All under one roof – a pioneering vision for research success

Togetherness is a beautiful thing – **Professor Jim Hughes** and **Steve Taylor**, two senior biomedical researchers from the MRC Weatherall Institute of Molecular Medicine at Oxford, discuss the joy of collaboration. In their latest shared project, the pair are developing software that will enable scientists to visualise the three-dimensional structure of a DNA molecule inside a cell.

Sir David Weatherall founded Oxford University's Institute of Molecular Medicine in 1989 with what was then a pioneering vision of getting scientists and clinicians to work together, all under one roof. At the time, the two specialists worked separately – scientists in the lab, doctors in the clinic. Housing them in a joint location allowed the relationship between the two fields to be fully appreciated and provided new opportunities for the cross-fertilisation of ideas. Twenty-eight years on and the renamed MRC Weatherall Institute of Molecular Medicine (MRC WIMM) is now recognised the world over for its innovative biomedical research. With a mixture of scientists and clinicians working side by side, the institute continues to reap the rewards of a collaborative approach to tackling some of the most important questions in medical research. And Sir Weatherall's pioneering vision has become a winning formula adopted by research centres across the globe.

MRC WIMM researchers Professor Jim Hughes (Group Leader of the Genome Biology Group) and Steve Taylor (Head of the Computational Biology Research Group) met up with *Research Features* to discuss life at the institute and to share news on where their own investigative partnership is taking them.

Hello Jim and Steve, welcome to Research Features! Could you tell us briefly about the origins of the MRC WIMM.
Jim: The MRC WIMM was set up in 1989 by Sir David Weatherall. David had a clinical background but he also understood the power of molecular biology and genetics. He thought that clinicians should be tapping

into the kind of scientific developments that were going on, and they needed a way to be trained in proper scientific methods. He also thought that basic biologists should be using the kind of knowledge that could be derived from human disease. He had the idea to house them in one place where they could work coherently together, located beside a hospital, to keep clear clinical links. It was very novel at the time and I think it's proven to be a very successful model. Many institutes around the world now work in the same way. What makes us unique, I think, is that we were the first and we've developed it a long way. It can take a lot to get clinicians and scientists to work together effectively and I think we've got a very good model for doing that.

Could you tell me about your roles at the institute? Let's start with Steve.
Steve: I head up the Computational Biology Research Group, which is a core bioinformatics group. I've been here for about 12 years now, in Oxford, and before that I was in industry. I've always been involved, in some shape or form, in setting up bioinformatics infrastructure and advising people how to do the best bioinformatics analysis.

So what does the Computational Biology Research Group do?
Steve: My team supports scientists, primarily at the Weatherall Institute, but others too. There's a lot of small research groups who don't have bioinformatics support and almost every biological experiment these days will require some informatics input. We work in lots of different fields, such as developing databases and custom software. We also deal with proteomics, microarrays and

next-generation sequencing. Some people imagine that as a core bioinformatics group we are in a kind of service role, browbeaten by the scientists, but actually it's a very collaborative relationship. We work with people to find a solution, and then we get co-authorship on the end result. We also get to collaborate and make interesting tools that we publish in our own right, such as Zegami. In that example, I started work on image analysis for a project at the MRC WIMM, and we needed tools to help us manage large amounts of images. The Zegami software began as a collaboration with Roger Noble, a computer scientist in Australia, and we published and then ended up setting up a business around it. So essentially the MRC WIMM is a space for coming up with new ideas and getting involved in projects.

CSynth allows molecules to be visualised in three dimensions.

images, documents, movies, 3D objects or dashboards in a single field of view that they can search, sort, filter or group in real time. It's a very broad project. It's not just applicable in biology, it's applicable to any area where you have lots of images and metadata and hence is now being used in areas from plant phenomics to human resource management. It's very difficult to get funding in certain areas, especially if you want to develop a new user interface, because you need to see the software in action to understand its value. A while ago, I was giving a talk about an early version of Zegami and someone from Oxford University Innovation saw me doing a demo and asked me if I was interested in doing a spin-out business based on it. It's something I hadn't considered before, but it allowed us to get funding to develop the idea further. Ultimately, we got full funding for the company (http://zegami.com) and there's a lot of advantages in that. Often in informatics, in the academic field, you will begin developing a software tool, initially backed by funding, which then runs out so you can't support it even though lots of people may be using it. But in industry, you're generating revenue and so the situation can be a lot more stable. Oxford University Innovation have organised it so that some of the money from our development projects flows back to the institute, so that should help fund future projects as well.

So developing your research into a business model can actually provide support for your project and help fund future research?
Steve: I think we're going to see a lot more of this way of working. I also think that we will have to be more creative about how we get funding in a post-Brexit UK, as I imagine it will become a lot more restricted.

Jim, could you tell me a bit more about your role and what you're currently working on?
Jim: My role in the MRC WIMM is more of a traditional research group leader. By traditional, I mean in the way that the group is funded, but it's also pretty non-traditional ▶

> **It can take a lot to get clinicians and scientists to work together effectively and I think we've got a very good model for doing that – Professor Jim Hughes**

Because we're the front line of bioinformatics, we tend to get thrown the new, interesting projects. Often we'll get data that people don't really know what to do with, because we've got a whole range of expertise here, especially in the new MRC WIMM Centre for Computational Biology funded by the Medical Research Council (MRC).

You mentioned Zegami, a spin-out company from your research. I'm interested in the relationship between the research that you're doing and how that can be applied and turned into a business model. Could you tell me more about that?
Steve: One of the things I'm keen on is transparency in data analysis and providing tools for people to enable that. Zegami allows users to display tens of thousands of

in that it's a fusion. There's a large part of me that's computational, but I actually have an equal presence at the bench as well. I was a bench scientist for 20 years but then, out of frustration, I learnt to code to become a bioinformatician.

I continued working at a bench, but I also had a desk in what was the original Computational Biology Research Group. That was how I met Steve. It turned out to be a fantastic collaboration which still exists now because we co-manage this facility with the head of another computational group. MRC WIMM's crossover of skills allowed us to translate data into code very quickly and I think that's where our big advantage has been. Outside of this kind of setting, it has been very hard for traditional groups to get their bench data analysed effectively. Often they would have to go and find somebody, try to convince them to analyse their data, they'd analyse it in a way that was not totally correct and then they'd go back; they'd have long conversations to try and literally understand what each other was talking about.

Because I bridged the two fields, my group didn't have those complications. We developed work at the bench inspired by our computational projects; and then we developed computational projects inspired by the fact that we understood exactly how the assays worked.

So you're able to bridge both of those areas that have historically been quite difficult?
Jim: Yes. A lot of my work has been about communication. I have sat in a room with biologists and computational people and I literally translate both ways. Barriers have broken down a little further now, but that's what my original role was in bringing the computational and bench sides together. I was successful in setting up a group in the MRC WIMM itself, a specialist genomics group, interested in how genes are regulated. That's a very basic scientific question. However, because we're in the MRC WIMM, we're very, very conscious of the clinical side of this. I'm funded to try and work out how gene regulation intersects with disease propensity. Essentially, we're using our technology to try and understand the mechanisms behind anaemia, diabetes, and autoimmune diseases. I'm very conscious of how my technologies and abilities intersect with the clinical community, particularly in trying to diagnose mutations which are very inaccessible at the moment. So it's a mixture of basic science, but informed to help and

intersect with clinical science which is what the MRC WIMM's remit is.

You're both involved in the CSynth (http://www.csynth.org) project. Developed in collaboration with Goldsmiths University in London, CSynth is a new interactive software which allows users to visualise the three-dimensional structure of a DNA molecule inside a cell and will integrate genomics data, super resolution microscopic images and polymer modelling. Could you tell me more about that?
Jim: We were trying to understand how long-range gene interactions happen and so had developed the Capture-C technology. Using Capture-C allows researchers to take hundreds of high-resolution pictures of the interactions within a given region of DNA, so a much sharper overall picture of the interaction landscape can be built up. Once we'd got that far, it became very obvious that to try and get any traction on the questions we had, we'd have to try and understand the three-dimensional structure of the nucleus. It's very hard to understand any biological question in 2D; we're essentially 3D animals. For example, it would be very hard to understand how your heart worked in a two-dimensional way. Because Steve and I work very closely together, he was aware of what I was doing on this project.

Steve: I was keen to get involved because this project focuses on visualisation, which I'm really interested in. I'm a microbiologist originally so I'm very interested in biological properties, but I'm also really interested in providing tools that make the science clearer and more obvious.

I attended a visualisation conference a couple of years ago, where Frederic Fol Leymarie was talking about a tool called FoldSynth, which is protein folding software developed by the team at Goldsmiths University. The interesting thing about it was the way the protein was shown as folding dynamically and also the way you could actually see how the 2D interactions related to the 3D space. I got talking to Frederic and that led onto the CSynth collaboration with Goldsmiths. Their group

is involved in computer games programming and I've always thought computer games is a brilliant subject for informatics because the interfaces are often very well thought out and making useable tools for complicated data is key. We had a prototype viewer mocked up using FoldSynth and we loaded some data into it to see what would happen.

Jim: Essentially, if you can model the 3D contacts within a protein, there's no reason you can't model any other 3D contacts. The principles are the same. The aim of our experiment was to see what uses it could be put to. We didn't have any great expectations about how it would work, but the result has been transformational. One problem we had to overcome was that the data was represented statically and of course the system itself is not static; it's highly dynamic and so we needed some way to try and bring dynamics into it. FoldSynth already did that, so we took the cue from there. That allowed us to try and visualise dynamics in a way that the human mind could interact with. We could turn it around and watch it change. After we started up the collaboration, the first thing they did was try to think of ways to visualise these data in a coherent way, and I think that was really successful, although we've got a long way to go in both data generation and code development.

Steve: Yes, integrating all the data together is going to be really instrumental in

This is the sort of research we're really interested in – trying to push those boundaries to enable scientists to work together – Steve Taylor

understanding new things. I'm quite excited that we're going to be using virtual reality, creating a new way of interacting with data that we haven't really done before. You'll literally be able to walk round the molecule and look at it from different angles. You'll also potentially be able to collaborate remotely on the same 3D model. You'll be able to walk round something that looks physically present in the room and say, "What do you think of this structure? Do you think this interaction is correct?" I think that could be fantastically powerful. This is the sort of research we're really interested in – trying to push those boundaries to enable scientists to work together.

How do you see it evolving over the next few years?
Jim: I think we're probably one of the first to try and do this. There are a few other instances, but instead of being very competitive, we've been very collaborative. It needs to be solved, but it's not going to be straightforward or easy and CSynth will help clarify the needs and the goals and promote interest.

The Weatherall Institute as a whole really focuses on collaboration. How significant has collaboration been to your work?

had diseases themselves, trying to understand where the research is going and how it will impact on them. The interest we got was eye-opening.

Jim: I have to agree. I felt quite touched a few times. As Steve said, people would come up and talk to us about certain diseases they had. It always throws you slightly, but if you give your knowledge freely and try to explain what the context is, I think it helps people understand why their life is the way it is. To try and get that kind of information from webpages or books is very hard, whereas if you're just standing there as a human explaining what genetic mutation is, why that would give them a disease, why it may or may not pass on to their children, all those things, just talking about it in a very human way, I think they found that very valuable. It also affected us and clearly brought home to us the human impact of our work.

Steve: I think it was an inspiring event for the next generation of people interested in the sciences. I've been doing computing for nearly 40 years now and when I started I would never have dreamt that I could merge computers and biology. A lot of people were really interested in hearing about that, with regards to their own career paths. I remember it was my nirvana when I found that you can actually do those two things at once.

I first became aware of CSynth at the New Scientist Live Exhibition. You said it was a very hectic and quite exhausting four days, but how important do you think that kind of outreach work with the public is?
Steve: I think it's really important. Gone are the days where you've got people in white coats just beavering away, working on a research paper that's going to be published in some academic journal. I think we've got to get out there and show the public what their money is being spent on. When we explained CSynth to the visitors at the New Scientist Live Exhibition, in London, we got some fantastic comments back from people from all different walks of society. That really makes you appreciate what sort of impact you're having. We had people asking about how they could get into the fields of biology and mathematics, and then at the other end of the spectrum we had people come up to us who

Left: The entrance to the Weatherall Institute.
Right: Professor Jim Hughes (left) and Steve Taylor (right) with some of the equipment from the CSynth project.

Contact

The MRC Weatherall Institute of Molecular Medicine,
University of Oxford,
John Radcliffe Hospital,
Headley Way,
Oxford,
OX3 9DS

T: 01865 222443
W: www.imm.ox.ac.uk