Editor's Note: PSC Science Highlights is an all-electronic magazine (e-magazine) available at www.psc.edu/sciencehighlights. The document you are holding is a printed summary of what's contained in our e-magazine, including a feature article and pointers to other articles highlighted here.

Welcome to the Spring 2019 issue of *PSC Science Highlights*!

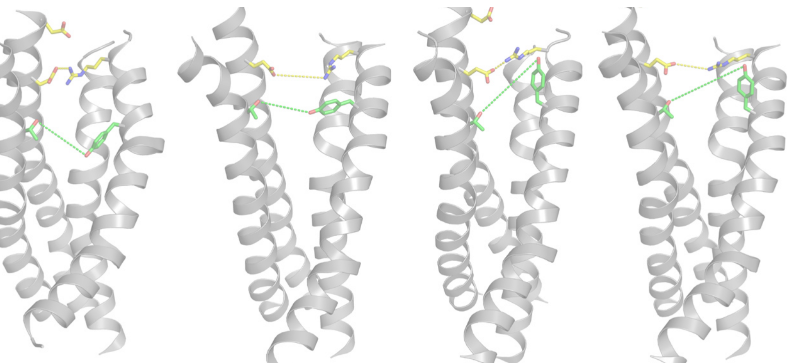
Over the last six months we have been busy upgrading PSC's capabilities in Big Data and artificial intelligence (AI). Arguably our most significant event was delivery of the first instance of the NVIDIA DGX-2 enterprise research AI system available to open research.

The DGX-2 is part of Bridges-AI, a $1.8-million, NSF-funded upgrade to our Bridges system that has added 88 NVIDIA Volta GPUs to an already successful CPU/GPU AI platform. Among other users, our Open Compass program will leverage Bridges-AI to help AI researchers across the nation tackle the most difficult problems in medicine, security, energy usage, business and other fields.

With the CMU Libraries, we are also organizing the first Artificial Intelligence for Data Discovery and Reuse conference at Carnegie Mellon. AIDR2019, to be held in May, will provide a platform for AI/machine learning researchers, data professionals and scientists to address data challenges and facilitate the next breakthroughs in science and technology using the power of AI and scientific data.

Featured in this issue is a Pennsylvania State University group that used Bridges' CPUs for machine-learning recognition of impending extreme weather. Next, they hope to leverage the deep-learning power of Bridges-AI to improve on their results. In a very recent AI success, the Libratus imperfect-information-game-playing AI designed at the CMU School of Computer Science (and which, as you have read in previous issues, runs on Bridges), has brought $10 million in Department of Defense funding to Pittsburgh.

Along the other axis of PSC's major emphasis—data science—is a new collaboration with the Compact Muon Solenoid (CMS) experiment at Fermilab. You can read about their use of Bridges to simulate possible signals in the "High-Luminosity LHC," a 10-fold power upgrade to the Large Hadron Collider (www.psc.edu/hadroncollider).



In bioinformatics, PSC received two major grants. PSC and the Department of Biomedical Informatics at the University of Pittsburgh's School of Medicine have received an initial $598K in NIH funding as part of the Human BioMolecular Atlas Program (HuBMAP). Another three-year $289K NSF grant will further our work as part of the National Center for Genome Analysis Support (NCGAS). This issue features a genomics story in which a UC Davis team used Bridges to identify genes associated with surviving climate change in birds.

Two molecular dynamics stories in this issue leveraged two different PSC systems with complementary strengths. A University of Delaware group used Bridges' ability to simulate extremely large molecular assemblies to uncover how a monkey protein prevents infection by HIV. A Cornell University collaboration employed the D.E. Shaw Research Anton 2 system hosted at PSC to simulate large molecules for long time periods. This capability shed light on how the TMEM16 scramblase proteins reorganize the cell membrane, a process critical to programmed cell death, blood coagulation and membrane repair.

Two articles reflect more traditional high performance computing successes. Georgia Institute of Technology researchers used Bridges to illuminate the physical chemistry of the two major pathways of electron transport. Bridges also helped a team at the University of Pittsburgh design a metal-oxide framework material that can turn effluent carbon dioxide into an industrially useful feedstock.

Since our last issue, our award-winning STEM education efforts have offered five Microsoft-funded GCode—"Girls Code"—workshops to expose young women to potential careers in programming and computer science. Our Wide-Area Classroom of HPC workshops for researchers and computing professionals served 571 in-person and remote students between December 2018 and January 2019, putting us on track to exceed the previous year's total of 3,015.

Finally, we're proud to note that at the SC18 high performance computing conference in Dallas in November, PSC received a best-ever six *HPCwire* Awards for accomplishments in the field. You can find the details in this issue.

Carnegie Mellon University | PSC PITTSBURGH SUPERCOMPUTING CENTER | University of Pittsburgh

About PSC

PITTSBURGH SUPERCOMPUTING CENTER (PSC) unites the most advanced computing technologies with multidisciplinary expertise and data to drive discovery for the benefit of society.

Established jointly with Carnegie Mellon University and the University of Pittsburgh, PSC creates and provides advanced computing ecosystems and solutions scalable to the greatest challenges. It engages in transformative, collaborative, cross-disciplinary research. It provides and supports education for academia, industry and the wider community. It promotes the transformative contribution of advanced computing to people's lives. And it engages with business and industry to drive economic and workforce development.

PSC was established in 1986 and is supported by several federal agencies, the Commonwealth of Pennsylvania and private industry.

Major Computing Resources

Bridges – a uniquely capable resource for empowering new research communities and bringing together HPC, AI and Big Data. Bridges is designed to support familiar, convenient software and environments for both traditional and non-traditional HPC users.
Anton 2 – a special-purpose supercomputer for biomedical simulation designed and constructed by D. E. Shaw (DESRES). A successor to Anton, Anton 2 is a 128-node system, made available to PSC by DESRES without cost for non-commercial research use by U.S. universities and other not-for-profit institutions. It is hosted by PSC with support from the National Institute of General Medical Sciences.

Thanks for your Support

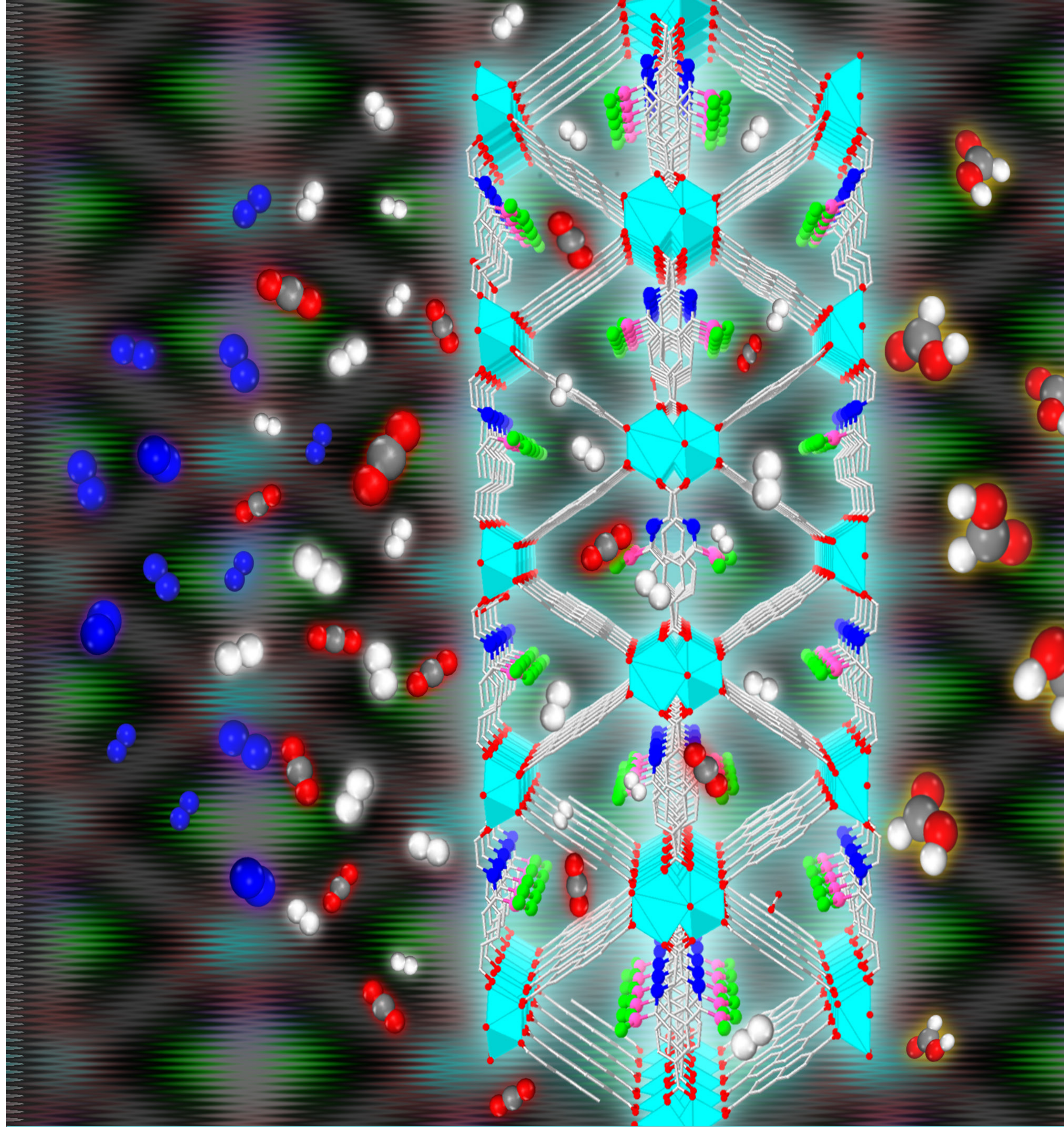
PSC gratefully acknowledges significant support from the following:

The National Science Foundation
The National Institutes of Health
The U. S. Department of Defense
The Commonwealth of Pennsylvania
D. E. Shaw Research
Microsoft
University of Pittsburgh

On the cover: *The simulated candidate metal oxide framework (MOF) converts carbon dioxide molecules (gray and red balls) to formic acid by reducing them with hydrogen (white balls). Image reproduced by permission of Karl Johnson and The Royal Society of Chemistry from Catal. Sci. Technol., 2018, 8, 4609-4617, DOI: 10.1039/C8CY01018H.*

Feedback: We would like to hear any feedback you have on our work or this publication. You can send any comments or suggestions via our feedback page at https://www.psc.edu/feedback. You can contribute to PSC's nonprofit, academic mission at https://www. psc.edu/donate.

Pittsburgh Supercomputing Center | 300 S. Craig Street, Pittburgh, PA 15213 | (412) 268-4960 | WWW.PSC.EDU



PITTSBURGH SUPERCOMPUTING CENTER
SCIENCE HIGHLIGHTS

SPRING 2019

FEATURE

DANGEROUS CURVES AHEAD

MACHINE LEARNING ENABLES SCIENTISTS TO SPOT "COMMA-SHAPED CLOUDS," EXTREME WEATHER

Between 1980 and 2018, according to the National Oceanic and Atmospheric Administration, 238 particularly severe U.S. weather and climate disasters—including severe storms, floods and other weather problems—caused more than $1.5 trillion in damages. Severe storms alone kill hundreds of people a year. It's safe to say quicker storm warnings would save lives and money.

When meteorologists started using satellite photos to help them predict the weather in the 1970s, they got a surprise. Before a storm developed, they saw the clouds in that area took on the shape of a comma. These comma-shaped clouds varied in size from tiny to enormous, with some spanning the continental U.S. from north to south. Before long, though, the meteorologists discovered they had too much of a good thing. The data from satellite photos expanded enormously, quickly overwhelming any ability for humans to monitor them in real time. The big comma-shaped formations were obvious—but smaller ones could still threaten life and property and were easy to miss.

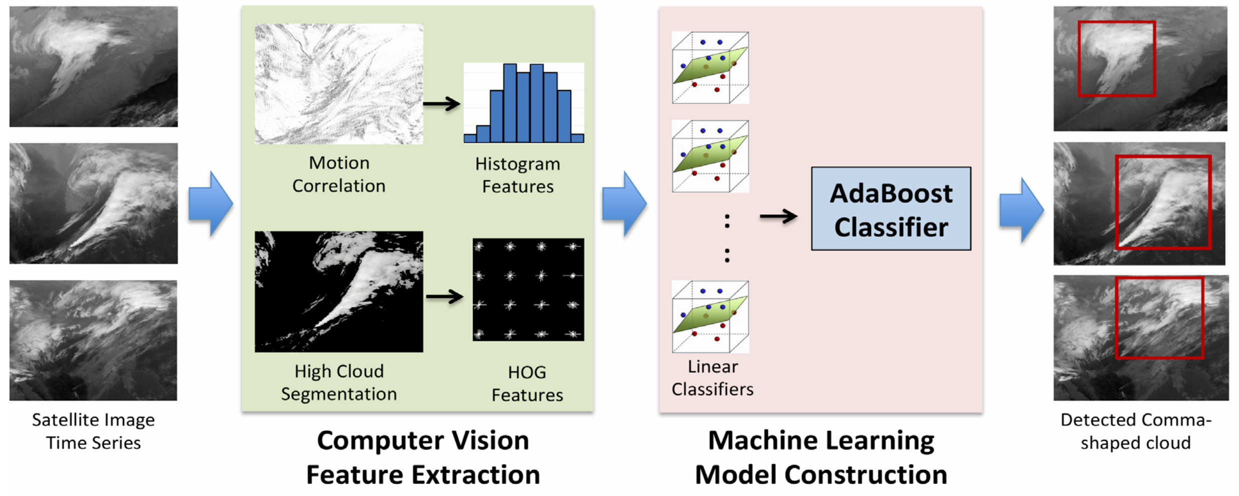
Enter James Wang of Pennsylvania State University. Wang is a computer scientist who uses Artificial Intelligence (AI) to solve "Big Data" problems.

Around 2008 he was looking for a good scientific application to inspire and test machine-learning technology, in which the computer starts from data that has been labeled by humans, and experiments many, many times on new data until it can reproduce the humans' performance.

> *"As computer scientists, we were really fascinated by this new field of data science, which we could leverage to make the computers more efficient in solving real-world problems. The issue was to define a problem that had enough data to work with. But not just a lot of data. The data needed to be manually labeled in order for the computer to learn … Meteorology is one of the rare fields where you can find a huge amount of data that are so complex people don't know how to leverage it—and the data are indexed [labeled]."—James Wang, Penn State*

Detecting severe thunderstorms was a perfect problem. The data are enormous, with tens of thousands of weather stations worldwide and satellites generating gigabytes of high-resolution images, radar data and other data every day. But they have also been carefully reviewed and labeled with "ground truth" by expert meteorologists. Plenty of accurate examples of developing storms were available for the computer to start with.

Read more at: WWW.PSC.EDU/AISTORMS



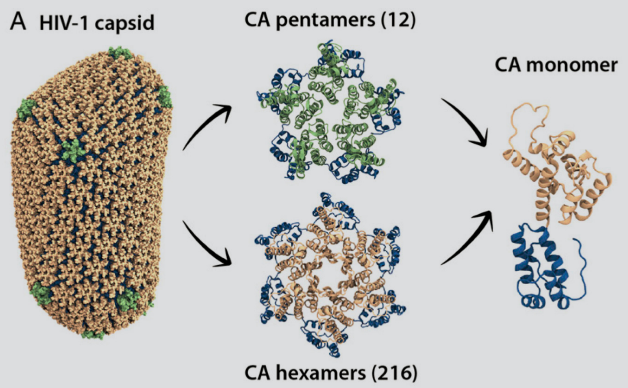
The comma-cloud-detecting AI takes a time series of satellite images (left), first extracting critical features (center left), and then using the AdaBoost classifier to convert those data (center right) into a detection (right).

HIGHLIGHTS

Stopping HIV in Its Tracks

A monkey "host restriction" protein works by destabilizing HIV as it enters the cell, according to scientists at the University of Delaware and the University of Pittsburgh (Pitt). This destabilization, revealed with simulations on PSC's Bridges system, shows why HIV, the virus that causes AIDS, can't infect monkey cells. Understanding how the monkey protein stops infection, and why the human version of the protein doesn't, promises a completely new avenue for protecting people against the virus.

WWW.PSC.EDU/MONKEYPROTEIN

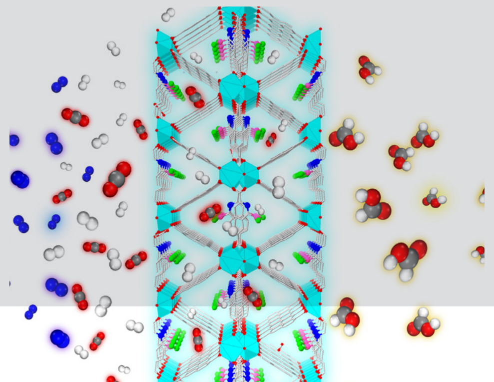


The HIV capsid's oblong shape (left) is made possible by the CA protein (right) sometimes assembling in five-membered pentamers (center, top), and sometimes in six-membered hexamers (center, bottom). The monkey TRIM5a protein disrupts the capsid by making the pentamers more rigid than they need to be to function.

Capture and Convert

A new material may be able to capture carbon dioxide and turn it into a commercially useful substance, according to a team at the University of Pittsburgh. Using PSC's Bridges system, they simulated two "metal oxide framework" materials that simulated removal of carbon dioxide from exhaust gas. Better, the material also converted it into formic acid, which can be used to make products like methanol fuel. If the material works as well in the lab and factory as it does in the computer, it could fundamentally alter the economics of limiting human CO_2 release and avoiding climate change.

WWW.PSC.EDU/CLIMATECHANGE

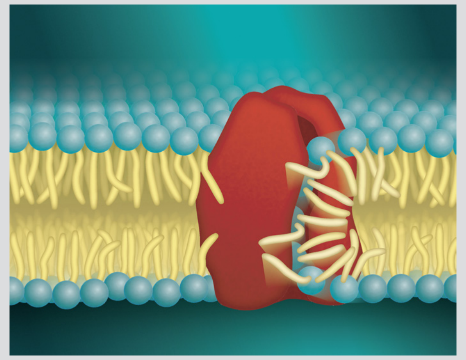


The simulated candidate metal oxide framework (MOF) converts carbon dioxide molecules (gray and red balls) to formic acid by reducing them with hydrogen (white balls). Image reproduced by permission of Karl Johnson and The Royal Society of Chemistry from Catal. Sci. Technol., 2018, 8, 4609-4617, DOI: 10.1039/C8CY01018H.

Phospholipids at the Gate

A protein that scrambles the contents of the inner and outer surfaces of the cell membrane has an unexpected gate at its center, discovered by Weill Cornell Medical College scientists using Anton 2 supercomputer simulations at PSC. The slot suspected of allowing membrane components to move between the membrane's inner and outer surfaces works much like a credit card reader, as scientists had thought from previous work. But thanks to the uniquely long time periods and large molecular complexes that Anton 2 can simulate, the Weill group has discovered an unforeseen additional mechanism that could provide drug targets in cancer, heart disease, stroke and blood disorders.

WWW.PSC.EDU/PHOSPHOLIPIDS

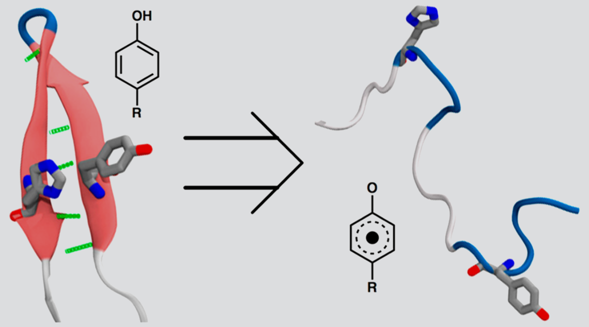


The "credit card" model for TMEM16. The water-loving heads of the phospholipid (blue) swipes through the slot in the scramblase (red), hiding the head from the fat-loving tails in the interior of the membrane (yellow) as the phospholipid flips from one membrane leaflet to the other. By Adam Steinberg; reproduced with permission.

A Basic Switch

Large changes in protein-chain folding accompany a pH-dependent switch between the two modes of an important life process called "electron transport", Georgia Tech scientists have suggested. Using simulations on the Bridges system at PSC, the team has explained otherwise mysterious lab results. The finding sheds light on a process central to plants building sugars from sunlight, higher organisms using those sugars for energy and the development of some human diseases.

WWW.PSC.EDU/ELECTRONTRANSPORT



At high pH, the simulations show, a tight, hairpin-shaped bend in the protein (left) gives way to a wobbly, random coil (right). Reprinted with permission from J. Phys. Chem. B 2017,121,15,3536-3545. Copyright 2017 American Chemical Society.

The Fingerprints of Survival

A UCLA-led team has found they can predict whether birds are vulnerable to climate change by comparing their genes to their changing environments. Using PSC's Bridges system, Rachael Bay and her colleagues working with the Bird Genoscape Project (birdgenoscape.org) found they could identify genes in the yellow warbler and the willow flycatcher that associate with recent declines in local populations of each bird. The work is a first step in understanding climate sensitivity of many species that the researchers hope will offer new tools to conservation managers.

WWW.PSC.EDU/WARBLER

A yellow warbler photographed during breeding season (Credit: Joe Ellis)

PSC NEWS IN BRIEF

- PSC Supplies Computation to Large Hadron Collider Group
- PSC Receives Best-Ever Six 2018 HPCwire Awards
- Microsoft Supports PSC's GCode Workshops
- PSC Receives $289K for Expansion of Genome Biology Research
- PSC & University of Pittsburgh (Pitt) Build Infrastructure for Human Tissue Atlas

WWW.PSC.EDU/SPR19BRIEFS