Neil Swainston

BSc (Hons) MSc PhD MRSB MRSC

✉ 127 Buckingham Road, Chorlton, Manchester M21 0RG
@ neil.swainston@gmail.com

Summary

- Bioinformatics, cheminformatics, scientific software engineering, machine learning, computational systems and synthetic biology in academic and industrial environments.
- BSc Chemistry with Industrial Experience (first); MSc Computing Science; PhD Computer Science.
- Professional software engineering skills.
- Team Leading and Project Management in industry.
- International research experience in Germany and France.
- Lecturing and undergraduate and postgraduate student supervision.
- Well published and cited (52 peer-reviewed articles, 11 with >100 citations; h-index: 26).
- Conference organising experience.
- Multiple successful grant applications (>£1.9M funding).

Statement

I am a computational biologist, with experience in both the commercial and academic sectors. With interests in 'omics data analysis, systems and synthetic biology, I have published over 50 peer-reviewed papers in these fields.

My interests are in industrial biotechnology applications, including the integration of multiscale approaches across host, metabolic pathway and enzyme engineering approaches to enable sustainable chemical production through microbial cell factories. I have, and continue to develop, computational approaches to support the full design-build-test-learn cycle of synthetic biology, covering DNA design and optimisation, lab automation, 'omics data analysis and machine learning. These approaches have been applied towards generating cell factories for a wide range of chemicals covering diverse chemical classes.

During my research career, I have worked exclusively in interdisciplinary teams, working across traditional disciplines including biology, chemistry, mathematics and computer science.

I am also involved in teaching and PhD student supervision responsibilities and have successful grant writing experience. I have received funding of over £1.8 million for projects in cheminformatics, metabolic modelling, text mining and enzyme engineering with a range of partners including the European Bioinformatics Institute, the National Centre for Text Mining and industrial collaborators.

Employment

07/19 -   **University of Liverpool** Senior Research Associate
            Institute for Integrative Biology

Development of improved bioinformatics methods in the full Design-Build-Test-Learn cycle applied to directed evolution in enzyme engineering. Application of machine learning approaches to 'omics data analysis in metabolomics and drug transport.

04/06 – 06/19   **Manchester Institute of Biotechnology** Senior Experimental Officer
                  Centre for Synthetic Biology of Fine and Specialty Chemicals (SYNBIOCHEM)

Recent work is in computational synthetic biology, gene design and synthesis, and directed evolution as applied to enzyme optimisation and biocatalysis. This work covers computational aspects including DNA design optimisation algorithms, lab automation for DNA assembly and construct validation with next generation sequencing. Further work involves application of machine learning applied to DNA and amino acid sequence analysis.

My experience in systems biology involved the co-leading of an international community focussed on the development of comprehensive predictive models of metabolism in human, yeast and other organisms, and this work has been both highly cited and publicised. Further work included the improvement of metabolic modelling predictions through integration of experimental data.

Other areas of expertise in experimental data analysis and integration cover next-generation sequencing, quantitative proteomics, metabolomics and enzyme kinetics.

04/99 – 04/06   **Waters Corporation** Bioinformatics Team Manager

Commercial software development, as applied to bioinformatics, proteomics and mass spectrometry.

My role progressed from software engineer, through to a team leader and ultimately a project manager, directly managing a group of five people and being responsible for a software release involving the work of nine developers.

This role was interdisciplinary, covering all facets of the software development life cycle, from requirements gathering, software design, development and documentation, though to testing and support. Due to the focus on requirements gathering from users, and providing software support, I was frequently involved in customer site visits, user training, and conference speaking; tasks which greatly developed my skills of collaborating with biologists and chemists.

10/98 - 04/99   **AstraZeneca** Graduate Trainee

Graduate Trainee in IT problem management and data analysis.

09/94 - 08/95   **Dow Chemical Company, Stade, Germany** Student Placement

Year long placement in a varied role involving polycarbonate synthesis and analytical chemistry. Duties involved lab research, plant visits, provision of analytical chemistry support, and development of a searchable database resource of spectra of chemical standards, presentation giving in both English and German.

Education

03/12	**PhD** "Systems biology informatics for the development and utility of genome-scale metabolic models". *University of Manchester*
09/97 – 10/98	**MSc Computing Science** *University of Newcastle-upon-Tyne; IRISA, Rennes, France*. Year long conversion course including Masters project in bioinformatics (DNA sequence analysis), undertaken in Rennes, France.
09/92 – 06/96	**BSc (Hons) Chemistry with Industrial Experience** *University of Manchester* First class honours; industrial experience in analytical chemistry with Dow Deutschland Inc., Stade, Germany.

Selected publications

- A repository for quality-assured data for enzyme activity. Swainston N, Kettner C. *Nature*. 2018, **556**, 309.
- PartsGenie: an integrated tool for optimising and sharing synthetic biology parts. **Swainston N**, et al. *Bioinformatics*. 2018, **34**, 2327-2329.
- Engineering the "Missing Link" in Biosynthetic (−)-Menthol Production: Bacterial Isopulegone Isomerase. Currin A, et al. *ACS Catal*. 2018, **8**, 2012–20.
- Recon 2.2: from reconstruction to model of human metabolism. **Swainston N**, et al. *Metabolomics*. 2016, **12**, 1-7.
- Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. Currin A, **Swainston N**, Day PJ, Kell DB. *Chem Soc Rev*. 2015, **44**, 1172-239.
- GeneGenie: optimised oligomer design for directed evolution. **Swainston N**, et al. *Nucleic Acids Res*. 2014, **42**:W395-400.
- A community-driven global reconstruction of human metabolism. Thiele I, **Swainston N**, et al. *Nat Biotechnol*. 2013, **31**, 419-25.
- Improving metabolic flux predictions using absolute gene expression data. Lee D, Smallbone K, Dunn WB, Murabito E, Winder CL, Kell DB, Mendes P, **Swainston N**. *BMC Syst Biol*. 2012, **6**:73.
- A consensus yeast metabolic network obtained from a community approach to systems biology.

Herrgård MJ, **Swainston N**, et al. *Nat Biotechnol*. 2008, **26**, 1155-1160.

Funding

- MCR BioFactory – a synthetic biology engine for the rapid discovery and engineering of high value proteins for the biotechnology industry. *Innovate UK; Innovation to Commercialisation of University Research (ICURe) Programme; 2019*. **£35,000**.
- GeneORator: a novel and high-throughput method for the synthetic biology-based improvement of any enzyme. *BBSRC; Follow on Fund; 2019*. **£249,678**. (BB/S004955/1; Co-Investigator)
- GeneORator – MCR Biofactory. *University of Manchester Intellectual Property (UMIP) Proof-of-Principle, with AB Vista; 2019*. **£178,260**. (20160050; Co-Investigator).
- Enriching Metabolic PATHwaY models with evidence from the literature (EMPATHY). *BBSRC; Responsive mode; 2014*. **£659,535**. (BB/M006891/1; Co-Investigator).
- Modelling and sensitivity analysis of metabolic networks in diabetic neuropathy. *Faculty of Life Sciences, University of Manchester, Interdisciplinary Projects in Quantitative Biology; 2014*. **£46,798**.
- Hackathon on Resources for Modelling in Biology 2014 (HARMONY 2014). *BBSRC; International workshops; 2014*. **£9047**. (BB/L026325/1; Co-Applicant).
- The relationship of clusters of gene expression associated with development in childhood disease in the ageing adult. *Manchester Institute for Collaborative Research on Ageing (MICRA) Seedcorn Funding; 2014*. **£6000**.
- Towards an integrated model of human metabolism, cell signalling and gene expression. *University of Manchester; Faculty of Engineering and Physical Sciences Strategic Fund; 2014*. **£4000**.
- Continued development of ChEBI towards better usability for the systems biology and metabolic modelling community. *BBSRC; Bioinformatics and biological resources fund; 2013*. **£682,950**. (BB/K019783/1; Co-Investigator).
- Iron metabolism and its role in neurodegenerative disease. *University of Manchester; Faculty of Engineering and Physical Sciences Strategic Fund; 2013*. **£5200**.

Teaching and student supervision

- **Teaching**: Bioinformatics, Proteomics, Systems Biology. Systems Biology Doctoral Training Centre, MRes Translational Medicine, University of Manchester, 2006-10, 2015. Systems Biology. MSc (Res) Translational Oncology, University of Sheffield, 2014-. Data integration and Interaction Networks. MSc Applied Bioinformatics, MSc Molecular Medicine, Cranfield University, 2011-15. Industrial Biotechnology, Coursera, 2017. (https://www.coursera.org/learn/industrial-biotech). PythonClub, University of Manchester, 2018-19.
- **Tutor**: In Silico Systems Biology: Network Reconstruction, Analysis and Network-based Modelling. EBI-EMBL, Hinxton, Cambridge, 23-26 May 2011.
- **Student supervisor**: Systems Biology Doctoral Training Centre, University of Manchester, 2007-12; MSc Applied Bioinformatics, Cranfield University, 2013; MSc Advanced Computer Science, University of Manchester, 2014; ALM Biotechnology, Harvard University Extension School, 2014.
- **Advisor**: University of Manchester iGEM Advisor, 2013-15.

Scientific responsibilities

- Lead conference organiser: HARMONY 2014, Manchester, 22-25 April 2014; STRENDA 2018, Manchester, 18-19 September 2018.
- Associate Editor, BMC Systems Biology.
- Invited member, Synthetic Biology Open Language (SBOL) Developers Group.
- Invited member, STRENDA Commission.
- Session chair. COMBINE 2010, Edinburgh.
- Scientific Committee member, International Symposium on Integrative Bioinformatics, 2010-14.
- Course Office member, FEBSX-SysBio2011: From Molecules To Function. Innsbruck, Austria.
- Nominated for election to position of SBML Editor, 2010-12 and 2011-13.

- Journal reviews, including: ACS Synthetic Biology, Bioinformatics, Nature Biotechnology, Nature Protocols, PLoS Computational Biology.
- Funding reviews: BBSRC, Breast Cancer Campaign.
- PhD external examiner: Kings College, London; University of Sheffield.

Computational and software development skills

- Languages: Java, Python, Matlab, C/C++.
- Web development: Javascript, ajax, JQuery, AngularJS, Bootstrap, Google Web Toolkit, Flask.
- Data science: numpy, scipy, scikit-learn, pandas, keras, Tensorflow.
- Cloud computing: Google Compute Engine.
- Data management: XML, JSON, relational, XML and graph (neo4j) databases.
- Semantic web: RDF.
- Software engineering: source code control (svn, git), build and deployment scripts (ant, Maven, Docker), software design with UML.
- Operating systems: OS X, Windows, Linux.

Courses and training

- **University of Manchester**: Academic career planning; Practical project management for PIs and academics; Collaboration: the hidden research skill; Communicating with your teams; Getting the best out of yourself and others; Coaching and mentoring skills for supervisors; Practical application of supervision skills; How to shine at fellowship interviews.
- **Coursera**: Introduction to Genetics and Evolution; Calculus One; Astrobiology and the Search for Extraterrestrial Life; Machine Learning.
- **Peter Kenyon**: The Front Line Manager.
- **Pentland Training**: Project Management Fundamentals.

References

Prof Douglas B Kell CBE
✉ University of Liverpool, Liverpool, UK
☎ +44 (0)151 795 7772
@ dbk@liv.ac.uk

Prof Markus Herrgard
✉ Novo Nordisk Foundation Center for Biosustainability, DK
☎ +45 24 92 17 80
@ herrgard@biosustain.dtu.dk

Prof Pedro Mendes
✉ UConn Health, Farmington, CT 06030-6033, USA
☎ +1 860 679 7632
@ pmendes@uchc.edu

Prof Hans V Westerhoff
✉ University of Amsterdam, 1090 GE Amsterdam, NL
☎ +31 (0)20 525 5150
@ h.v.westerhoff@uva.nl

Publications

1. GeneORator: an effective strategy for navigating protein sequence space more efficiently through Boolean OR-type DNA libraries. Currin A, et al. *ACS Synth Biol.*, 2019, **8**, 1371-8.
2. Machine Learning of Designed Translational Control Allows Predictive Pathway Optimization in Escherichia coli. Jervis AJ, et al. ACS Synth Biol., 2019, **8**, 127-136.
3. Multifragment DNA Assembly of Biochemical Pathways via Automated Ligase Cycling Reaction. Robinson CJ, et al. *Methods Enzymol*. 2018, **608**, 369-392.
4. Fast and Flexible Synthesis of Combinatorial Libraries for Directed Evolution. Sadler JC, et al. *Methods Enzymol*. 2018, doi: 10.1016/bs.mie.2018.04.006.
5. Rationalizing context-dependent performance of dynamic RNA regulatory devices. Kent R, et al. *ACS Synth Biol*. 2018, doi: 10.1021/acssynbio.8b00041.
6. An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. Carbonell P et al. *Nat Commun Biol*. 2018, **1**, 66.
7. A repository for quality-assured data for enzyme activity. **Swainston N**, Kettner C. *Nature*. 2018, **556**, 309.
8. STRENDA DB: enabling the validation and sharing of enzyme kinetics data. **Swainston N**, et al. *FEBS J*. 2018, **285**, 2193-2204
9. PartsGenie: an integrated tool for optimising and sharing synthetic biology parts. **Swainston N**, et al. *Bioinformatics*. 2018, **34**, 2327-2329.
10. Selenzyme: enzyme selection tool for pathway design. Carbonell P, et al. *Bioinformatics*. 2018, bty065.
11. Engineering the "Missing Link" in Biosynthetic (−)-Menthol Production: Bacterial Isopulegone Isomerase. Currin A, et al. *ACS Catal*. 2018, **8**, 2012–20.
12. biochem4j: integrated and extensible biochemical knowledge through graph databases. **Swainston N**, et al. *PLoS ONE*. 2017, **12**:e0179130.
13. CodonGenie: optimised ambiguous codon design tools. **Swainston N**, et al. *PeerJ Computer Science*. 2017, **3**:e120.
14. Identifiers for the 21st century: How to design, provision, and reuse persistent identifiers to maximize utility and impact of life science data. McMurry JA, et al. *PLoS Biol*. 2017, **15**:e2001414.
15. SpeedyGenes: Exploiting an Improved Gene Synthesis Method for the Efficient Production of Synthetic Protein Libraries for Directed Evolution. Currin A, et al. *Methods Mol Biol*, 2017, **1472**, 63-78.
16. SYNBIOCHEM Synthetic Biology Research Centre, Manchester–A UK foundry for fine and speciality chemicals production. Le Feuvre RA, et al. *Biochem Soc Trans*. 2016, **44**, 675-7.
17. Recon 2.2: from reconstruction to model of human metabolism. **Swainston N**, et al. *Metabolomics*. 2016, **12**, 1-7.
18. SYNBIOCHEM-a SynBio foundry for the biosynthesis and sustainable production of fine and speciality chemicals. Carbonell P, et al. *Biochem Soc Trans*. 2016, **44**, 675-7.
19. Bioinformatics for the synthetic biology of natural products: Integrating across the Design-Build-Test cycle. Carbonell P, et al. *Nat Prod Rep*. 2016, **33**, 925-32.
20. libChEBI: an API for accessing the ChEBI database. **Swainston N**, et al. *J Cheminform*. 2016, **8**:11.
21. ChEBI in 2016: Improved services and an expanding collection of metabolites. Hastings J, et al. *Nucleic Acids Res*. 2016, **44**, D1214-9.
22. SBOL Visual: A Graphical Language for Genetic Designs. Quinn JY, et al. *PLoS Biol*. 2015, **13**:e1002310.
23. Membrane transporter engineering in industrial biotechnology and whole-cell biocatalysis. Kell DB, et al. *Trends Biotechnol*. 2015, **33**, 237-246.
24. RobOKoD: microbial strain design for (over)production of target compounds. Stanford NJ, et al. *Front Cell Dev Biol*. 2015, **3**:17.
25. Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. Currin A, et al. *Chem Soc Rev*. 2015, **44**, 1172-239.
26. A 'rule of 0.5' for the metabolite-likeness of approved pharmaceutical drugs. O'Hagan S, et al. *Metabolomics*. 2015, **11**, 323-339.
27. SpeedyGenes: an improved gene synthesis method for the efficient production of error-

corrected, synthetic protein libraries for directed evolution. Currin A, et al. *Protein Eng Des Sel*. 2014, **27**, 273-80.

28. GeneGenie: optimised oligomer design for directed evolution. **Swainston N**, et al. *Nucleic Acids Res.* 2014, **42**, W395-400.

29. Path2Models: Large-scale generation of computational models from biochemical pathway maps. Büchel B, Rodriguez N, **Swainston N**, Wrzodek C, et al. *BMC Syst Biol.* 2013, **7**:116.

30. An analysis of a 'community-driven' reconstruction of the human metabolic network. **Swainston N**, et al. *Metabolomics*. 2013, **9**, 757-764.

31. A model of yeast glycolysis based on a consistent kinetic characterization of all its enzymes. Smallbone K, Messiha H, et al. *FEBS Lett.* 2013, **587**, 2832-41.

32. A community-driven global reconstruction of human metabolism. Thiele I, **Swainston N**, et al. *Nat Biotechnol*. 2013, **31**, 419-25.

33. Improving metabolic flux predictions using absolute gene expression data. Lee D, Smallbone K, Dunn WB, Murabito E, Winder CL, Kell DB, Mendes P, **Swainston N**. *BMC Syst Biol*. 2012, **6**:73.

34. The SuBliMinaL Toolbox: automating steps in the reconstruction of metabolic networks. **Swainston N**, et al. *J Integr Bioinform.* 2011, **8**:186.

35. Sustainable Model Building: The Role of Standards and Biological Semantics. Krause F, et al. Methods Enzymol. 2011, **500**, 371-95.

36. A community effort towards a knowledge-base and mathematical model of human pathogen Salmonella Typhimurium LT2. Thiele I, et al. *BMC Syst Biol.* 2011, **5**:8.

37. A QconCAT informatics pipeline for the analysis, visualization and sharing of absolute quantitative proteomics data. **Swainston N**, et al. *Proteomics* 2011, **11**, 329–333.

38. Systematic integration of experimental data and models in systems biology. Li P, et al. *BMC Bioinformatics* 2010, **11**:582.

39. Further developments towards a genome-scale metabolic model of yeast. Dobson PD, Smallbone K, et al. *BMC Syst Biol.* 2010, **4**:145.

40. Enzyme kinetics informatics: from instrument to browser. **Swainston N**, Golebiewski M, et al. *FEBS J.* 2010, **77**, 3769–3779.

41. Integrative Information Management for Systems Biology. **Swainston N**, et al. *In proceedings of the 7th International workshop on Data Integration in the Life Sciences 2010 (DILS'10), Gothenburg, Sweden. Lecture Notes in Computer Science* 2010, **6254**, 164-178.

42. Integration of metabolic databases for the reconstruction of genome-scale metabolic networks. Radrich K, et al. *BMC Syst Biol.* 2010, **4**:114.

43. Towards a genome-scale kinetic model of cellular metabolism. Smallbone K, et al. *BMC Syst Biol.* 2010, **4**:6.

44. Information management for high content live cell imaging. Jameson D, et al. *BMC Bioinformatics* 2009, **10**:226.

45. Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. Brown MC, et al. *Analyst* 2009, **134**, 1322–1332.

46. libAnnotationSBML: a library for exploiting SBML annotations. **Swainston N**, Mendes P. *Bioinformatics* 2009, **25**, 2292–2293.

47. A consensus yeast metabolic network obtained from a community approach to systems biology. Herrgård MJ, **Swainston N**, et al. *Nat Biotechnol.* 2008, **26**, 1155-1160.

48. Capture and analysis of quantitative proteomics data. Lau K, et al. *Proteomics* 2007, **7**, 2787-99.

49. Growth control of the eukaryote cell: A systems biology study in yeast. Castrillo JI, et al. *J Biol.* 2007, **6**:4.

50. An informatic pipeline for the data capture and submission of quantitative proteomic data using iTRAQ. Siepen JA, **Swainston N**, et al. *Proteome Sci.* 2007, **5**:4.

51. Model-driven User Interfaces for Bioinformatics Data Resources: Regenerating the Wheel as an Alternative to Reinventing It. Garwood K, et al. *BMC Bioinformatics* 2006, **7**:532.

52. Synthesis and redox properties of the cycloheptatrienylmolybdenum complexes [MoX(N-N)(h-C_7H_7)]$^{z+}$, (N-N = 2,2¢-bipyridine or 1,4-Bu$_2$t-1,3-diazabutadiene; z = 0, X = Br or Me; z = 1, X = NCMe, CNBut or CO). Disley SPM, et al. *J. Organomet. Chem*. 1998, **566**, 151-158.