Institution: WestCHEM
Unit of Assessment: Sub-panel 8 – Chemistry
Title of case study: A unique computer technology for the accelerated discovery of drugs that "shape-shift" proteins refocuses and expands a U.S. Drug Discovery company

1. Summary of the impact

A computer technology has been invented to accelerate drug discovery. It predicts locations in disease-associated biomolecules where drug molecules could bind, induce shape changes, and thereby bring the activity of the biomolecule under control. A U.S. drug discovery company, Serometrix, has exclusively licensed this technology and incorporated it within their core discovery process. The impact upon them has been:

- A step change in their technical approach towards small molecule drug discovery,
- Attraction of $227k venture capital funding for the new company direction,
- Expansion of workforce (four new personnel by end of 2013),
- "Shaving years off" their discovery development programme,
- Promising new drug leads,
- Planned reduction of early trial compounds "from millions to hundreds".

2. Underpinning research

Context

Proteins are a major class of biomolecule that are central to virtually all life processes. Our aim was to learn more about protein evolution and how shape and function might be predictable through analysis of the growing databases of sequence and 3D structure. We investigated a range of contrasting theoretical approaches to the data analysis as reported in references 1-4. The outcomes eventually led us to focus on the information content of just the three-dimensional path adopted by the molecular backbone in the folded protein. These fold shapes encode the fundamental characteristics relevant to maintaining functional capability, molecular responsiveness and adaptability through the evolutionary process. From the very large range of possible chain fold shapes, only relatively few have persisted in living organisms, apparently only those fit and "smart" enough for purpose.

The originality of this investigation lay in its focus purely upon the enduring characteristics of the chain folding shapes rather than on the atomic detail of the amino acid residue chemistries along their length.

Key findings

Our results from this research strategy in the late 1990s/early 2000s indicated that a computer technology for predicting important interactive sites in proteins from the chain-fold motif *alone* was possible. The first attempt at a systematic algorithm for site location in enzymes based on the new principle was published in 2003 (Reference 5). Subsequently, after the prototype program was modified, refined and extended in its analytical procedure, given a versatile graphical user interface and made able to process large numbers of protein structures (achieved through collaboration with Dr John Wilson in the Department of Computer and Information Science at Strathclyde), it was realised that the output from the program was also succeeding in predicting additional sites where small molecule binding was known to induce changes in protein shape that regulated activity. As no reliable method existed for predicting sites for shape-altering drugs (also describable as "shapeshifting", topomorphic or allosteric drugs), it was realised that our technology, would have considerable commercial value as an aid to industrial drug discovery, and also be a game-changer in terms of the discovery logic and the quality of the end product.

The majority of small molecule drugs on the market are **competitive** of a protein's key functional site and basically stop the protein working altogether. However, shape-shifting drugs can bind away from the key site (i.e. they are termed non-competitive or uncompetitive) and exert more subtle influences on the bioactivity, bringing about only small, non-critical, changes to the shape of the key site. Subtle adjustment of such sites is the way natural evolution adjusts protein activity via side-chain mutations remote from active sites, while natural allosteric effectors exert remote control of the sites during the lifetime of the protein. Put simply, it is like the difference between an on/off switch and a dimmer switch in lighting control.

This computer technology was one of the very first to offer a readily accessible solution to the problem of systematically detecting shapeshifting sites in proteins. The desirability of moving towards drugs that have an allosteric/topomorphic action has only come to the fore in the last few years as successful examples of such drugs have been brought to market. Thus our technology was brought to fruition at a key point in the changing strategic direction of the drug discovery industry. After a long period of testing, refinement and preparation for high throughput industrial use, the technology was publicised (Reference 6) and championed with support from the University's Research and Knowledge Exchange Services.

Key researchers
Lead researcher: Dr Mark Dufton, Senior Lecturer, WestCHEM (appointed 1984, Senior Lecturer from 1994 - present).

Collaborating researcher: Dr John Wilson, Senior Lecturer, Department of Computer and Information Science, University of Strathclyde (appointed 1985, Senior Lecturer from 2002 - present).

3. References to the research

The three references that best illustrate the quality of the research are numbers 2, 4 and 5 below.

[1] Foci of amino acid residue conservation in the 3D structures of the Kunitz BPTI inhibitors. Cardle, L. & Dufton, M. (1997) *Protein Engineering,* 10, 131-136. DOI:10.1093/protein/10.2.131

[2] Evolutionary trace analysis of the Kunitz/BPTI family of proteins. Pritchard, L. & Dufton, M. (1999) *Journal of Molecular Biology,* 285, 1589-1607. DOI: 10.1006/jmbi.1998.2437.

ISSN: 0022-2836

[3] Do proteins learn to evolve? The Hopfield Network as a basis for the understanding of protein evolution. Pritchard, L. & Dufton, M. (2000) *Journal of Theoretical Biology,* 202, 77-86.
DOI: 10.1006/jtbi.1999.1043

[4] Evaluation of a novel method for the identification of coevolving protein residues. Pritchard, L., Bladon, P., Mitchell, J. & Dufton, M. (2001) *Protein Engineering,* 14, 549-555.
DOI: 10.1093/protein/14.8.549

[5] Simple intrasequence difference (SID) analysis: an original method to highlight and rank sub-structural interfaces in protein folds. Pritchard, L., Cardle, L., Quinn, S. & Dufton M. (2003) *Protein Engineering,* 16, 87-101. DOI: 10.1093/proeng/gzg012

[6] SID Technology software, as reported in: The Drug Discovery Portal: A Computational Platform for Identifying Drug Leads from Academia. Clark, R., Johnston, B., Mackay, S., Breslin, C., Robertson, M., Sutcliffe, O., Dufton, M., & Harvey, A. (2010) *Current Pharmaceutical Design*, 16, 1697-1702. DOI: 10.2174/138161210791164018

4. Details of the impact

In 2011, the U.S. Drug Discovery company, Serometrix, was attracted by the University's championing of our computer technology for the accelerated discovery of allosterically acting drugs. The company operates internationally between other university research teams and the very large pharmaceutical companies to identify promising drug leads. At the time, the company was focused on developing drugs for the lowering of cholesterol levels and had been working on one particular protein target. This target was proving intractable to the usual drug discovery process and Serometrix had already concluded that some kind of "shapeshifting" (i.e. allosterically acting) drug was required.

When we were invited to analyse the target, the company already had its own confidential direct experimental data about the location of sites where small molecules could bind and produce "shapeshifting" (i.e. they challenged us with a "blind trial"). On receiving our theoretical predictions as to suitable site locations (calculated by our technology in seconds) and comparing them with their own findings (achieved after significant expenditure of their time and money on conventional investigation), the company immediately opened negotiations with the university to develop a research contract, exclusively license our technology and obtain more calculations on other targets (our technology is applicable to all 86,000 structures in the Protein Data Bank).
The CEO of the company said at the time of the initial trial work (Source 1):

> *"What I find extremely exciting about this entire process is that if you look at the time required to identify the spot to begin working, through to our current understanding as of today – it is mind boggling! If you compare what we have jointly accomplished with these tools, as compared to any other drug discovery platform known to man at this time, we may have just improved the efficiency of this process by several orders of magnitude. It is clear to me that your program is providing accurate predictions about regions where these sites might exist."*

Since then, the company has concluded our technology is "best in class" (Source 3) and has:

- Contracted with the University of Strathclyde for exclusive access to the technology (Sources 3, 6 and 7). The contract and license includes a multi-million buyout option and requires the university to provide the company with secure on-line access to the technology for 20 years.
- Used its exclusive access to the new technology to help secure U.S. venture capital funding of $228k in June 2012 (Source 4) and pursued discussions leading to further venture capital funding, final agreement for which is expected to be concluded shortly (Source 2).
- Hired a new Business Development Manager and Medicinal Chemist to take full advantage of our technology as quickly as possible (two further appointments are due in late 2013).
- Applied the technology to accelerate progress on its current portfolio of drug targets with the intention of eventually applying it to *all* proteins of interest as drug targets (including revisiting those that have been previously investigated by the pharmaceutical industry).
- "Shaved years off" the first stage of its core discovery process and reduced its planned number of early trial compounds "from millions to hundreds" (Source 3).
- Discovered new active lead compounds via the first site predictions provided by the university in 2011. This work is focused on PCSK9, a high value commercial target for treatment of hypercholesterolemia and arthrosclerosis (Source 2).
- More widely, the company has established primary areas of therapeutic interest including oncology, cardiovascular, metabolic, central nervous system and infectious disease and have identified opportunities for the following families of compounds (Source 5):
  - SX-ARPC is a novel family of AR pathway antagonists that have potential as a novel therapeutics for Prostate Cancer.
  - SX-RDS1 is a novel set of DNA repair inhibitors discovered to enhance the effectiveness of radiation therapy in a non-toxic manner.

- SX-HIV1 is a family of retroviral integrase inhibitors with applications for a range of human retroviruses including HIV.

Drug site predictions from our technology have been used by the company to fashion actual drug lead compounds that target the sites. This means that the company's employees and their collaborators are now actively engaged in experimental work for which both the direction and application have been determined by our technology. Their intention is to develop drug leads that can be licensed forward to the larger companies and then pursue new targets/leads. As this collaboration matures, the reach will extend beyond the company to larger concerns and eventually to those who will benefit from the drugs.

Our tool for the discovery of **drugs that "shapeshift" their molecular targets** is showing how to up-regulate and down-regulate protein/enzyme action *without abolishing it*, and this is a paradigm shift from classic drug action. The newer drugs can also be *much more selective* amongst a family of related targets because they are not restricted to highly conserved sites, leading to lower dosing requirements and fewer toxic side-effects. This innovative industrial collaboration and on-line service comes at a time when much of the drug discovery industry is embroiled in a continuing saga of expensive failure and temptation to fraud through its adherence to conventional wisdom in terms of drug action.

By helping to automate the process, bringing costs down and prompting production of less risky drugs, our technology breaks the current log-jam in the drug discovery market and makes it more likely that small market drugs (i.e. for small groups of sufferers or developing world populations) will become viable.

5. Sources to corroborate the impact

1. Statement from the CEO of US drug discovery company Serometrix corroborates his comments following the initial trials undertaken.

2. Statement from the CEO of Serometrix corroborates the claims regarding the company's licencing and implementation of the technology within the company including initial drug targets, impact on company activities and benefits arising.

3. http://www.strath.ac.uk/press/newsreleases/headline_737125_en.html corroborates the partnership between the university and Serometrix and the comments by the Serometrix CEO on the impact that the technology has on their business.

4. Who Got Funded web-site (www.whogotfunded.com) corroborates the venture capital funding raised by Serometrix in June 2012 (Registration is required to access this web-site; a print out of the relevant information is available from the university).

5. http://www.serometrix.com/pipeline.html corroborates the company's primary areas of therapeutic interest.

6. http://pharma.financialexpress.com/latest-updates/2574-scientists-from-university-of-strathclyde-invents-technology-to-treat-major-illnesses corroborates the take-up of the technology by Serometrix.

7. http://drugdiscovery.pharmaceutical-business-review.com/news/strathclyde-researchers-invent-computer-program-for-drug-discovery-280813 corroborates the take-up of the technology by Serometrix.