



# W4DC

WOMEN IN HIGH  
PERFORMANCE  
COMPUTING



## PRACE Women in HPC Magazine 2015



ALL RIGHTS RESERVED. This report contains material protected under International and European Copyright Laws and Treaties. Any unauthorized reprint or use of this material is prohibited. No part of this report may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system without express written permission from the author / publisher.

The copyrights of the content and images contained within this report remain with their original owner(s). The PRACE name and the PRACE logo are © PRACE.

Front and back covers © PRACE.

## **PRACE Women in HPC Magazine 2015**

[www.prace-ri.eu](http://www.prace-ri.eu)

### **Editor-in-Chief**

Marjolein Oorsprong – PRACE Communications Officer

### **Editors**

Alison Kennedy – EPCC, United Kingdom

Toni Collis – EPCC, United Kingdom

Sylvie Joussaume – INSU/CNRS, France

Tiina Leiponen – CSC, Finland

### **Proof Reading**

Frances Smolinski

### **Success Story Editors**

Harry O'Neill

Bill Sherwood

### **Graphic Design**

Crystal Moody

Published by Insight Publishers, Bristol, UK

The Partnership for Advanced Computing in Europe (PRACE) is an international non-profit association with its seat in Brussels. The PRACE Research Infrastructure provides a persistent world-class high performance computing service for scientists and researchers from academia and industry in Europe. The computer systems and their operations accessible through PRACE are provided by 4 PRACE members (BSC representing Spain, CINECA representing Italy, GCS representing Germany and GENCI representing France). The Implementation Phase of PRACE receives funding from the EU's Seventh Framework Programme (FP7/2007-2013) under grant agreement RI-312763 and from the EU's Horizon 2020 research and innovation programme (2014-2020) under grant agreement 653838.

**Published:** May 2015

**ISBN:** 9789082169423



# Women in HPC 2015

## Table of Contents

<b>Editorial .....</b>	<b>4</b>
<b>Women in HPC Network.....</b>	<b>5</b>
<b>Projects .....</b>	<b>6</b>
Tackling turbulence to minimise emissions.....	6
Natural strength .....	8
Quarks and Gluons in Action! .....	10
Understanding a new generation of solar cells .....	12
Computers in the fight against cancer.....	14



# Editorial

## PRACE Women in HPC

I am delighted to present you this pre-edition of the PRACE Digest showcasing Women in HPC. PRACE, in collaboration with the Women in HPC initiative, is embracing a year celebrating the contribution women make to HPC and computational science. This magazine puts the spotlight not just on the scientific advances made by PRACE, but also on the women that contribute to make PRACE a world-leading force in HPC-enabled science and in the move towards Exa-scale computing.



PRACE has always aimed to improve European competitiveness. HPC is a key economic driver in Europe, providing the ability for scientific breakthroughs that are impossible without leading edge computing resources. Enabling equality of opportunity and participation is also important to increase this competitiveness. PRACE's year of Women in HPC aims to emphasize the work that PRACE is already doing to improve the opportunities for women, and the input that they provide to the scientific community. By publishing the first magazine of this kind in the field of HPC, Europe is leading the way to highlight the contribution women make in this rapidly developing field.

By showcasing the talents of women, we hope this publication will provide inspirational models to the next generation of PRACERs. HPC is uniquely positioned to attract women as its users are from a broad range of fields, and HPC already has a growing proportion of female users. At international events, such as ISC and SC, the Women in HPC initiative draws great interest from the HPC and scientific community.. The number of women working in HPC is growing day by day. The HPC community is strengthening so that women no longer feel that they are the exception but instead are part of the norm.

This publication and the PRACE year of Women in HPC, in collaboration with the Women in HPC initiative are steps to promote diversity in HPC. We hope that this year raises the profile of female professionals in HPC and engages the community, challenging the stereotype of an HPC scientist. Until diversity is embedded in the community and equality of opportunity and ability becomes commonplace, please enjoy reading about the women who are already changing the face of HPC.

**Dr Sylvie Joussaume**

*Chair of the PRACE Scientific Steering Committee*

# Women in HPC

Women in HPC was set up to address the gender imbalance in HPC through support, research and initiatives to bring women into the community.

The Women in High Performance Computing (WHPC) Network works to raise the professional profile of Women in HPC by:

- Bringing together female HPC scientists, researchers, developers, users and technicians
- Raising the professional profile of women working in HPC
- Increasing the participation of women in outreach activities
- Assessing the influence of equality initiatives on the HPC community

To encourage women to participate we aim to:

- Increase the visibility of experienced female role models in HPC
- Provide opportunities for networking amongst women in HPC, both peer-to-peer and peer-to-role model
- Raise awareness that gender-balancing of research groups improves scientific output

- Discuss how to balance careers with other commitments
- Highlight best practice by institutions that are already achieving gender-balance in HPC
- Create and disseminate an action plan that quantifies and addresses the gender imbalance in HPC
- Initiate a continuing programme of surveys of HPC facility users to quantify how many women continue in HPC-related careers and to discover if (and why) they leave them
- Provide quantifiable evidence to those attempting to assess the efficacy of schemes such as Athena Swan, Project Juno and the 'gender dimension' in the European funding scheme 'Horizon 2020'.

## Upcoming Events

[www.womeninhpc.org.uk/events](http://www.womeninhpc.org.uk/events)

In 2015 we will be running a series of events including:

**16 July 2015:** Women in HPC Workshop at ISC 2015 (more details below)

**4 September 2015:** Women in HPC UK Workshop, British Computing Society, London

**19 November 2015:** Women in HPC at SC15

**Women in HPC Workshop In collaboration with ISC 2015 14.00-18.00, Thursday 16 July 2015 Kilobyte Room, Frankfurt Am Main, Germany**

This event will bring together female early-career researchers with a focus on European participation, providing them with an opportunity to showcase their work and to meet role models and peers in an environment designed to move beyond the stereotype of HPC as a male-dominated field. It will open with an introduction to current research by the Women in HPC network, the demographics of the HPC community and the different experiences of men and women in the field.



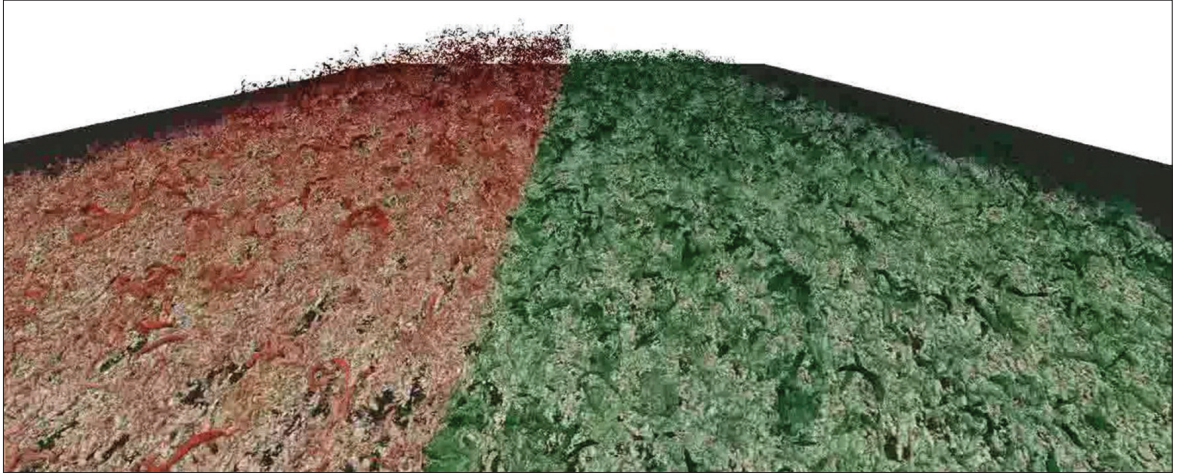
Building on our first workshop at SC14, we will also host a panel discussion, focusing on "How can women fulfill their dreams in HPC, combining work with having a family and working in a male-dominated environment?".

Call for papers: Deadline for submission: 1 June 2015, 5pm GMT Full details:

[www.womeninhpc.org.uk/isc15#Call](http://www.womeninhpc.org.uk/isc15#Call)

# Tackling turbulence to minimise emissions

Advances in modern supercomputers are for the first time allowing researchers to get to grips with the physics of wall-bounded turbulence in adverse pressure gradient (APG) environments, a leap that could mean big changes for the future of our energy consumption



Above: Spatial development of turbulent structures along a flat plate subjected to adverse pressure gradient (left) and zero pressure gradient (right). Courtesy of Vassili Kitsios.

**T**urbulence is a messy, convoluted phenomenon. With its characteristic swirling motions comprising an array of interacting length scales at several different orders of magnitude, it is no surprise that the physics behind it has been an ongoing challenge to understand. With wall-bounded turbulence and adverse pressure gradients (APG), however, things get even messier.

Turbulence has a huge impact on energy efficiency, directly affecting the design of aircraft wings, wind turbine blades or any type of turbomachinery, therefore knowing how to minimise drag is crucial for conserving energy. Throw a wall into the equation and the nature of turbulence is suddenly changed on a fundamental level, making it much harder to grasp the physics. But if you can't properly understand the nature of the flow in an APG environment then you can't properly design efficient machines such as jet engines and turbines.

With massive parallel computing resources available to European scientists through PRACE, the possibility of finding solutions to these challenges is closer at hand than ever before. At the Leibniz Supercomputing Centre (LRZ) in Germany, the high performance SuperMUC computer is being used to run direct numerical simulations (DNS) of wall-bounded turbulence in APG environments. Dr Ayşe Gül Güngör, assistant professor in the Faculty of Aeronautics and Astronautics at Istanbul Technical University, is working in collaboration with Javier Jimenez and Julio Soria on the investigation into flow dynamics. By enabling more energy efficient design, their research aims at cleaner and more efficient power generation and a reduction in fuel consumption to minimise our emissions of CO<sub>2</sub>.

The literature currently available focuses on simpler theoretical problems, because APGs introduce other parameters into the flow, these must be understood before it is possible to develop and bring in the control



Dr Ayşe Gül Güngör

*“By enabling more energy efficient design, this research aims at cleaner and more efficient power generation and a reduction in fuel consumption to minimise our emissions of CO<sub>2</sub>”*

techniques that impact the design process. “Understanding the flow is the main issue,” explains Güngör. “A clear and agreed understanding of which parameters are paramount for the development of the APG boundary layer does not yet exist.”

To help overcome turbulence, active and passive flow control systems can be introduced. For example, a passive control technique could be the roughness or smoothness of an aeroplane wing. When an aeroplane is landing, taking off or in flight, it will experience different flow regimes, which need to be controlled actively during the aeroplane’s operation. Air-blowing devices, like sweeping jet actuators, can be installed to receive information from the flow and adjust themselves accordingly. Active flow control therefore helps to reduce drag by counteracting the flow separation that occurs in APG environments.

Güngör, however, is attempting to do more than reduce flow separation. With the aid of SuperMUC, she is hoping to obtain an equilibrium state of flow. The equations that govern any type of fluid motion are called Navier-Stokes equations, so they can be used in a broad range of modelling scenarios including ocean currents and fluid transport through pipelines. One way to solve Navier-Stokes equations is through DNS calculations. Unlike other methods of computational fluid dynamics, DNS solves these equations without recourse to further modelling, meaning that the entire range of spatial and temporal scales needs to be resolved. These kinds of calculations can produce a full set of data at any point within the flow and at any time with the simulation. If you take out the modelling then you don’t need to rely on approximations. “DNS provides access to the full spatial and temporal velocity field, while real experiments are limited by experimental

facilities and measure a sparse subset of the flow field,” states Güngör.

Despite advances in high performance computing, it’s not yet possible to solve all of flow dynamics with DNS, but things are moving in the right direction. As it currently stands there is a tremendous amount that can be tackled, as Güngör relates: “There are many turbomachinery applications that you can do with DNS right now but not without these computers. Without PRACE this project would be impossible.”

Along with access awarded by PRACE to the supercomputing facilities, the experimental water tunnel at Monash University’s Laboratory for Turbulence Research in Aerospace and Combustion (LTRAC) has also been made available. Here, Güngör will eventually be able to conduct a series of experiments, but some agreement on the DNS is needed first.

For Güngör this project is just the beginning of a long-term plan to revolutionise the design of energy generation and transport platforms in APG conditions. Taking its cue from previous investigations into non-equilibrium boundary layers and running concurrently with a similar project using national (Tier-1) computing facilities, Güngör is confident that more are sure to follow.

Reducing energy consumption is a major challenge the world over, but there is currently little activity in this area of research. “Knowledge on this type of flow is scarce, and we are really working hard to obtain some new insights into their physical nature and structure,” states Güngör. When the experiments eventually take place at LTRAC, the results will lead to new control strategies and new theories that could potentially be applied in any industrial environment.



## Natural strength

Composite materials made up of inorganic and biological matter present remarkable properties including fracture resistance, toughness and strength. **Professor Frauke Gräter** has been investigating the mechanical properties of nacre, a material that possesses great stability due to its elaborate hierarchical nanostructures



Professor Frauke Gräter

**T**he iridescent shimmer found on the inside of seashells and the outside of pearls is caused by a substance called nacre. It is one of the strongest and most resilient materials found in nature due to its composite structure, pairing stiff but brittle calcium carbonate crystals with soft layers of protein that impart it an elevated resistance against fracture. Biomaterials such as nacre are highly organised at the nanoscale, so by studying this organisation it may be possible to elucidate the source of their remarkable mechanical performance, which often outstrips synthetic analogues.

Structurally, the majority of nacre is made up of calcium carbonate crystals

that are several tens of nanometres thick. These are stacked in a very organised manner similar to a traditional ‘bricks and mortar’ arrangement. Intersecting these crystals in parallel stripes are layers of protein and chitin, only a few nanometres thick, which provide extra mechanical stability.

Professor Frauke Gräter of the Heidelberg Institute for Theoretical Studies has been leading a project investigating the separate components of nacre under physical loads. “The calcium carbonate found in nacre comes in the form of aragonite, a crystalline structure commonly found in nature,” she says. “A wider debate in materials science at present is the role that flaws — tiny errors in the molecular structure — have on mechanical performance of crystals. Experimental data shows us that these flaws exist in aragonite, so we want to determine whether or not these molecule-sized gaps in structure cause the material to fail earlier under stress.”

*“Using 11.5 million core hours on the HERMIT supercomputer, we have extensively simulated crystals of different sizes, with different sized flaws and at differing loading conditions”*

Behaviour at the nanoscale is notoriously hard to predict, and as such it does not necessarily follow that flaws in a structure at this level will confer the same type of structural weakness that flaws of a similar proportion do at the macroscale. However, Gräter and her team have found that these flaws do indeed impact on the structural mechanics of aragonite. “Using 11.5 million core hours on the HERMIT supercomputer, we have extensively simulated crystals of different sizes, with different sized flaws and at differing loading conditions,” she says. “We found that even the smallest flaw of a few atoms causes the crystals to fail earlier.”

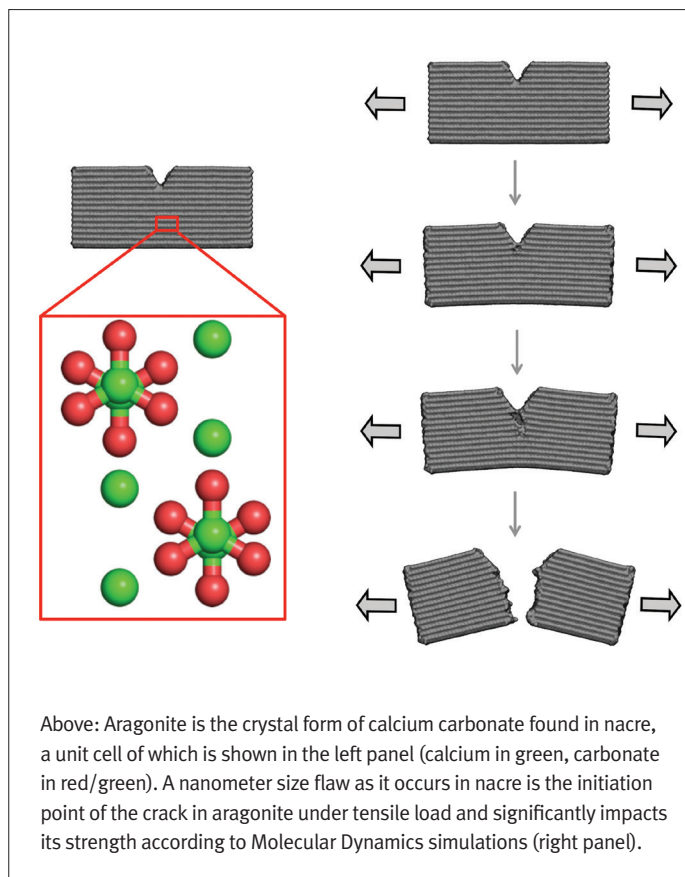


*“Our programme highlighted the areas at which stress concentrates when placed under load, and it was clear to see that even tiny flaws cause stress to concentrate at that point”*

As part of their work, the group developed a method called force distribution analysis. When engineers want to test the stability of a structure, they use crash test simulators. Force distribution analysis does the same thing but at a molecular scale. It exists as a patch for GROMACS, a widely used molecular dynamics code, and has been integral to understanding the role of flaws in the crystals. “Our programme highlighted the areas that stress concentrates when placed under load, and it was clear to see that even tiny flaws cause stress to concentrate at that point,” Gräter states.

This discovery has forced the team to look for alternative sources for nacre’s strength. They now believe that the protein component may compensate for flaws in the crystalline structure of aragonite. “If the protein is able to fill in the gaps of the crystals, it may be acting as a shock absorber, dissipating the forces and stress concentrations and thus allowing the crystal to behave as if it is flawless,” explains Gräter.

The team have already carried out some preliminary studies on the peptides present in the protein layer of nacre and their interaction with the aragonite crystals. The peptides, glutamate and aspartate, have negatively charged side chains that bind to



the positively charged calcium ions present in the crystal. The study involved binding these peptides to the crystal surface and then pulling them off, and it was shown that the bond is very strong.

The next step for the researchers will be to simulate the two components in action together. “The idea that the protein layer acts as a sort of shock absorber in nacre has been around for a while, but what that idea is really missing is actual data about how the forces are propagated from one layer to the other at the molecular scale,” Gräter states. “We’re aiming to provide a solid experimental basis for what has already been speculated.”

Biocomposite materials such as nacre are found everywhere in nature, including in the human body in the case of teeth and bones. Similar arrangements of stiff crystalline blocks and thin soft organic layers confer them similar structural integrity. At present, materials used to replace bone and teeth such as ceramics are fully inorganic, but Gräter’s work may provide a push towards creating similarly structured synthetic materials that closely mimic the mechanical stability of the original skeletal tissue.

## Quarks and Gluons in Action!

The standard model is sometimes described as a “theory of almost everything”, but some phenomena remain unexplained by it. **Professor Constantia Alexandrou** of the University of Cyprus and The Cyprus Institute has been using ab initio simulations of quantum chromodynamics in order to calculate fundamental properties of protons – that could provide a link towards probing the hidden physics of dark matter

**T**he strong interaction is one of the four known fundamental forces in the universe. It holds together nucleons (i.e. protons and neutrons) within the nucleus of atoms, and arises from the exchange of gluons between the quarks that make up these nucleons. Gluons act as “exchange particles” for a property called colour between quarks (colour relates to the strong interaction similarly to the way that charge relates to electromagnetism).

The theory of strong interactions and the way that quarks and gluons interact is known as quantum chromodynamics (QCD). Due to the nature of strong interactions, the analytical methods used to look at other fundamental forces such as electromagnetism do not apply. Other methods are thus needed in the study of the complex phenomena arising from the strong interactions described by QCD.

The only approach to directly study quantum chromodynamics in a non-perturbative manner is known as Lattice QCD. It is a discretised version of QCD, meaning that it is formulated on a grid of points in space and time. Through simulating quantum chromodynamics in this way, researchers are able to calculate quantities such as the ab initio mass of a proton. Professor Constantia Alexandrou explains more: “In this theory, the fundamental degrees of freedom are the quarks and the gluons, which come together

to form the proton. Using QCD, we can calculate from first principle the mass of the proton, which makes up most of the visible material we see in the universe.”

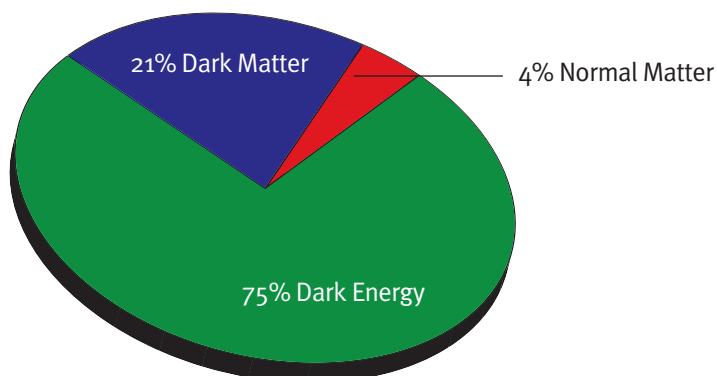
In Alexandrou’s work millions of integrals are performed, using the “Hybrid Monte Carlo” method that involves the use of sophisticated methods and extensive amounts of core hours on Tier-0 supercomputers. Calculating the inverse of matrices having billion by billion elements gives the quark propagators that form the basic ingredients to calculate the mass of the proton and other particles.

“Carrying out ab initio calculations such as these require huge computational resources,” says Alexandrou. “When I got my job at the University of Cyprus, this was a huge problem because we didn’t have access to these kinds of resources. I have always been supportive of the existence of a Pan-European HPC infrastructure such as PRACE for this reason, as it means that



Professor Constantia Alexandrou

*“The question is: can we calculate observables that can give us hints to what dark matter is?”*

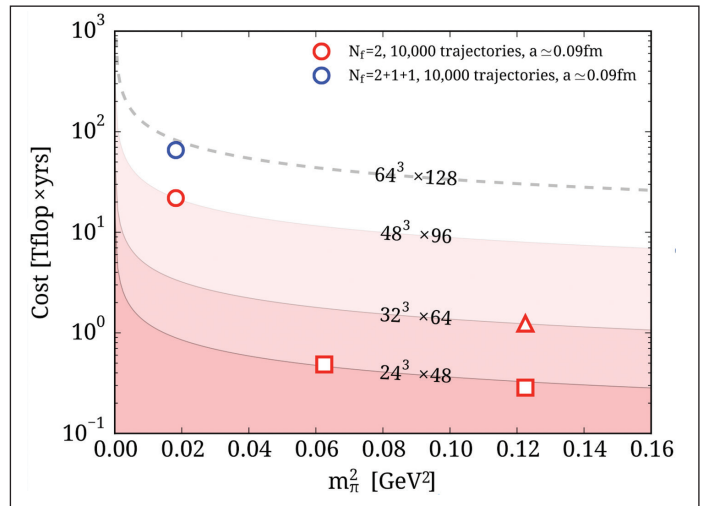


no matter where you are from in Europe, your research can still be competitive. And empowering all scientists in Europe to produce excellent science makes Europe more competitive.”

The codes used for lattice QCD calculations are in fact quite simple compared to the vastly complicated codes used in areas such as climate modelling. This allows researchers to rewrite their codes every time new computer architectures are introduced, enabling them to achieve good scaling and performance on computers relatively easily. Due to this fact, lattice QCD has often been used as a prototype application for new computer architectures. “Many people in our field are also experts in HPC technologies. Lattice QCD scientists helped to develop the Blue Gene supercomputer architecture, and so it is a field that not only benefits from supercomputers but also contributes to their development,” explains Alexandrou.

QCD has entered an exciting era as, for the first time, it has become possible to simulate the theory with physical parameters that connect to the real world. “This is the moment we have all been waiting for,” says Alexandrou. “We will of course need theoretical and algorithmic developments to pursue calculations of more complex quantities that hold the promise of probing physics beyond the standard model of Particle Physics. It is a great challenge and a thrilling journey.”

The discovery of the Higgs boson at CERN completed all aspects of the standard model. However, this by no means indicates that a complete picture of the world around us has been formed. Only about 5 per cent of the matter in the universe can be explained with current theoretical knowledge. The remaining 95 per cent is made up of the elusively named duo of dark matter and dark energy



– elusively named because we know so little about what they actually are.

“The question is: can we calculate observables that can give us hints to what dark matter is?” says Alexandrou. “To be able to make predictions at this level, we first need to be able to reproduce from first principles calculations what we do know. This is our task at the moment: computing some of the known quantities in order to benchmark our new methods, while at the same time developing these methods so that we can look for more complicated quantities that might allow us to glimpse at entirely new phenomena of physics.”

As HPC technology and our methodology develop in the coming years, the more we will be able to predict using lattice QCD. “We want to be able to provide experimentalists with some clues as to how they can find out more about dark matter. Technological and algorithmic developments will help drive this. Solving QCD, unlike other theories of fundamental physics, is deeply intertwined with HPC technologies, and so the hours allocated to us by PRACE are absolutely essential to make progress in our work. This area of research cannot exist without access to supercomputers.”

Above: Cost to simulate lattice QCD for a given lattice spacing as a function of the pion mass and different lattice volumes. Simulations are performed by the European Twisted Mass Collaboration.



## Understanding a new generation of solar cells

The unprecedented rise of organometal halogen perovskites has put them at the forefront of promising solar cell technologies yet a great deal about them remains unknown. With the aid of high-performance computing resources, researchers from EPFL want to understand more



Dr Ursula Roethlisberger

Considering how recently perovskite solar cell technology appeared on the photovoltaic scene, the near doubling of its energy conversion efficiencies in just a few years has been nothing short of breathtaking. Not least because the performance potential of perovskite cells isn't even close to being tapped out. Originally gaining wide interest for their exceptional light absorbing properties, halogen perovskites ( $ABX_3$ ) have only become more attractive and discussed among scientists since their abilities as charge carriers and transporters were noted as well. By exploiting these properties it may be possible to develop simpler and even more efficient solar cells, but nobody really knows what fundamental mechanisms are responsible for these desirable attributes.

Leading a concerted effort to uncover these mechanisms is professor Dr Ursula Roethlisberger, head of the Laboratory of Computational Chemistry and Biochemistry

(LCBC) at the Swiss Federal Institute of Technology in Lausanne (EPFL). Roethlisberger aims to elucidate the mechanisms of perovskite cells through direct simulation in order to learn how perovskite qualities such as structure and composition affect its transport properties.

PRACE allocated the project over 17 million core hours CPU time on SuperMUC, one of the Tier-0 supercomputers housed at the Leibniz Supercomputing Centre (LRZ) in Germany. "In terms of these types of explicit simulations we really need massive resources. We couldn't have performed them on any type of machine here at EPFL," states Roethlisberger. Through PRACE, projects like this are able to achieve in a couple of months what would otherwise take a whole year.

*"Through PRACE, projects like this are able to achieve in a couple of months what would otherwise take a whole year"*

At their most basic, photovoltaic cells require light absorption that leads to electronic excitation. This generates electron-hole pairs or excitons. When electrons are excited they are generally more mobile which means they can be made to create a current or voltage by transferring them to an electrode. To do this, one needs to be able to separate the opposing types of charge carriers and collect them separately at the electrode. The more efficient this transport is, then the fewer losses there will be in the conversion process. Recent experimental work points towards efficient perovskite transport of holes and electrons that could make them ideal solar cell material. Consequently, the clamour to investigate and uncover its properties means the field is developing incredibly fast, with Roethlisberger heading just one of many groups focusing on its potential.

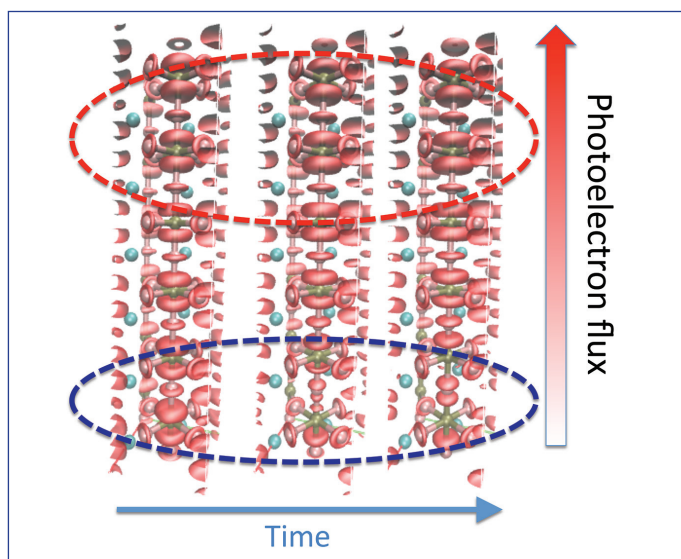
To understand the basic mechanisms, Roethlisberger's group is carrying out direct simulations of the charge carrier transport in the excited state. First, semi-classical models allow them to characterise the transport properties in terms of effective masses for electron and holes. Then, the composition or the structure can be changed in order

to see how these transport properties are affected. “We’ve got a good knowledge of how the structure and chemical composition affect the electronic properties, and we understand how they impact electron-hole masses too,” she says.

With this knowledge it is possible to predict just how efficient a perovskite cell will be if it uses, for example, different halogens or metals. But at the highest level of these simulations the project aims to go beyond this simpler type of characterisation. Then it will be possible to really see how the charge is transported for the holes and the electrons. It is likely that electron-hole transport in perovskites is anisotropic, meaning that the material does not conduct equally well along the different crystallographic axes. Although in principle perovskites can be used for both light absorption and for charge carrier transport, in most instances solar cells of this kind actually consist of an electron conducting material, such as titanium oxide, then the perovskite and then a hole transporting material. With multiple interfaces and anisotropy to deal with, it is crucial that the perovskite is oriented appropriately with respect to the electron conducting material.

Roethlisberger aims to investigate this anisotropy by applying an external electric field both parallel and orthogonal to the perovskite structure’s tetragonal axis to see if its transport properties are indeed dependent on direction. Once determined, Roethlisberger’s results will allow her to see which prospective solar cell materials are sensitive in the way they are coated on the electron conducting material and in which direction they should be oriented.

She has to work fast to stay ahead of the game, however. The rapidity at which the field of halogen perovskites is developing is



almost impossible to exaggerate despite some of the basic physics still being unsolved. With their potential to exceed 20 per cent conversion efficiency, the race is on to uncover the ways in which this can be made a reality. Currently Roethlisberger’s team is about half way through its allocation time with SuperMUC, a resource which has proved absolutely vital, as she says: “In this field, so many papers are published and the technology is developing so fast that you have to be very efficient to get new results. So for us, PRACE is very important.”

With this advantage, the team is now preparing one of the first papers to come out of the project to discuss the largely unreported “hysteresis effect”. Hysteresis describes the output of a system as being dependent on both current and past inputs, a characteristic which can lead to miscalculations of energy efficiency values and might also affect long-term stability. But the origin of the observed hysteresis in J-V curves has been unknown until now. “We basically found out that this has something to do with ionic transport in the material,” explains Roethlisberger, “so this is the first paper we are writing up using the experimental results.”

Though a major scoop for the team, there is still plenty to do in terms of computation and theoretical characterisation of perovskite systems as Roethlisberger states: “This is just the tip of the iceberg. We started with one of the most difficult simulations but there are so many different issues that need to looking into.” As interest in these materials continues to soar in both the academic and the industrial world, research like Roethlisberger’s is proving to be a vital step toward the development of highly efficient and potentially low-cost renewable energy technology.

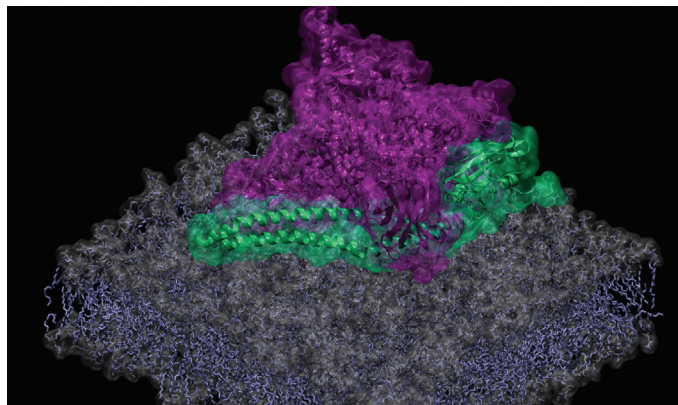
# Computers in the fight against cancer

Understanding the way in which mutant proteins cause the onset of cancer opens up the possibility of designing drugs that can stop this process. **Dr Zoe Cournia** and her team have been using the Curie supercomputer to investigate two mutant proteins often found in a number of cancers

**T**he protein PI3K $\alpha$  is an enzyme involved in fundamental cellular processes such as cell growth, cell division and formation of new blood vessels. All of these processes can aid cancer cell survival. Consequently, mutations in this protein often lead to cancer. Specifically, they are found in 29% of breast cancer patients, 35% of endometrial cancer patients and 18% of colon cancer patients.

Using computer simulations, Dr. Zoe Cournia, Dr. Paraskevi Gkeka, and Dr. Hari Leontiadou of the Biomedical Research Foundation Academy of Athens have been investigating how mutations in the PI3K $\alpha$  protein can lead to cancer. “If we can understand the mechanisms of how specific mutations of this protein cause cancer, we can design drugs that target them,” they explain. With the dawn of personalised medicine, this kind of research could lead to drugs specifically designed for the mutations the cancer of a particular patient.

Studying proteins at this level of detail is easier said than done. The PI3K $\alpha$  protein is a huge molecule made up of about 15,000 atoms, forming ~1,500 amino acids encoded in the DNA. The number of atoms is dramatically increased to 350,000 atoms when water molecules are added around the protein, which helps mimic the physiological environment in which PI3K $\alpha$  is found in the cell. One of its most common oncogenic mutants, E545K, is caused by a change to just one of these amino acids (E545K literally means that the 545th amino acid has been changed from an “E” – glutamic acid – into a



“K” – lysine). To find out how such a small but significant difference in structure affects the function of PI3K $\alpha$ , highly detailed simulations have to be performed to show how both the normal (wild-type) and mutant proteins act variably within the cell.

It is thought that the E545K mutant has a unique mechanism of over activation, as Dr Hari Leontiadou explains: “We believe that one domain of the protein becomes dislocated from the rest of the protein and this triggers over activation of its enzymatic activity. Using long molecular dynamics simulations, we can describe the structural and dynamical differences between the wild type and mutant proteins. Then, using metadynamic simulations, we can calculate the free energy surface of the two variants. This will allow us to identify and characterise the domains or residues that play a key role in the mechanism of over activation.” So far, structural and dynamical differences have been found between the wild type and mutant proteins. The researchers are in the process of investigating these differences and performing the metadynamic

In this figure, PI3K $\alpha$  is shown attached to a model membrane. The catalytic and regulatory subunits of the protein are shown in purple and green, respectively and the membrane is represented in grey. The image was generated with VMD and GIMP.





Dr Zoe Cournia

*“Using the  
PRACE HPC  
resources  
has been  
critical for the  
advancement  
of our work”*

simulations. Initial indications show that the free energy barrier for dislocating one domain is significantly lower in the mutant than in the wild type, meaning that it can be removed more easily and cause over activation.

Using data acquired from these simulations allows the researchers to investigate potential mechanisms by which candidate drugs can bind to the protein. By locating putative allosteric pockets, areas on the protein on which drugs can potentially bind, they can use a method called virtual screening to analyse millions of compounds in order to evaluate which chemicals have the potential to bind to these pockets and cause inhibition of the PI3K $\alpha$  overactivated enzyme. “One allosteric pocket has already been located close to the domain that we think might be causing over activation,” says Dr. Cournia. “We have now identified a few compounds that could potentially inhibit this activity. However, for a compound to be a real candidate drug, a required feature is to be selective only for the mutant protein, leaving the normal protein unaffected.”

The method has already been used successfully in a previous project studying another oncogenic mutant of PI3K $\alpha$ , H1047R, where the 1047th amino acid is changed from a histidine (H) to an arginine (R). “After finding significant structural and dynamical differences between the wild-type form and the mutant, we were able to identify together with our experimental collaborators four compounds that selectively inhibited the over activation of the mutant protein,” says Dr. Paraskevi Gkeka. “We are now in the process of publishing these data, and are hoping to achieve similar success with the PI3K $\alpha$  E545K mutant.” Both studies of the PI3K $\alpha$  protein, are part of a large consortium (POM; PI3K $\alpha$  Oncogenic Mutations), which pursues the research Program “PIK3CA Oncogenic Mutations in Breast and Colon Cancers: Development of Targeted Anticancer Drugs and Diagnostics). The program involves

experimental and theoretical scientists from a variety of disciplines and is funded by the Greek Secretariat for Research and Technology. Once candidate drugs are identified, they are put through a number of optimisation cycles. Candidate drug properties such as potency, the duration of presence in the body or the rate of metabolic turnover can all be manipulated to the desired level through chemical modifications of the structure of the molecule in collaboration with experimental groups. POM is in the process of optimising candidate drugs for the PI3K $\alpha$  H1047R mutant and the results look promising so far.

New data from other researchers on crystal structures of PI3K $\alpha$  are becoming available while the project is running; analysing these data with the help of computers will shed more light on exactly how the protein acts within the body. “We are always looking to exploit new information so that we can build a more complex and realistic representation of the system that we are studying,” says Dr. Cournia. “We will now be applying for more computational time to look at the evolution of the new crystal structures in order to understand how PI3K $\alpha$  retrieves the substrate from the cell membrane”.

“Using the PRACE HPC resources has been critical for the advancement of our work. Recent advances in computer-aided drug design allow us to develop drugs specifically designed for a given protein, shortening the time for development of new drugs. However, such studies cannot progress at a useful speed without supercomputing facilities. Understanding the detailed underlying molecular and atomic interactions involved in drug-protein interactions has increased the efficiency and decreased the cost in the drug discovery process. I believe that our work is a good example of how computers could help develop candidate drugs that have the potential to save millions of lives worldwide.”

