's Private Collection

Einstein's youthful curiosity has always been an inspiration for me. The animus for this color study derives from Einstein's words, "Imagination is more important than knowledge. Knowledge is limited. Imagination encircles the world."

Bryan Norton
Art Center College of Design

Not an alien life form or the latest video game.

It's what you become when you've made the right career choice.

Science & engineering opportunities for students & grads that are out of this world.