



“In order to translate the academic paper into a working model we needed a partner that could compliment our knowledge of pharmaceutical research with proven expertise in scientific software”.

Darren Green,
Director of
Computational
Chemistry,
GSK’s Technology
Development
Department

GlaxoSmithKline optimize drug discovery with Tessella software expertise

When GlaxoSmithKline (GSK) needed to turn academic research on genetic algorithm optimization into a working model they called on Tessella’s scientific software expertise. The resulting solution has the potential to help reduce the time and effort required to identify new drug candidates, and has enabled GSK to contribute to further research into algorithm-directed lead optimization through the publication of a study and full data set which compares the results achieved by the algorithm versus manual synthesis.

Background and Challenge

GlaxoSmithKline (GSK) is a world-leading research-based pharmaceutical company, with a mission to improve the quality of human life through the development of innovative medicines, vaccines and other healthcare products. Headquartered in the UK, with operations in the US, GSK has an estimated seven per cent share of the world’s pharmaceutical market and invests more than \$500,000 dollars every hour into the research and development of new medicines for asthma, virus control, infections, mental health, diabetes and digestive conditions. GSK is also the company behind many market leading brands, such as all, Gaviscon, Panadol, Aquafresh, Sensodyne, Lucozade and Nicorette.

Darren Green, Director of Computational Chemistry at GSK’s Technology Development Department explains further, “At the heart of our business is the need to discover compounds that are biologically

active in order to develop them into potential new drug candidates. This led us to investigate new platforms for the synthesis and bioassay of novel compounds. These platforms were based on microfluidic technology, and hence consumed much smaller quantities of materials than conventional systems allowing the results to be obtained more rapidly.”

In order to create a fully-integrated closed-loop synthesis-assay system, an optimization module was required, which would contain a selection algorithm that would decide which compound – of a potentially very large pool of ‘virtual’ compounds – to make in the next synthesis/assay iteration.

The algorithm would also be able to base its selections on data from all previous iterations, and thus discover compounds with high biological activity as quickly as possible.



Solution and Benefits

Darren continues, "Based on existing academic literature on optimizing searches in combinatorial chemistry, we chose to use a genetic algorithm as the basis for one of the optimization modules."

Genetic algorithms use strings of letters to represent individuals in a population, and mimic biological processes by crossing the strings of individuals and applying other genetic processes like mutation. The result is that each generation adapts and becomes fitter than the previous, by a process similar to natural selection.

Having tested a number of products their fitness can be assessed by the experimental results and the genetic algorithm used to produce the next generation of products to be tested. As the process is repeated the level of measured biological activity will increase.

Darren continues, "In order to translate the academic paper into a working model we needed a partner that could compliment our knowledge of pharmaceutical research with proven expertise in scientific software." Tessella were selected because of their experience of developing similar optimization systems.

They worked closely with the GSK team to rapidly build on the basic algorithms documented in the paper. The resulting solution is a Genetic Algorithm Optimizer that not only meets the original use case, but also provides a flexible optimizer which can be applied to other scientific problems. The system is able to optimize a wide range of search domains including discrete, continuous and grouped variables, allowing searches to be optimized for both combinatorial and parallel synthesis.

Darren continues, "No optimization system can replace the expertise and creativity of an experienced chemist. However, the Genetic Algorithm Optimizer has been designed to allow chemists to enhance their searches by providing chemically relevant information about the reagents and products. The system also provides feedback that assists the chemists to understand better the

underlying chemistry and why certain products are biologically active."

The close collaboration between the Tessella and GSK teams means the system has a number of innovative and flexible features, including the ability for chemists to identify significant groups of reagents; to specify the level of similarity between reagents or between products; and to manage experimental errors. In addition, the system automatically and dynamically adjusts its search strategy to maximise efficiency.

Results

After validation runs using historic or in silico generated data, the new software module was used in a real optimization experiment to iteratively select ten generations of compounds for synthesis and assay.

Darren continues, "In addition, we conducted a comparison study of the results from the optimizer software against a full combinatorial library that was synthesised manually and tested independently. The results show that the compounds selected by the algorithm are heavily biased toward the more active regions of the library. The algorithm is also robust to both missing data (where synthesis failed) and in active compounds. All this demonstrates that, in combination with suitable hardware, statistical analysis and visualization tools, the Genetic Algorithm Optimizer is a very effective tool for lead optimization."

GSK have published the study, including the full combinatorial and biological data sets, into the public domain as a paper entitled "Automated Lead Optimization of MMP-12 Inhibitors Using a Genetic Algorithm" in order to help further advance research into algorithm-directed lead optimisation.

Tessella plc 26 The Quadrant, Abingdon Science Park, Abingdon, Oxfordshire OX14 3YS, UK
T: +44 (0)1235 555511 | F: +44 (0)1235 553301 | E: info@tessella.com

Tessella Inc 233 Needham Street, Suite 300, Newton, MA 02464, USA
T: 1 617 454 1220 | E: info@tessella.com

Tessella – successfully delivering IT and consulting services to world leaders in R&D, science and engineering.

Founded in 1980, Tessella is the international provider of science powered technology and consulting services. World leading organizations choose our unique blend of science, engineering and sector expertise to deliver innovative and cost-effective solutions to complex real-world commercial and technical challenges. Our people are high achievers from leading universities and are passionate about delivering value to clients. We are proud that our work makes the world a better place to live in: developing smarter drug trials; preserving the digital heritage of nations across the globe; minimizing risk in oil and gas exploration; controlling the orbit and attitude of satellites; researching fusion energy.

Copyright © Tessella plc 2011, all trademarks acknowledged. Issue: V1.R0.M0 | January 2011



www.tessella.com