Beilstein-Institut

Systems Chemistry, May 26th – 30th, 2008, Bozen, Italy

Modelling for Regenerative Medicine: Systems Biology Meets Systems Chemistry

David A. Winkler*, Julianne D. Halley, Frank R. Burden

CSIRO Molecular and Health Technologies,
Private Bag 10, Clayton 3168, Australia

E-Mail: *dave.winkler@csiro.au

Received: 5th September 2008 / Published: 16th March 2009

Introduction

Complex systems science is making substantial contributions to the study of biological systems, and has made a substantial contribution to the new field of systems biology. Systems biology focuses on the systematic study of complex interactions in biological systems using an integrative rather than reductionist perspective. One of the goals of systems biology is to study, model, and understand new emergent properties of biological systems from a complex systems perspective [1, 2]. This integrative approach to biology is generating substantial benefits in facilitating study of larger more complicated systems, providing improved understanding of nonlinear system properties, and provides an ability to model systems at appropriate levels of detail where the model is matched to data density and research questions. Various aspects of systems biology have been reviewed recently [3 – 11]. Chemistry has lagged behind most other disciplines in adopting complex systems approaches, possibly because it has largely been a reductionist science, and reductionist approaches have been very successful. Adopting a complementary complex systems approach to chemistry will build on this success to study more complex matter.

Paradigm Shifts

Chemistry focuses on the synthesis and properties of relatively small sets of molecular species but is now increasingly embracing the generation and study of larger, more diverse systems of molecules [12]. Chemistry is becoming increasingly multidisciplinary, embracing important new research fields where chemistry, physics, biology and materials science

overlap. Emerging research fields such as nanomaterials and self-assembly will benefit from an integrative, complex systems approach that is complementary to the reductionist methods that characterize the discipline of chemistry [13]. Reductionist and deterministic concepts like proteins and ligands as rigid 'lock and key' systems, cells as molecular machines, and proteins as relatively rigid structures are increasingly at odds with new spectroscopic data, computational experiments, and high throughput experimental results. An adaptive, dynamic description of small molecules (chemistry), and larger molecules (biology but now increasingly chemistry) is required. As Kurakin elegantly states [14], there is a global crisis of the mechanistic, deterministic paradigm in life sciences that a complex systems description of molecular processes can resolve. In his classic work on the nature of scientific revolutions, Kuhn states that the success of a dominant paradigm brings about its own crisis and necessitates a paradigm shift, as the advances in technology and methods lead to a widespread accumulation of experimental data that cannot be readily explained and accommodated within the existing conceptual framework [15]. Kurakin's papers give some diverse and important examples of these [14, 16 – 20].

This paper is not intended to be an exhaustive review of complex systems and their application to chemistry, rather a summary of the main concepts, important areas of application, and examples in an important area of overlap between chemistry and biology. The first section reviews the major elements of complex systems science from a chemistry perspective, summarizes the relatively small research effort in chemical complex systems, and indicates where a complex systems approach is likely to provide substantial benefit to new and existing chemical and biological challenges. By analogy with systems biology, we and others define the integrative study of very complicated chemical systems using the tools and concepts of complex systems science as *systems chemistry,* a term first employed by von Kiedrowski to define the chemical origins of biological organization [21]. Ludlow and Otto published the first review of systems chemistry as a subject in 2008 [22], although Whitesides noted the need to apply complexity concepts to chemistry in 1999 [12]. The second section of the paper focuses on the importance of scale in efficient modelling of complex systems, and discusses how Bayesian methods can be used to choose an appropriate level of scale for models. One of the most exciting new areas of chemical and biological research is the understanding and molecular control of stem cell fate for use in regenerative medicine. Complex systems tools and concepts are providing a novel, complementary approach to understanding and controlling stem cell fate. The third part of the paper illustrates how this complexity can be applied to modelling processes driving stem cell fate decisions.

Complex Systems

Complex systems science is a rapidly growing new paradigm for understanding and modelling extremely complicated systems (physical, social, economic, biological etc), and for discovering common mechanisms in apparently diverse phenomena. They generally exhibit

nonlinear dynamic behaviour caused by many interactions among the system components, and have so-called emergent properties that are difficult or impractical to understand or predict from knowledge of the components from which the system is constructed. Some complex systems can evolve in ways that are very sensitive to initial conditions. This behaviour can be described by nonlinear differential equations. Examples of complex systems are found in virtually all disciplines of science – from molecular self-organization and self-assembly to weather patterns, social systems, ecosystems, electricity grids, economies, and more esoteric properties like self-awareness, emotion, and life itself. Complex systems have been reviewed extensively in the recent literature [23 – 39]. Although there is still some debate about properties exhibited by complex systems, most exhibit a number of important properties that distinguish them from simple systems.

Criticality and nonlinearity

Complex systems often undergo an abrupt change in properties when a threshold value of some system parameter is crossed. This event is analogous to a phase transition that results from component interactions within the system switching from local connectivity to global connectivity. Many complex systems exhibit this type of behaviour, which is characterized by order and stability below the critical point and instability and chaos above the critical point. Phase transitions, crystallization, and alignment of magnetic domains in materials are common examples of systems exhibiting critical behaviour. Cramer and Booksh published one of the few reviews on chaos theory applied to chemistry, and list several interesting classes of chemical systems in which chaotic behaviour occurs [40].

Autocatalytic systems, which abound in chemistry and biology, are examples of nonlinear systems. The most familiar example of this type of behaviour is the complex spiral pattern observed in the BZ (Belousov-Zhabotinsky) reaction (Fig. 1). Although often observed as temporally varying series of patterns, reactions of this type also generate spatial patterns. Reaction-diffusion processes have been proposed as a mechanism for pattern formation in animals (e. g. tiger stripes), as a mechanism that generates body plans during embryogenesis.

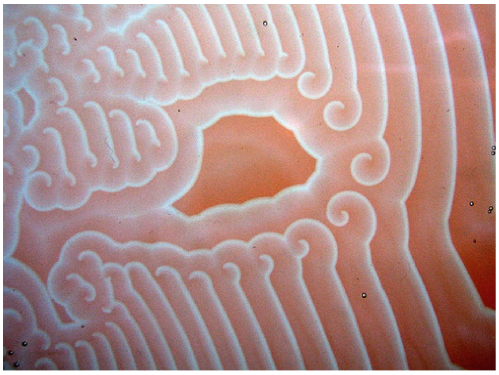

Figure 1. Spiral waves in BZ reaction

Self-organization and self-assembly

Interactions between large numbers of components may generate self-organized and self-assembled structures. Unfortunately, these terms are used in different ways in different scientific disciplines, and are even used interchangeably. A simple way of distinguishing them is to define self-assembly as a process that generates a structure that is in static equilibrium and is thermodynamically more stable than its components, and the assembly is driven by this energy difference [41]. Self-organizing systems increase their internal order over time implying a decrease in system entropy, so that self-organizing systems cannot be closed. The second law of thermodynamics requires energy transfers across the boundary and there is a corresponding increase in the entropy of the environment in which the system is embedded [33]. The self-organized structure is maintained in dynamic equilibrium and decays if the energy source is removed. A recent paper discusses the relationship between self-organization and self-assembly and gives examples of how these definitions apply across multiple disciplines [41]. Examples of systems that are difficult to categorize and that cross boundaries are also given. Understanding the mechanisms that underlie self-assembly and self-organization will enhance our ability to design materials and new technology for important fields such as nanotechnology, tissue engineering, and biomaterials. Complex systems concepts and modelling tools provide additional methods for understanding the behaviour of these systems.

Emergence and scale

Reductionist methods typically attempt to explain how a system behaves by considering the properties of the system components and their interactions. Complex systems approaches highlight emergent properties of a system and allow very complicated systems to be viewed at multiple levels of scale. Appropriate scales capture the emergent macroscopic properties of interest, without requiring a full understanding of mechanisms at finer levels of detail. In most cases, an ultimate fine-grained model of a complicated system is not possible because of the difficulty of modelling at this level of scale and the large amount of information required to build and validate such models. Generally, time, cost, or other resource limitations prevent this information becoming fully available. Complex systems often exhibit properties that are self-similar or scale-invariant (fractal) [42 – 45]. Power law behaviour in which a log-log plot of frequency versus magnitude of the property is linear, is observed in a diverse range of physical and biological complex systems [36, 46, 47]. The slopes of the power law graphs (the exponent of the power law) for different systems tend to cluster, generating *Universality Classes* that may have common underlying mechanisms [48 – 50]. Clearly, emergence, scale and complexity are interrelated and, as they are difficult concepts to define precisely, there is considerable latitude in how the concepts are employed across multiple disciplines. Recent work has attempted to generate more precise, widely applicable definitions for terms such as emergence, scale, self-assembly and self-organization [13, 41, 51, 52].

Emergent, holistic properties of complex systems are usually the ones that we observe and attempt to engineer. As they are generated by the relatively simple nonlinear interactions of many system components, these emergent properties are to some extent unexpected or surprising. Examples of emergent properties include the self-organized convection cells in heated liquids (Fig. 2), the oscillation of species in predator-prey systems, and weather patterns. Emergence and self-organization are intimately related, and recent work has described how emergence may be characterized in terms of self-organization [52].

Figure 2. Hexagonal, self-organized convection cells (Bénard cells) in a heated fluid.

Emergent properties can vary in their 'depth' [41, 52]. Simple emergent properties include gas pressure resulting from the interaction between atoms or molecules and their environment. More deeply emergent chemical properties include the formation of entirely new molecular species with novel properties when molecules participate in chemical reactions. Physical properties of molecules (e. g. aromaticity) can also be considered emergent, as it is very difficult or impossible to predict them from first principles given the properties of their components (atoms) and the interactions in which these components participate. However, emergent properties can often be predicted at a different level of scale (a coarser level of description). For example, the pressure of a gas can be calculated using a dynamical simulation that accounts for the speed and direction of many atoms or molecules, but is more readily calculated at a coarser level of scale using the gas law [53].

Networks and interactions

Networks are very important paradigms for understanding interactions in complex systems as they are a compact way of describing interactions between components. They have been extensively studied in many fields other than chemistry. If a complex chemical system consists of components that interact in a directed or non-directed way, this can be easily presented as a network or graph. Although graph theory has been long been applied to chemical systems such as molecules, relatively little work has been done on how network representations of molecules and chemical reactions can yield new insight into the properties of molecules. In systems biology, network paradigms are widely used to understand genetic regulation, signalling, and protein-protein interactions. There has been recognition recently that an understanding of network properties is necessary to properly exploit biological vulnerabilities and target drugs more effectively. Network concepts are therefore becoming increasingly important in medicinal chemistry and their applications to drug design have been reviewed recently [54 – 58].

Systems Chemistry

Compared to systems biology, the application of complex systems science to chemistry is almost non-existent. This is surprising given that Ilya Prigogine was awarded the Nobel Prize in Chemistry in 1977 for work on nonlinear dynamical systems in chemistry that helped to establish the foundations of complex systems science. Several other Nobel Prizes have been awarded for research relating to self-assembling systems (Crick, Watson and Wilkins, Medicine 1962; Krug, Chemistry 1982, Lehn, Chemistry, 1987) that build on complex systems frameworks. Clearly, Watson, Crick and Wilkins' work on understanding the structure and self-assembly of DNA led to a paradigm shift in our understanding and revolutionized the fields of biology and medicine. However, their work had a relatively smaller impact on chemistry, a discipline that is well equipped to describe the interactions that control DNA's structure. A few other chemists, notably Testa, Kier and Belousov and Zaboutinsky, explored complex behaviour in chemistry but the area did not develop in a sustainable way and declined in prominence. In a sense, the field was ahead of its time, and did not offer solutions to problems that reductionism could not tackle adequately. Within the past decade, interest in applying complexity to chemistry has increased, and several workers, notably Whitesides, have written seminal papers. Whitesides proposed that almost everything of interest in chemistry is complex, by the above definitions of complexity [12]. For example a single step in a multistep synthesis involves ~10^{22} molecules of several types, a myriad of interacting nuclei and electrons and ~10^{24} molecules of solvent. As such real systems cannot be dealt with at the molecule level, synthetic chemistry has employed approximations, rules, and heuristics. Complex systems paradigms can provide additional, complementary methods for understanding and dealing with such immense complexity.

Contemporary Chemical Research Where a Complex Systems Approach May Be Relevant

Most of the interesting new areas of chemistry are discipline spanning, and will benefit from a complementary complex systems approach.

Nonlinear dynamical systems and pattern formation

Nonlinear dynamical systems can exhibit chaotic behaviour, often modelled using differential equations akin to those describing chemical kinetics. Spatial and cellular automata methods are being applied increasingly to dynamical systems. As discussed, some chemical reactions can behave in this manner. Although the most recognised example is the BZ reaction [59] (see Figure 2), many other dynamic, pattern-forming reactions such as the Briggs-Raucher reaction are known. Recently, reaction-diffusion or oscillating reactions have been used to generate novel materials such as a self-walking gel [60, 61].

Fractal materials

Systems that exhibit fractal or self-similar behaviour are found widely in nature – snowflakes, lightning, crystals, river networks, ferns, blood vessels and coastlines. Real world objects show fractal behaviour over an extended, but finite, scale range [108]. Fractal behaviour is common in nano-structured materials, where self-similarity provides unique properties such as giant enhancement of nonlinear optical properties, and related effects that underlie surface-enhanced Raman scattering [62]. Fractal effects are also important in electrochemistry [63].

Supramolecular polymers

Conventional polymers are linked by permanent covalent bonds. Supramolecular polymers are a new class of polymers in which the linkages are reversible. The dynamic properties of these materials open up the prospect of many new applications. The most important step in developing practical supramolecular polymers was the discovery of the 2-ureido-4[1 H]-pyrimidinone unit that could form four hydrogen bonds. This monomer has self-complementary allowing it to self-assemble into polymers with novel properties, such as the ability to self-heal when cut [64].

Self-replicating systems

Self-replicating systems clearly exist in biology and there has been much speculation on the types of autocatalytic systems that may have arisen on the primordial earth that may have led to self-replicating molecules and life. Kauffman argued that a sufficiently large pool of chemicals may be intrinsically capable of generating autocatalytic cycles and self-replicating

molecules [65]. Seminal work by Wintner *et al.* showed how simple organic structures can catalyse their own formation. Self-complementarity is the key to this autocatalytic behaviour [66]. Design of self-replicating molecules is very difficult and progress in the field has been slow. The recent review by Paul and Joyce summarizes progress in the area [67].

Artificial life

Artificial life and the closely related field of synthetic biology covers a range of research areas, many of them intimately involved with chemistry. It spans synthesis of unnatural organic molecules that function in living systems, through biomimetic chemistry that produces synthetic molecules that recapitulate the behaviour of natural analogues, to the greatest challenge of recreating synthetic chemical systems that exhibit inheritance, genetics and evolution. Recent reviews and papers by Benner, Rosen, and Cornish-Bowden provide a summary of progress in artificial life [68 – 70].

Dynamic combinatorial libraries

Dynamic combinatorial chemistry is defined as combinatorial chemistry under thermodynamic control, in which all constituents are in equilibrium. The library members are therefore constantly interconverting via reversible chemical reactions. The composition of the library can therefore be altered by changing factors in the environment [71]. For example, dynamic libraries have been exploited in drug discovery. When members of the library interact favourably with a biological target, the composition of the library shifts and results in enrichment of compounds that bind to the target protein. The environment can therefore 'select' or enrich the library with compounds that are more 'fit'. Given the enormous size of drug-like chemistry space, dynamic combinatorial libraries offer a more efficient method of exploring this space to identify novel drug leads than covalently bonded combinatorial libraries. The technique is currently limited by the small number of reversible reactions that are biocompatible.

Evolutionary methods

Genetic algorithms and genetic programming are also evolutionary, adaptive tools that are finding increased applications in chemistry and biology. Genetic algorithms are very efficient methods for exploring vast search spaces such as those involved in protein folding, or drug-like chemical space. When applied to chemistry, they involve a string representation of a molecule or molecular property (e. g. a SMILES string, torsion angle list etc), a mutation operator (e. g. swap one atom type or functional group for another, add or delete an atom or functional group, change a torsion angle etc), and a fitness function (e. g. biological activity, desired protein fold etc). The algorithm generates families of mutants that are assessed by the fitness function. The fittest are retained and less fit discarded. The cycle repeats starting with the fittest individuals from the previous cycle, and stops when some criterion is

matched. They have been used to find optimum sets of molecular descriptors for modelling of drug activity, for searching chemical databases for leads [73 – 75], and for optimising combinatorial libraries [76], discovery and optimisation of new catalytic materials [77 – 79], design of peptide mimetopes [78, 79], multiobjective optimisation [80, 81], synthesis planning [82 – 85], and in many other chemical applications.

Drug targeting

Interactions between molecules (e. g., drugs and proteins, proteins and DNA, etc.) in living systems can be described by networks. Very recently, network approaches have been applied to understand why the efficacy of new drug discovery is declining [54, 55, 86]. It has long been recognized that almost all drugs hit multiple targets, which is problematic as it leads to off-target effects and toxicity. Drugs are often designed to hit a specific target that is presumed to induce the desired phenotypic change, but it has only recently been appreciated that the robustness of biological networks resists such perturbations and may allow systems to adapt to minimize the impact. If we understand the robustness of biological networks, we could better engineer drugs to hit multiple intentional targets to affect the best outcome.

Supramolecular, self-organizing and self-assembling systems

Self-assembling, supramolecular, and self-organizing materials are currently of considerable interest to chemists [87 – 90]. New research fields of nanoscience and smart materials aim to design materials and molecular devices that can be assembled from the 'bottom up' rather than 'top down'. Chemical or biological materials have been shown to undergo self-assembly, although a complete understanding of assembly mechanisms is still lacking, making design challenging. Self-assembly has been applied to electronic device manufacture [91], viral capsid modelling [92], nanoparticle assembly [93 – 95], peptides self-assembly, nanotubes, glycolipids, carbohydrates, polymers, crystals [94, 96 – 101], and many other materials[30]. Aberrant self-assembly provides the mechanism for many protein folding diseases such as prion and Alzheimer's diseases, amyloidoses, cataracts, and diabetes [102, 103]. New complexity-based tools, such as agent-based models, have an increasing important role to play in design of self-assembling materials [13, 104].

Self-organization also applies to chemistry but is less prominent in the chemical literature largely due to the tendency of many authors to describe self-assembly as self-organization[39]. Most examples of true self-organization involve biological systems because they are invariably open [33, 105, 106]. A chemical example (although biology-based) is the dynamic instability of microtubules [107]. Formation of microtubules involves a spontaneous self-assembly of tubulin, but the dynamics of the process also involves disassembly using energy librated when tubulin hydrolyses GTP (tubulin is an enzyme as well as structural protein) [108]. An example of a self-organized tubulin structure is given in Figure 3.

Cram and Lehn conducted seminal work on self-assembly, establishing the field of supramolecular chemistry involving the generation of complex structures from reversible non-covalent bonds These structures exploit hydrogen bonding, metal coordination, hydrophobic forces, van der Waals forces, π-π interactions and electrostatic effects [109].

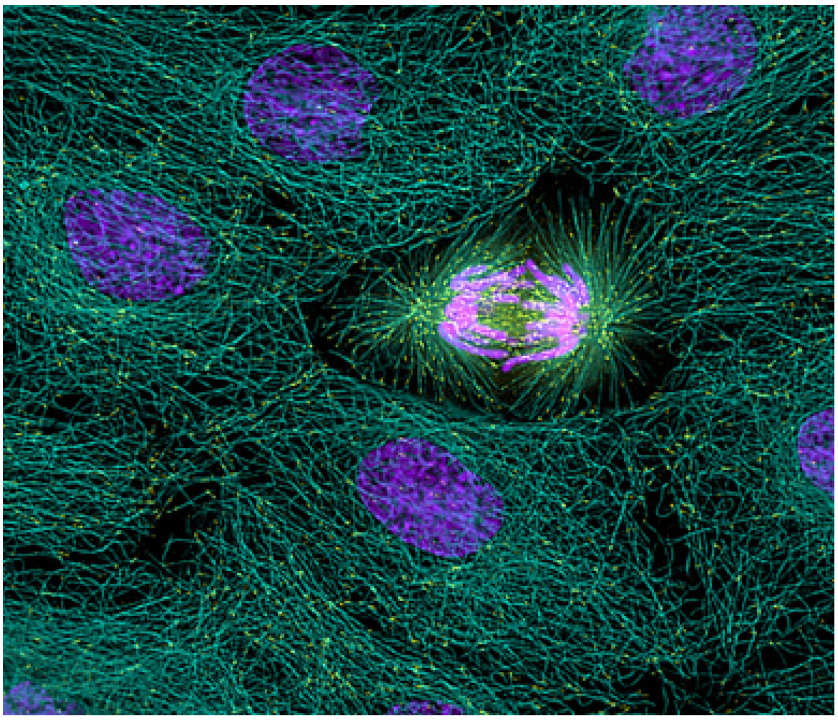

Figure 3. A self-organized self-assembly of microtubules.

Mescoscale Modelling – Choosing the Correct Scale

Some complex systems appear to be 'computationally irreducible', raising the question of how we model emergent properties of such systems. Recent work on diverse complex systems suggests sparse, mesoscale, or 'coarse grained' models, that omit most of the detail and concentrate only on the essential pattern forming process, can provide a path forward. Choosing too low a level of detail (too coarse) results in models that have a limited ability to generalize, and can predict only broad features of the system. Choosing a too detailed level of scale produces more complicated models but requires much more data. As many modelling exercises are data limited, overly complex models pose problems for parameterisation, validation, and they are prone to overfitting.

Biologically, coarse graining of sensory information is essential so that organisms can function without being overloaded. For example, in the visual system the eye does substantial feature selection and classification before information is sent to the brain. If mesoscale descriptions can adequately model the essential (emergent) properties of complex systems, how do we find the correct level of scale that matches the usefulness of the predictions to the availability of data? We adopt two approaches, one specific for biology and one more generally applicable.

Sparse feature selection and modelling

Generically, mesoscale models can be generated using sparse feature selection methods that select in a context-dependent manner, the most relevant aspects of the system under study. Prominent of these are Bayesian methods like expectation maximization algorithms, described elegantly by Figeuiredo [110].

Neural networks are another adaptive, learning method that is finding increased application in chemistry and biology [58, 111 – 116]. Neural networks are algorithms that learn patterns or associations between mathematical representations of objects and emergent properties of systems that are comprised of, or interact with these objects. An important area where neural networks have contributed to complexity-based modelling of chemical systems, is in QSAR and chemometrics. Neural networks have been used widely in chemistry and pharmacology for finding relationships between molecules and their biological or physicochemical properties. They have an inherent ability to model a complex system at a sufficiently coarse grained scale that matches the model complexity to the availability of data. This is particularly important in medicinal chemistry, which is often data-poor [12]. Neural networks are platform technologies that have been used to model a myriad of emergent properties of chemical and biological complex systems, such as drug target activity [117 – 121], drug-like (ADME) properties [122 – 126], immune modulation [127, 128], physicochemical properties [129 – 136], toxicity [112, 137 – 142].

Mesoscale models of regulatory interactions

Although there is vigorous debate [143] about the architecture of biological networks, many are scale-free due to the intrinsic robustness of such architectures. Mesoscale models of cell regulatory networks can be derived by focusing on key network hubs and genes controlled by these hubs that are relevant to the biological process under study. Mesoscale regulator models can represent many types of interacting components in biology. In gene regulation, the hubs represent the set of genes most relevant to a particular biological process such as differentiation, apoptosis or self-renewal. Mesoscale regulatory networks can be deduced by a ‘bottom up’ reductionist analysis of gene microarray data, or data from other experiments such as chromatin immunoprecipitation (ChIP). An alternative ‘top down’ approach is to generate ‘blank’ mesoscale networks that are subsequently trained to recapitulate

experimental data or predict the outcomes of new experiments. The latter approach was first described by Geard and Wiles, and is a substantially novel analogue applied to modelling of quantitative gene regulation in *C. elegans* embryogenesis [144, 145]. The approach is illustrated in Figure 4.

This coarse-grained approach to cell differentiation allows us to rise above the details of biochemical experiments to produce a map or 'fate matrix' of cell behaviour that reflects the integration of multiple layers of complexity in cellular processes. Given a set of target genes that characterize each cell embryogenesis pathway, a mesoscale regulatory network model can be trained using gene microarray data, so that its output mimics changes in gene activity as cells divide and/or differentiation. Once trained, the network architecture and weight matrix represent a coarse-grained model of the regulatory network controlling cell fate. Each gene in the model interacts with the weight matrix, reflecting the fact that the transcription of a gene is the result of integrating the cell's biochemical state rather than the action of any single gene product. One set of master genes (i.e. one expression state) regulates the transcription of the cell's genes, leading to a new state. We aim to annotate these mesoscale models with experiments identifying regulatory interactions, to produce a model of sufficient detail to understand and control cell fate [146].

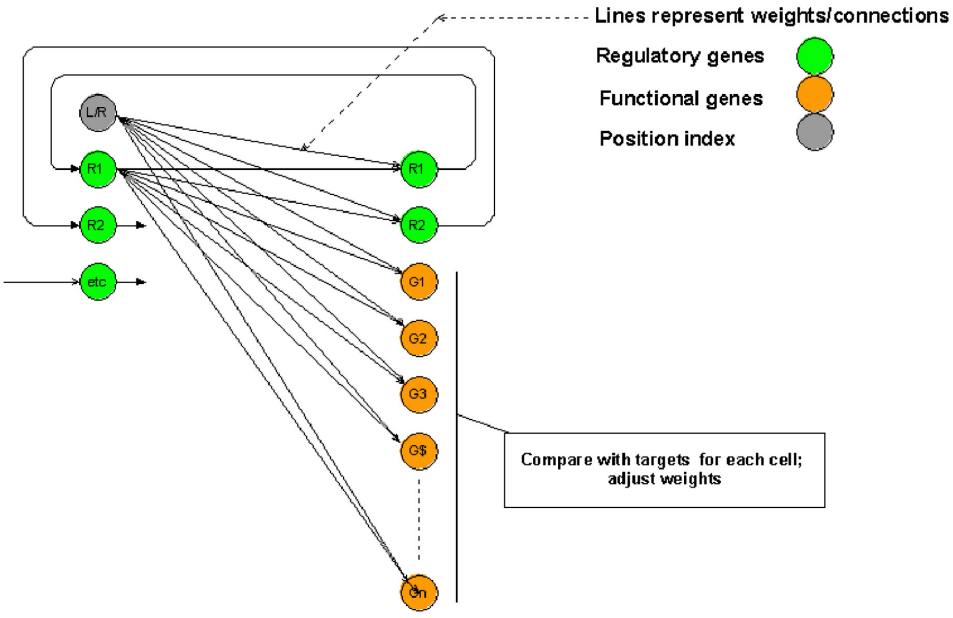

Figure 4. Recursive neural network structure, showing how the regulatory genes and relative position index control the expression levels of genes in the various cells in the lineage.

Modelling Stem Cell Fate Decisions

Complex systems science coupled with interdisciplinary approaches provide scope to integrate the many different layers of complexity that characterise complex biological systems. Hence, important areas of research such as stem cell biology a growing literature views stem cell function within the framework of a complex adaptive network, rather than the traditional deterministic and reductionist approach that focuses on single stem cells [147]. We summarize several areas where sparse methods have been applied to build mesoscale models of properties of stem cell or related multipotent cells.

Stemness genes in human embryonic stem cells (hESCs)

Laslett *et al.* isolated sub-populations of embryonic stem cells that differed slightly in their degree of 'stemness' (pluripotency) [148]. The experiments aimed to identify genes that maintained pluripotency in hESC and those that drive early commitment to other lineages. Flow cytometry was used to select a small set of pluripotent hESCs using fluorescent antibody markers whose expression correlated with the pluripotency marker Oct4. This very stemlike subpopulation was further sorted into hESC populations at four, graded, very early stages of differentiation (P4-P7) (Fig. 5). Gene expression microarrays were generated from each of the populations. Sparse Bayesian feature selection algorithms identified a small number of markers that correlate with the pluripotency of the hESCs, and were classifier genes with a putative key role in the maintenance of pluripotence. They could achieve this classification with high statistical significance, suggesting the genes played a prominent role in stemness.

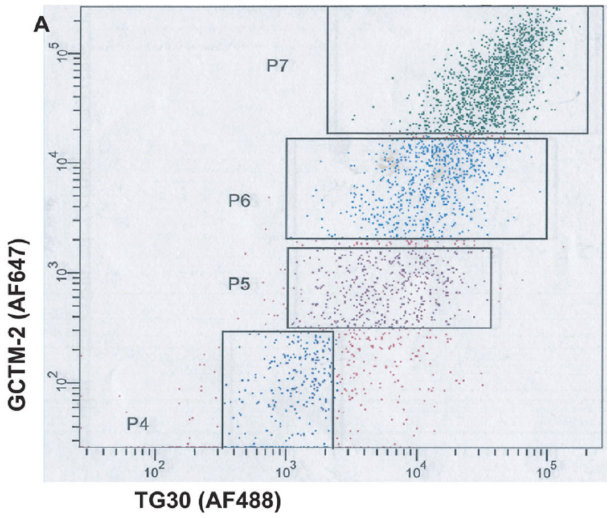

Figure 5. Flow cytometric experiments identifying the four classes of hESCs.

Erythroid differentiation

Welch and co-workers recently reported experiments on estradiol-induced differentiation along the erythroid pathway over 30 hours [149]. This trajectory starts from the late burst-forming unit-erythroid stage and progresses to the orthochromatic erythroblast stage. Estradiol was used to trigger the expression of the gene GATA-1, initiating synchronous differentiation in the population. Microarrays were generated for populations of cells at each time point during the experiment. Five classifier genes were found that correlate with the stage of differentiation down the lineage. These genes were highly reliable markers for the differentiation process, and are likely to play an important role in regulation of the erythroid pathway.

Mesoscale network modelling of mouse embryogenesis

We modelled murine embryogenesis from the zygote to the blastocyst stage (Fig. 6) using a 'top down' mesoscale regulatory modelling framework.

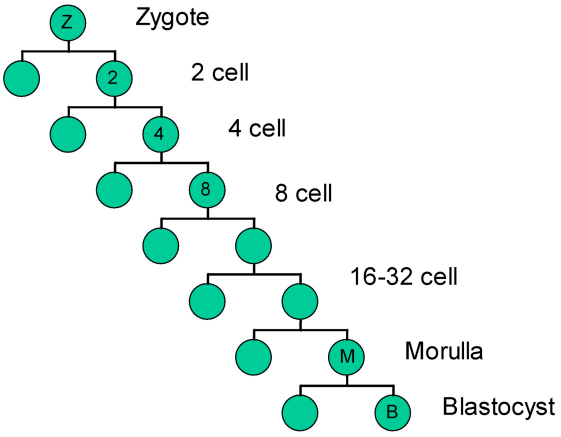

Figure 6. The lineage diagram for mouse embryogenesis. Cell labels are shown.

Hamatami *et al.* measured the expression level of genes in carefully prepared samples of cells from the unfertilized egg to the blastocyst using gene microarrays [150] (see Figure 6). We modelled 135 genes from the mid-preimplantation gene activation (MGA) cluster for each cell type in the embryogenesis pathway. We first trained the model using the gene expression targets for each of the 135 genes in the cell types at the seven time points. We then assessed the ability of the model to predict gene expression in cells in the pathway. The mesoscale regulatory network model recapitulated quantitative gene expression profiles within the cell lineage surprisingly high fidelity. We also generated models that predicted expression in cells at an intermediate point in the lineage (interpolation) and at the end of the lineage (extrapolation) when these were not used in training.

Figure 7 summarizes these predictions. Figure 7A shows how well the model is able to recapitulate the quantitative expression levels of all genes in all cells in the training set. Figure 7B shows the most stringent test where the expression levels of the last three cell divisions in the trajectory are predicted from a model trained on the first four cells in the lineage. While the error is substantially higher than in the single cell expression prediction, the predictions nevertheless have high statistical significance.

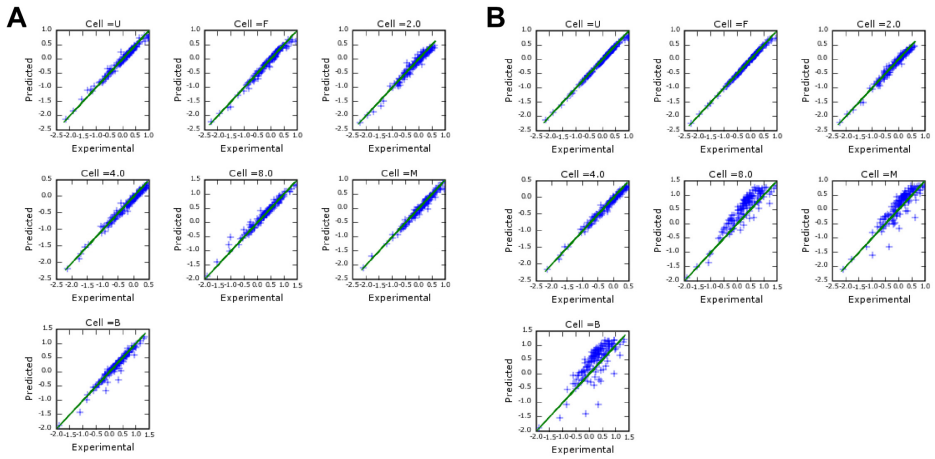

Figure 7. Predicted versus experimental log2 fold expression levels for genes in the seven cell types. (**A**) Training set. (**B**) prediction of 8-cell, morula and blastocyst expression for model trained on first 4 cells in embryogenic pathway.

Why sparse models work

Our sparse feature selection methods and mesoscale regulatory models support the common observation that a relatively small number of factors control the fate of stem cells – as few as three to six. Hart *et al.* also illustrated a similar sparse transcription factor requirement when modelling the yeast cell cycle transcription network [151]. They initially used 204 transcription factors to model the temporal variation of genes across the cell cycle. They pruned their network models progressively, removing the least relevant transcription factors in turn, and monitored the performance of the model. They found a gradual and limited degradation in model performance until five transcription factors remained, after which the performance of the model collapsed (Fig. 8).

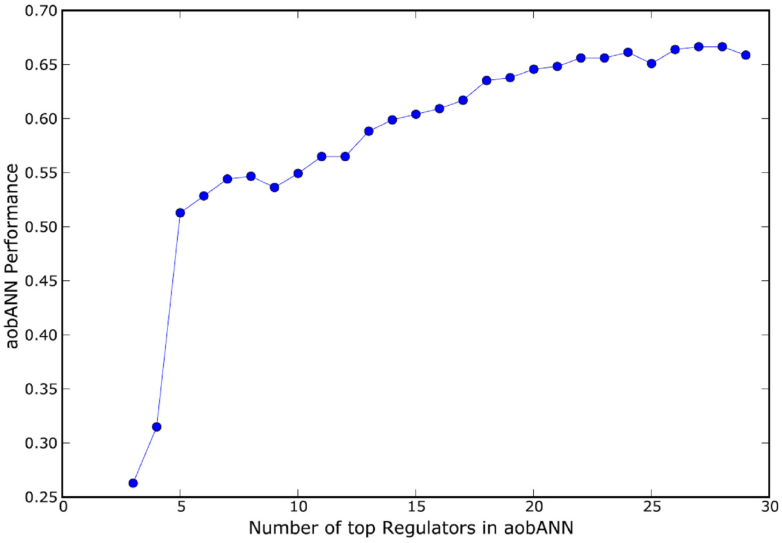

Figure 8. Performance of neural network model of temporal variation in gene expression during cell cycle. Note the dramatic drop in performance when the number of transcription factors is 5 or less.

Conclusions and the Future

The evolution of chemistry away from manipulation of single molecules towards description and manipulation of complex systems of molecules is one of the key driving forces for embracing systems chemistry. In this way, the move to the study of more complex matter may counter intuitively generate new types of problems that are sufficiently simple when viewed in a complexity framework that they can be tackled in an analytical sense [12]. We have arrived at Kuhn's crisis of the dominant paradigm in chemistry, a period that is always followed by an exciting period of extraordinary research and new discoveries [14]. However, we can only achieve this by recognizing the opportunity and acting on it. Ludlow and Otto's call for chemists to embrace complexity and recognize systems chemistry is a very timely challenge to expand this important new paradigm for understanding chemical problems [22].

Chemists may finally capitalize on the powerful but in some respects premature possibilities that the Nobel Laureates presented to us, and move more rapidly into the exciting worlds of new chemistries, materials, and biology.

References

[1] Kitano, H. (2002) Computational systems biology. *Nature* **420**:206 – 10.

[2] Kitano, H. (2002) Systems biology: a brief overview. *Science* **295**:1662 – 4.

[3] Way, J.C., Silver, P.A. (2007) Systems engineering without an engineer: Why we need systems biology. *Complexity* **13**:22 – 29.

[4] Fisher, J., Henzinger, T.A. (2007) Executable cell biology. *Nat. Biotechnol.* **25**:1239 – 49.

[5] Aigner, T., Haag, J., Zimmer, R. (2007) Functional genomics, evo-devo and systems biology: a chance to overcome complexity? *Curr. Opin. Rheumatol.* **19**:463 – 70.

[6] Marguet, P., Balagadde, F., Tan, C.M., You, L.C. (2007) Biology by design: reduction and synthesis of cellular components and behaviour. *J. Roy. Soc. Interface.* **4**:607 – 23.

[7] Quackenbush, J. (2007) Extracting biology from high-dimensional biological data. *J. Exp. Biol.* **210**:1507 – 17.

[8] Materi, W., Wishart, D.S. (2007) Computational systems biology in drug discovery and development: methods and applications. *Drug Disc. Today.* **12**:295 – 303.

[9] van Regenmortel, M.H.V. (2007) The rational design of biological complexity: A deceptive metaphor. *Proteomics* **7**:965 – 75.

[10] Lin, J., Qian, J. (2007) Systems biology approach to integrative comparative genomics. *Expert. Rev. Proteomics* **4**:107 – 19.

[11] Palumbo, M.C., Farina, L., Colosimo, A., Tun, K., Dhar, P.K., Giuliani, A. (2006) Networks everywhere? Some general implications of an emergent metaphor. *Curr. Bioinfo.* **1**:219 – 34.

[12] Whitesides, G.M., Ismagilov, R.F. (1999) Complexity in chemistry. *Science* **284**:89 – 92.

[13] Lehn, J.-M. (2002) Toward self-organization and complex matter. *Science* **295**:2400 – 3.

[14] Kurakin, A. (2007) Self-organization versus Watchmaker: ambiguity of molecular recognition and design charts of cellular circuitry. *J Mol Recog.* **20**:205 – 14.

[15] Kuhn, T. (1996) The Structure of Scientific Revolutions. Chicago: University of Chicago Press.

[16] Kurakin, A. (2004) Self-organization versus watchmaker: stochasticity and determinism in molecular and cell biology. Novato lectures.

[17] Kurakin, A. (2005) Self-organization vs. Watchmaker: stochastic gene expression and cell differentiation. *Dev. Genes Evol.* **215**:46 – 52.

[18] Kurakin, A. (2005) Stochastic cell. *IUBMB Life* **57**:59 – 63.

[19] Kurakin, A. (2005) Self-organization versus watchmaker: stochastic dynamics of cellular organization. *Biol. Chem.* **386**:247 – 54.

[20] Kurakin, A. (2006) Self-organization versus Watchmaker: Molecular motors and protein translocation. *Biosys.* **84**(1):15 – 23.

[21] Stankiewicz, J., Eckardt, L.H. (2006) Chembiogenesis 2005 and Systems Chemistry Workshop. *Angew. Chem. Int. Ed.* **45**:342 – 4.

[22] Ludlow, F.R., Otto, S. (2008) Systems Chemistry. *Chem. Soc. Rev.* **37**:101 – 8.

[23] Batty, M. (2008) The size, scale, and shape of cities. *Science* **319**:769 – 71.

[24] Luzzi, R, Vasconcellos, A.R., Ramos, J.G. (2007) Non-equilibrium statistical mechanics of complex systems: An overview. *Rivista Del Nuovo Cimento* **30**:95 – 157.

[25] Mara, A., Holley, S.A. (2007) Oscillators and the emergence of tissue organization during zebrafish somitogenesis. *Trends Cell Biol.* **17**:593 – 9.

[26] Weber, B.H. (2007) Emergence of life. *Zygon*. **42**:837 – 56.

[27] Foote, R. (2007) Mathematics and complex systems. *Science* **318**:410 – 2.

[28] Eckmann, J.P., Feinerman, O., Gruendlinger, L., Moses, E., Soriano, J., Tiusty, T. (2007) The physics of living neural networks. *Physics Rep-Rev Section Phys Lett.* **449**:54 – 76.

[29] Rammel, C., Stagl, S., Wilfing, H. (2007) Managing complex adaptive systems – A co-evolutionary perspective on natural resource management. *Ecolog. Econom.* **63**(1):9 – 21.

[30] Balazs, A.C. (2007) Modelling self-assembly and phase behaviour in complex mixtures. *Ann. Rev. Phys. Chem.* **58**:211 – 33.

[31] Almaas, E. (2007) Biological impacts and context of network theory. *J. Exp. Biol.* **210**:1548 – 58.

[32] Mitchell, M. (2006) Complex systems: Network thinking. *Artific. Intell.* **170**:1194 – 212.

[33] Newth, D., Finnigan, J. (2006) Emergence and self-organization in chemistry and biology. *Aust. J. Chem.* **59**:841 – 8.

[34] Markose, S.M. (2005) Computability and evolutionary complexity: Markets as complex adaptive systems (CAS). *Econom. J.* **115**:F159-F92.

[35] Amaral, L.A.N., Ottino, J.M. (2004) Complex networks – Augmenting the framework for the study of complex systems. *Eur. Phys. J. B.* **38**:147 – 62.

[36] Gisiger, T. (2001) Scale invariance in biology: coincidence or footprint of a universal mechanism? *Biol. Rev.* **76**:161 – 209.

[37] Hess, B. (2000) Periodic patterns in biology. *Naturwiss.* **87**:199 – 211.

[38] Wales, D.J., Scheraga, H.A. (1999) Review: Chemistry – Global optimization of clusters, crystals, and biomolecules. *Science* **285**:1368 – 72.

[39] Halley, J.D., Winkler, D.A. (2006) Classification of self-organization and emergence in chemical and biological systems. *Aust. J. Chem.* **59**:849 – 53.

[40] Cramer, J.A., Booksh, K.S. (2006) Chaos theory in chemistry and chemometrics: a review. *J. Chemom.* **20**:447 – 54.

[41] Halley, J.D., Winkler, D.A. (2008) Consistent concepts of self-organization and self-assembly. *Complexity* **14**(2):10 – 17.

[42] Kenkel, N.C., Walker, D.J. (1996) Fractals in the Biological Sciences. *COENOSES* **11**:77 – 100.

[43] Mandelbrot, B.B. (1977) Fractals: Form, Chance and Dimension. San Francisco: W. H. Freeman.

[44] Mandelbrot, B.B. (1982) The fractal geometry of nature. San Francisco: W. H. Freeman.

[45] Sornette, D. (2000) Critical Phenomena in Natural Sciences. Chaos, Fractals, Self-Organization and Disorder: Concepts and Tools. Berlin: Springer.

[46] Avnir, D., Bihan, O., Malcai, O. (1998) Is the geometry of nature fractal? *Nature* **279**:39 – 40.

[47] Schroeder, M. (1991) Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise. New York: W. H. Freeman and Company.

[48] Hughes, D., Paczuski, M. (2002) Large Scale Structures, Symmetry, and Universality. *Phys Rev Lett.* **88**:054302.

[49] Ward, M. (2001) Universality: the underlying theory behind life, the universe and everything. London: Pan Books.

[50] Stanley, H.E., Amaral, L.A.N., Gopikrishnan, P., Ivanov, P.C., Keitt, T.H., Plerou, V. (2000) Scale invariance and universality: organizing principles in complex systems. *Physica A*. **281**:60 – 8.

[51] Ryan, A.J. (2007) Emergence is coupled to scope, not level. *Complexity* **13**:67 – 77.

[52] Halley, J.D., Winkler, D.A. (2008) A simple description of emergence and its relation to self-organization. *Complexity* **13**:10 – 15.

[53] Coleman, P. (2007) Frontier at your fingertips. *Nature* **446**:379.

[54] Hopkins, A.L. (2007) Network pharmacology. *Nat. Biotechnol.* **25**:1110 – 1.

[55] Hopkins, A.L., Mason, J.S., Overington, J.P. (2006) Can we rationally design promiscuous drugs? *Curr. Opin. Struct. Biol.* **16**(1):127 – 136.

[56] Kitano, H. (2007) A robustness-based approach to systems-oriented drug design. *Nat. Rev. Drug Disc.* **6**:202 – 10.

[57] Sharom, J.R., Bellows, D.S., Tyers, M. (2004) From large networks to small molecules. *Curr. Opin. Chem. Biol.* **8**:81 – 90.

[58] Winkler, D.A. (2008) Network models in drug discovery and regenerative medicine. *Biotech. Ann. Rev.* **14**:143 – 170.

[59] Zaikin, A.N., Zhabotin, A.M. (1970) Concentration Wave Propagation in 2-Dimensional Liquid-Phase Self-Oscillating System. *Nature* **225**:535.

[60] Maeda, S., Hara, Y., Sakai, T., Yoshida, R., Hashimoto, S. (2007) Self-walking gel. *Adv. Mat.* **19**:3480.

[61] Yashin, V.V., Balazs, A.C. (2006) Pattern Formation and Shape Changes in Self-Oscillating Polymer Gels. *Science* **314**:798 – 801.

[62] Shalaev, V.M., Markel, V.A., Poliakov, E.Y., Armstrong, R.L., Safonov, V.P., Sarychev, A.K. (1998) Nonlinear Optical Phenomena in Nanostructured Fractal Materials. *J. Nonlin. Opt. Phys. Mat.* **7**:131 – 52.

[63] Hibbert, D.B., Melrose, J.R. (1988) Copper electrodeposits in paper support. *Phys. Rev. A.* **38**:1036 – 48.

[64] de Greef, T.F.A., Meijer, E.W. (2008) Materials science: Supramolecular polymers. *Nature* **453**:171 – 3.

[65] Kauffman, S.A. (1995) At Home in the Universe. The Search for Laws of Self-Organization and Complexity. New York: Oxford University Press.

[66] Wintner, E.A., Conn, M.M., Rebek, J. (1994) Self-Replicating Molecules: A Second Generation. *J. Am. Chem. Soc.* **116**:8877 – 84.

[67] Paul, N., Joyce, G.F. (2004) Minimal self-replicating systems. *Curr. Opin. Chem.-Biol.* **8**:634 – 9.

[68] Benner, S.A., Sismour, A.M. (2005) Synthetic biology. *Nat. Rev. Genet.* **6**:533 – 43.

[69] Cornish-Bowden, A. (2006) Putting the Systems back into Systems Biology. *Persp. Biol. Med.* **49**:475 – 89.

[70] Rosen, R. (1991) Life itself:A comprehensive inquiry into the nature, origin, and fabrication of life. New York: Columbia Univ.Press.

[71] Corbett, P.T., Leclaire, J., Vial, L., West, K.R., Wietor, J.-L., Sanders, J.K.M., *et al.* (2006) Dynamic Combinatorial Chemistry. *Chem. Rev.* **106**:3652 – 711.

[72] Tan, D.S. (2005) Diversity-oriented synthesis: exploring the intersections between chemistry and biology. *Nat. Chem. Biol.* **1**:74 – 84.

[73] Neri, F., Toivanen, J., Makinen, R.A.E. (2007) An adaptive evolutionary algorithm with intelligent mutation local searchers for designing multidrug therapies for HIV. *Applied Intelligence* **27**(3):219 – 35.

[74] Grosdidier, A., Zoete, V., Michielin, O. (2007) EADock: Docking of small molecules into protein active sites with a multiobjective evolutionary optimization. *Proteins* **67**(4):1010 – 25.

[75] Jain, A.N. (2006) Scoring functions for protein-ligand docking. *Curr. Prot. Pept. Sci.* **7**:407 – 20.

[76] Jung, Y.S., Kulshreshtha, C., Kim, J.S., Shin, N., Sohn, K.S. (2007) Genetic algorithm-assisted combinatorial search for new blue phosphors in a (Ca,Sr,Ba, Mg,Eu)(x)ByPzO delta system. *Chem. Mat.* **19**:5309 – 18.

[77] Serra JM, Corma A, Valero S, Argente E, Botti V.(2007) Soft computing techniques applied to combinatorial catalysis: A new approach for the discovery and optimization of catalytic materials. *QSAR Combin. Sci.* **26**:11 – 26.

[78] Belda, I., Madurga, S., Tarrago, T., Llora, X., Giralt, E. (2007) Evolutionary computation and multimodal search: A good combination to tackle molecular diversity in the field of peptide design. *Mol. Divers.* **11**:7 – 21.

[79] Hohm, T., Limbourg, P., Hoffmann, D. (2006) A multiobjective evolutionary method for the design of peptidic mimotopes. *J. Comp. Biol.* **13**:113 – 25.

[80] Nicolaou, C.A., Brown, N., Pattichis, C.S. (2007) Molecular optimization using computational multi-objective methods. *Curr. Opin. Drug Disc. Dev.* **10**:316 – 24.

[81] Rusu, T., Bulacovsch, V. (2006) Multiobjective tabu search method used in chemistry. *Int. J. Quant. Chem.* **106**:1406 – 12.

[82] Lameijer, E.W., Kok, J.N., Back, T., Ijzerman, A.P. (2006) The molecule evoluator. An interactive evolutionary algorithm for the design of drug-like molecules. *J. Chem. Info. Mod.* **46**:545 – 52.

[83] Liao, C.Z., Liu, B., Shi, L.M., Zhou, J.J., Lu, X.P. (2005) Construction of a virtual combinatorial library using SMILES strings to discover potential structure-diverse PPAR modulators. *Eur. J. Med. Chem.* **40**:632 – 40.

[84] Gillet, V.J., Khatib, W., Willett, P., Fleming, P.J., Green, D.V.S. (2002) Combinatorial library design using a multiobjective genetic algorithm. *J. Chem. Inf. Comp. Sci.* **42**:375 – 85.

[85] Weber, L. (2002) Multi-component reactions and evolutionary chemistry. *Drug Dis. Today* **7**:143 – 7.

[86] Yildirim, M.A., Goh, K., Cusick, M.E., Barabasi, L., Vidal, M. (2007) Drug-target network. *Nat. Biotechnol.* **25**:1119 – 26.

[87] Whitesides, G.M., Boncheva, M. (2002) Beyond molecules: self-assembly of mesoscopic and macroscopic components. *Proc. Nat. Acad. Sci. U.S.A* **99**:4769 – 74.

[88] Lehn, J.-M. (1995) Supramolecular Chemistry. New York Weinheim.

[89] Lindoy, L.F., Atkinson, I.M. (2000) Self-assembly in Supramolecular Systems. Cambridge, UK: Royal Society of Chemistry.

[90] Whitesides, G.M., Ferguson, G.S., Allara, D., Scherson, D., Speaker, L., Ulman, A. (1993) Organized molecular assemblies. *Crit. Rev. Surf. Chem.* **3**:49 – 65.

[91] Black, C.T., Ruiz, R., Breyta, G., Cheng, J.Y., Colburn, M.E., Guarini, K.W., *et al.* (2007) Polymer self assembly in semiconductor microelectronics. *IBM J. Res. Dev.* **51**:605 – 33.

[92] Sitharam, M., Agbandje-Mckenna, M. (2006) Modelling virus self-assembly pathways: Avoiding dynamics using geometric constraint decomposition. *J. Comp. Biol.* **13**:1232 – 65.

[93] Fukushima, T., Jin, W., Aida, T. (2007) Graphitic nanotubes formed by programmed self-assembly. *J. Synth. Org. Chem. Jap.* **65**:852 – 61.

[94] Arumugam, P., Xu, H., Srivastava, S., Rotello, V.M. (2007) 'Bricks and mortar' nanoparticle self-assembly using polymers. *Polymer Int.* **56**:461 – 6.

[95] Shimizu, T. (2006) Self-assembly in discrete organic nanotubes. *Sen-I Gakkaishi* **62**:P114-P8.

[96] Corti, M., Cantu, L., Brocca, P., Del Favero, E. (2007) Self-assembly in glycolipids. *Curr. Opin. Coll. Inter. Sci.* **12**:148 – 54.

[97] Colombo, G., Soto, P., Gazit, E. (2007) Peptide self-assembly at the nanoscale: a challenging target for computational and experimental biotechnology. *Trends Biotech.* **25**:211 – 8.

[98] Kokkoli, E., Mardilovich, A., Wedekind, A., Rexeisen, E.L., Garg, A., Craig, J.A. (2006) Self-assembly and applications of biomimetic and bioactive peptide-amphiphiles. *Soft Matter* **2**:1015 – 24.

[99] Gray, D.G., Roman, M. (2006) Self-assembly of cellulose nanocrystals: Parabolic focal conic films. In: *Cellulose Nanocomposites: Processing, Characterization, and Properties*, (K. Oksman and M. Sain, Eds.) ACS Symposium Series 938, American Chemical Society, Washington DC, U.S.A., p. 26 – 32.

[100] Wang, X.S., Winnik, M.A., Manners, I. (2006) Synthesis, self-assembly, and applications of polyferrocenylsilane block copolymers. In: *Metal-Containing and Metallosupramolecular Polymers and Materials*, (U.S. Schubert, G.R. Newkome, I. Manners, Eds.) ACS Symposium Series 928. American Chemical Society, Washington DC, p. 274 – 91.

[101] Johnston, M.R., Latter, M.J. (2005) Capsules, cages and three-dimensional hosts: Self-assembly of complementary monomers. *Supramol. Chem.* **17**:595 – 607.

[102] Dumoulin, M., Kumita, J.R., Dobson, C.M. (2006) Normal and aberrant biological self-assembly: Insights from studies of human lysozyme and its amyloidogenic variants. *Acc. Chem. Res.* **39**:603 – 10.

[103] Gazit, E. (2005) Mechanisms of amyloid fibril self-assembly and inhibition. *FEBS J.* **272**:5971 – 8.

[104] Whitesides, G.M., Grzybowski, B. (2002) Self-assembly at all scales. *Science* **295**:2418 – 21.

[105] Pfeifer, R., Lungarella, M., Iida, F. (2007) Self-organization, embodiment, and biologically inspired robotics. *Science* **318**:1088 – 93.

[106] Miller, A.D. (2002) Order for free: Molecular diversity and complexity promote self-organisation. *Chembiochem* **3**:45 – 6.

[107] Gerhart, J., Kirschner, M. (1997) Cells, Embryos, and Evolution. MA: Blackwell Science.

[108] Halley, J.D., Winkler, D.A. (2008) Critical-like self-organization and natural selection: two facets of a single evolutionary process? *Biosystems* **92**:148 – 158.

[109] Lehn, J.-M. (2007) From supramolecular chemistry towards constitutional dynamic chemistry and adaptive chemistry. *Chem. Soc. Rev.* **36**:151 – 60.

[110] Figueiredo, M.A.T. (2003) Adaptive sparseness for supervised learning. *IEEE Trans. Patt. Anal. Mach. Intell.* **25**:1150 – 9.

[111] Winkler, D. (2001) The broader applications of neural and genetic modelling methods. *Drug Disc. Today* **6**:1198 – 9.

[112] Winkler, D.A. (2004) Neural networks in ADME and toxicity prediction. *Drugs Future* **29**:1043 – 57.

[113] Winkler, D.A. (2004) Neural networks as robust tools in drug lead discovery and development. *Mol. Biotech.* **27**:139 – 67.

[114] Winkler, D.A., Burden, F.R. (2000) Robust QSAR models from novel descriptors and Bayesian Regularised Neural Networks. *Mol. Sim.* **24**:243.

[115] Winkler, D.A., Burden, F.R. (2002) Application of neural networks to large dataset QSAR, virtual screening and library design. In: *Combinatorial Chemistry Methods and Protocols* (Ed. Bellavance-English, L.). Humana Press.

[116] Winkler, D.A., Burden, F.R. (2004) Bayesian Neural Networks for Modelling in Drug Discovery. *Biosilico* **2**:104 – 11.

[117] Polley, M.J., Winkler, D.A., Burden, F.R. (2004) Broad-based quantitative structure-activity relationship modeling of potency and selectivity of farnesyltransferase inhibitors using a Bayesian regularized neural network. *J. Med. Chem.* **47**:6230 – 8.

[118] Wang, H., Chen, B., Yao, S.Z. (2006) Quantitative structure-activity relationship modelling of angiotensin converting enzyme inhibitors by back propagation artificial neural network. *Chin. J. Anal. Chem.* **34**:1674 – 8.

[119] Fernandez, M., Carreiras, M.C., Marco, J.L., Caballero, J. (2006) Modelling of acetylcholinesterase inhibition by tacrine analogues using Bayesian-regularized Genetic Neural Networks and ensemble averaging. *J. Enz. Inhib. Med. Chem.* **21**:647 – 61.

[120] Fernandez, M., Caballero, J. (2006) Bayesian-regularized genetic neural networks applied to the modelling of non-peptide antagonists for the human luteinizing hormone-releasing hormone receptor. *J. Mol. Graph. Mod.* **25**:410 – 22.

[121] Arakawa, M., Hasegawa, K., Funatsu, K. (2006) QSAR study of anti-HIV HEPT analogues based on multi-objective genetic programming and counter-propagation neural network. *Chemom. Intell. Lab. Sys.* **83**:91 – 8.

[122] Polley, M.J., Burden, F.R., Winkler, D.A. (2005) Predictive human intestinal absorption QSAR models using Bayesian regularized neural networks. *Aust. J. Chem.* **58**:859 – 63.

[123] Winkler, D.A., Burden, F.R. (2004) Modelling blood-brain barrier partitioning using Bayesian neural nets. *J. Mol. Graph. Mod.* **22**:499 – 505.

[124] Jung, E., Kim, J., Kim, M., Jung, D.H., Rhee, H., Shin, J.M. Choi, K., Kang, S.K., Kim, M.K., Yun, C.H., Choi, Y.J., Choi, S.H. (2007) Artificial neural network models for prediction of intestinal permeability of oligopeptides. *BMC Bioinformatics* **8**:245.

[125] Di Fenza, A., Alagona, G., Ghio, C., Leonardi, R., Giolitti, A., Madami, A. (2007) Caco-2 cell permeability modelling: a neural network coupled genetic algorithm approach. *J. Comp.-Aided Mol. Des.* **21**:207 – 21.

[126] Chen, L.J., Lian, G.P., Han, L.J. (2007) Prediction of human skin permeability using artificial neural network (ANN) modelling. *Acta Pharm. Sinica.* **28**:591 – 600.

[127] Doytchinova, I.A., Flower, D.R. (2003) Towards the in silico identification of class II restricted T-cell epitopes: a partial least squares iterative self-consistent algorithm for affinity prediction. *Bioinfo*. **19**:2263 – 70.

[128] Winkler, D.A., Burden, F.R. (2005) Predictive Bayesian Neural Network Models of MHC Class II Peptide Binding. *J. Mol. Graph. Mod.* **23**:481 – 9.

[129] Batouche, S., Rebbani, N., Gheid, A. (2006) Artificial neural network and topological indices to predict retention indices in gas chromatography. *Asian J. Chem.* **18**:2623 – 36.

[130] Liu, G.S., Yu, J.G. (2005) QSAR analysis of soil sorption coefficients for polar organic chemicals: Substituted anilines and phenols. *Water Res.* **39**:2048 – 55.

[131] Tetko, I.V., Bruneau, P. (2004) Application of ALOGPS to predict 1-octanol/water distribution coefficients, logP, and logD, of AstraZeneca in-house database. *J. Pharm. Sci.* **93**:3103 – 10.

[132] Eros, D., Kovesdi, I., Orfi, L., Takacs-Novak, K., Acsady, G., Keri, G. (2002) Reliability of logP predictions based on calculated molecular descriptors: A critical review. *Curr. Med. Chem.* **9**:1819 – 29.

[133] Arupjyoti, S., Iragavarapu, S. (1998) New electrotopological descriptor for prediction of boiling points of alkanes and aliphatic alcohols through artificial neural network and multiple linear regression analysis. *Comp. Chem.* **22**:515 – 22.

[134] Zhang, R.S., Liu, S.H., Liu, M.C., Hu, Z. (1997) Neural network molecular descriptors approach to the prediction of properties of alkenes. *Comp. Chem.* **21**:335 – 41.

[135] Yan, A.X., Gasteiger, J., Krug, M., Anzali, S. (2004) Linear and nonlinear functions on modeling of aqueous solubility of organic compounds by two structure representation methods. *J. Comp.-Aided Mol. Des.* **18**:75 – 87.

[136] Huuskonen, J., Rantanen, J., Livingstone, D. (2000) Prediction of aqueous solubility for a diverse set of organic compounds based on atom-type electrotopological state indices. *Eur. J. Med. Chem.* **35**:1081 – 8.

[137] Burden, F.R., Winkler, D.A. (2000) A quantitative structure-activity relationships model for the acute toxicity of substituted benzenes to *Tetrahymena pyriformis* using Bayesian-regularized neural networks. *Chem. Res. Toxicol.* **13**:436 – 40.

[138] Melagraki, G., Afantitis, A., Sarimveis, H., Igglessi-Markopoulou, O., Alexandridis, A. (2006) A novel RBF neural network training methodology to predict toxicity to *Vibrio fischeri*. *Mol. Divers.* **10**:213 – 21.

[139] Caballero, J., Garriga, M., Fernandez, M. (2005) Genetic neural network modelling of the selective inhibition of the intermediate-conductance Ca^{2+}-activated K^{+} channel by some triarylmethanes using topological charge indexes descriptors. *J. Comp.-Aided Mol. Des.* **19**:771 – 89.

[140] Vracko, M., Mills, D., Basak, S.C. (2004) Structure-mutagenicity modelling using counter propagation neural networks. *Environ. Toxicol. Pharmacol.* **16**:25 – 36.

[141] Mazzatorta, P., Vracko, M., Jezierska, A., Benfenati, E. (2003) Modeling toxicity by using supervised Kohonen Neural Networks. *J. Chem. Info. Comp. Sci.* **43**:485 – 92.

[142] Admans, G., Takahashi, Y., Ban, S., Kato, H., Abe, M., Hanai, S. (2001) Artificial neural network for predicting the toxicity of organic molecules. *Bull. Chem. Soc. Jap.* **74**:2451 – 61.

[143] Keller, E.F. (2005) Revisiting "scale free" networks. *Bioessays* **27**:1060 – 1068.

[144] Geard, N., Wiles, J. (2005) A gene network model for developing cell lineages. *Artif. Life* **11**:249 – 67.

[145] Winkler, D.A., Burden, F.R., Halley, J.D. (2008) Using recursive networks to describe and predict gene expression and fate decisions during *C. elegans* embryogenesis. *Artif. Life* in press.

[146] Weaver, D.C. (1999) Modelling regulatory networks with weight matrices. *Pac. Symp. Biocomp.* **4**:112 – 23.

[147] Hussain, M.A., Theise, N.D. (2004) Post-natal stem cells as participants in complex systems and the emergence of tissue integrity and function. *Pediatric Diabetes* **5**:75 – 8.

[148] Laslett, A.L., Grimmond, S., Gardiner, B., Stamp, L., Lin, A., Hawes, S.M., *et al.* (2007) Transcriptional Analysis of Early Lineage Commitment In Human Embryonic Stem Cells. *BMC Dev. Biol.* **7**:12 – 30.

[149] Welch, J.J., Watts, J.A., Vakoc, C.R., Yao, Y. (2004) Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood* **104**:3136 – 47.

[150] Hamatani, T., Carter, M.G., Sharov, A.A., Ko, M.S.H. (2004) Dynamics of Global Gene Expression Changes during Mouse Preimplantation Development. *Dev. Cell* **6**:117 – 31.

[151] Hart, C.E., Mjolsness, E., Wold, B.J. (2006) Transcription network: Inferences from neural networks. *PloS Comp. Biol.* **2**:1592 – 607.