Systems biology meets chromatin function

Workshop on Nuclear Organization

Astrid E. Visser¹* & David A. Fell²

¹University of Leiden, Leiden, The Netherlands and ²Oxford Brookes University, Oxford, UK

The 4th EMBO Workshop on Nuclear Organization took place between 12 and 15 October 2006 in Gosau, Austria, and was organized by D. Jackson, R. van Driel, H. Lipps and H. Westerhoff.

Keywords: chromatin; systems biology; network; cell nucleus; dynamics

both in terms of the number of interacting components and the spatial heterogeneity. In other areas of cell function that are carried out by multi-enzyme networks, the question of how function arises from the interaction of the relevant components is being addressed by various combined theoretical and experimental approaches termed 'systems biology'. Such an approach could determine the extent to which known interactions explain chromatin function and identify what remains to be understood. Furthermore, the combination of mathematical and computational modelling with quantitative measurements fulfils an urgent requirement, as it has the rigour to derive robust predictions that can be precisely tested against experimental observations to support or disprove hypotheses. This meeting brought together experts in chromatin function and systems biology to inform each other of the approaches and issues in their respective fields, and to explore whether they are ready to interact to develop the systems biology of chromatin function. Inevitably, few of the participants could claim to span both fields, but the two opening plenary talks by H. Meinhardt (Tübingen, Germany) and A. Blais (New York, NY, USA) achieved this, and illustrated contrasting approaches to studying chromatin function through systems biology.

EMBO reports (2007) 8, 446–450. doi:10.1038/sj.embor.7400960

Introduction

It is clear that DNA sequence and transcription factor availability alone are not sufficient for effective gene regulation in eukaryotes. Epigenetic factors at various levels also have an essential role: for example, DNA methylation and histone modifications form the molecular basis of gene regulation by creating chromatin microenvironments that promote or prevent transcription. Gene activation requires complete cascades of chromatin modifiers to prepare the chromatin for transcription. In addition, the topology of the chromatin and the dynamic interactions between nuclear bodies influence genome function. How the appropriate and controlled expression of genes is achieved is clearly a complicated system,

Gene networks

The elucidation of gene networks can define components and their patterns of interactions as a basis for a systems biology description. Network models can be generated from top-down or bottom-up perspectives. Top-down networks are based on combining many sets of genome-wide data, as illustrated by Blais for skeletal muscle (Blais & Dynlacht, 2005). By using muscle-specific microarrays, messenger RNA (mRNA) expression and occupancy of muscle-specific transcription factors were assayed under various physiological conditions. These data were combined with bioinformatics descriptors, such as the Gene Ontology and predictions of transcription factor target sequences. The resulting muscle-specific network constructed by Blais and colleagues revealed new insights into processes involved in muscle-cell differentiation, exit from the cell cycle and how these cells cope with physiological stress.

An alternative is the bottom-up approach, in which models are built with the aim of determining the minimal requirements to describe a particular phenomenon or system. Meinhardt illustrated how relatively simple models containing nonlinear feedback and

¹Leiden Institute of Chemistry, University of Leiden, PO Box 9502, 2300 RA Leiden, The Netherlands

²School of Life Sciences, Oxford Brookes University, Headington, Oxford OX3 0BP, UK

*Corresponding author. Tel: +31 (0)71 527 4432; Fax: +31 (0)71 527 4340;

E-mail: a.visser@chem.leidenuniv.nl

Submitted 19 January 2007; accepted 14 March 2007; published online 13 April 2007

feed-forward loops can explain the basis of pattern formation during embryogenesis (Fig 1). The basis lies in a short-range—in either space or time—auto-amplifying activator in combination with a long-range inhibitor. The formal structure of these models is also valid for the description of gene regulation at the molecular level. Such models explain the counter-intuitive observation that during activation, inhibitors of the process can also become elevated. Such models do not aim to describe all reactions but rather focus on the minimum core components that are required to describe the behaviour of the system (Meinhardt, 1982).

Systems biology encompasses both top-down and bottom-up approaches. Although they require different experimental and theoretical techniques, they are complementary and should eventually meet to allow the formation of a merged description. Bottom-up descriptions provide a clear inferential chain from components and their properties to system behaviour, but are sensitive to errors of omission in model formulation if relevant components or interactions are overlooked. Top-down approaches attempt to identify all components associated with system behaviour, often at ever-increasing levels of granularity, but might not be revealing about the important properties and mechanisms of the components.

The rest of the talks clustered around three main themes: recent results on the experimental definition of the molecular components, interactions and dynamics involved in gene activation and chromosomal organization; the development of experimental techniques to determine the physical parameters of chromatin and the nucleus, to investigate the structural organization and heterogeneity within the nucleus, and to monitor the dynamics and interactions of individual components; and examples of modelling and analytical approaches applied to a range of multi-component networks including metabolism, signal transduction and the cell cycle.

Chromatin biology

A recurrent theme addressed by both experimentalists and modelers was the coordinated action and balance between activators and inactivators. G. Reid (Heidelberg, Germany) described the composition and sequential recruitment of chromatin remodelling complexes to the oestrogen-receptor-responsive gene *pS2*. Gene activation was dependent on both activating—such as histone acetyltransferases (HATs)—and silencing—such as histone deacetylases and methyl CpG-binding protein 2—complexes that set up cycles of transcription. Importantly, Reid showed that as with histone acetylation, DNA methylation can also be dynamic, as the promoter DNA was transiently methylated in between transcriptionally active phases. J. Mellor (Oxford, UK) reviewed the coordinated actions of chromatin remodellers and histone modifications in transcription initiation, elongation control and re-setting of the chromatin structure to prevent aberrant initiation from cryptic start sites. Mellor made the point that steady-state experiments will give only half the information, that is whether particular modifications occur, whereas chase experiments elucidate a strict order of sequential events (Mellor, 2005). New studies on yeast 14-3-3 homologues Bmh1 and Bmh2 further emphasized that histone modifications do not occur independently. Mellor showed that Bmh1, as with human 14-3-3 proteins, binds histone H3 phosphorylated at serine 10 (H3S10p) *in vivo*. Depletion of Bmh1 and Bmh2 resulted in a complete lack of acetylation of histone H3 at lysines 14 and 18 (H3K14ac and H3K18ac, respectively) and of trimethylation of lysine 4 (H3K4me3). Together with further

Fig 1 | An example of a nonlinear feedback-induced switch. Gene product *g* (red circle) has nonlinear feedback on the activation of its own gene. Morphogen *m* (green circle) initiates gene transcription. At low levels of *g* and *m*, *g* remains at its low steady-state level as an increase in *g* is countered by the degradation term (with rate constant μ). Adding morphogen induces additional synthesis of *g*, which is reinforced by the sigmoidal self-activation of *g* with rate constant *c* and affinity term κ , so that above a threshold value, synthesis dominates degradation and *g* increases to a high steady-state level that can be maintained when *m* is withdrawn. This sequence of events in the evolution of the concentration of *g* (red) in response to a morphogen gradient (green) is simulated in the lower panels, resulting in two states of *g* after the morphogen disappears. Image provided by H. Meinhardt.

studies, a hierarchical, one-way, modification cascade was proposed in which the 14-3-3 homologues bind to H3S10p and facilitate H3K14 acetylation, which, in turn, is required for H3K18ac and H3K4me3. Bmh1 and Bmh2 were speculated to act by configuring the histone tail for any HAT.

Similarly, F. Stewart (Dresden, Germany) showed, by using mass spectrometry of *Schizosaccharomyces pombe* histones, that several modifications occur only in combination. Furthermore, Stewart presented a model of how trithorax (TRX), which is required for the correct expression of homeotic genes, can make a reversible event irreversible. Initially, a TRX-responsive gene is activated by a transcription factor. Once TRX has been recruited, it can maintain gene activity after the transcription factor has been removed. TRX is an extremely large protein with many DNA-binding sites; therefore, Stewart speculated that TRX ‘caterpillars’ around DNA polymerase such that one end binds a duplicated DNA strand before the other end is displaced. TRX would therefore never be fully removed from TRX-responsive sequences, facilitating its stable localization during subsequent cell generations.

Higher-order structures of chromatin and its nuclear position also affect genome function (Lancôt *et al*, 2007). D. Rhodes (Cambridge, UK) reconstituted long nucleosomal arrays and determined their structure by using electron microscopy. Nucleosomes are known to coil into a 30 nm helical structure known as the 30 nm fibre. Rhodes

showed that these fibres are denser than originally believed and contain 10–12 nucleosomes per gyre rather than the formerly accepted six nucleosomes. Acetylation of 30% of the histone H4 lysine 16 (H4K16) residues was sufficient for fibre disruption. Curiously, reconstituting DNA with H4K16Q mutant histones, which mimic the electrostatic effect of acetylation, did not disrupt the fibres to the same extent as acetylation. S. Gasser (Basel, Switzerland) reported estimates of 10 nucleosomes per DNA gyre on the basis of the persistence length of chromatin in yeast *in situ*. The persistence length is the short distance over which a flexible polymer, such as DNA or chromatin, behaves as a stiff rod. The persistence length was determined by measuring the physical distances between distinct genomic loci in yeast nuclei. Gasser further showed that the rapid dynamics of genomic loci observed in yeast could be described as a 'random walk' movement in a constrained space. The constraint was shown to arise mostly because the loci were part of a long DNA fibre, the chromosome. When the loci were excised from the chromosome, the resulting plasmids moved throughout the entire nucleus except the nucleoli. Movement of chromosomal loci was dependent on ATP but not on RNA elongation. Importantly, the targeting of chromatin remodeller INO80 to the locus increased its mobility. This led Gasser to speculate that the chromatin remodelling that precedes transcription might change mobility by altering the persistence length of the chromatin.

Similar factors might also contribute to chromatin dynamics in mammalian cells. By using live cell fluorescence imaging, D. Jackson (Manchester, UK) showed that chromatin domains that had replicated during early S-phase (representing euchromatin) crossed distances of 0.3–0.75 μm during a 15-min period, which is similar to the mobility of chromatin in yeast. By contrast, domains labelled in mid-S-phase, which represent more condensed heterochromatin, were generally constrained to 0.1–0.4 μm and moved with half the speed. By using RNA fluorescence *in situ* hybridization and immunofluorescence imaging, P. Fraser (Cambridge, UK) showed that on activation, genes moved to pre-existing transcription factories across distances of at least 0.5 μm . Genes in the same metabolic pathway for haem synthesis were frequently transcribed in the same transcription factory. This colocalization was dependent on transcription and was confirmed by chromosome conformation capture (3C) experiments. Fraser proposed that when co-regulated genes require an overlapping set of transcription factors for activation, they frequently share the same transcription factory, even when the genes are located on different chromosomes. He speculated that this spatial self-organization is based on a local enrichment of concentration-limited transcription factors near a recently activated target gene and hence in a specific transcription factory. A simplified computer simulation visualized this concept and initiated a discussion emphasizing that quantitative measurements of, for example, diffusion rates, concentrations and *in vivo* affinity are essential for reliable modelling.

Aberrations in the nuclear lamina alter chromatin function and have been linked to several diseases (Broers *et al.*, 2006). Jackson reported that lamin B1 depletion—but not lamin B2 or lamin A/C depletion—inhibited RNA synthesis and DNA replication, and caused chromosomes to collapse towards the nuclear periphery. D. Vaux (Oxford, UK) performed comparative microarray analysis on mouse cells lacking lamin B or lamin B-processing enzymes such as Rce1. The genes that were subsequently dysregulated did not cluster functionally; however, three genes were spaced closely

on chromosome 18. This chromosome was shown to locate preferentially near the nuclear periphery and to lose this preferred position in cells lacking lamin B or Rce1. By contrast, the genes of chromosome 19 were not dysregulated by lamin B knockout, and this chromosome retained its central position independent of the lamin B status. Despite the differences between human and mouse cells, which might reflect differences in the relative composition of the lamina, these data indicate an involvement of organizational proteins such as lamin B in gene regulation. M. Fornerod (Amsterdam, The Netherlands) directly addressed which genomic regions were physically close to lamin B by using the Dam identification (DamID) technique in *Drosophila*; these were mainly widely spaced, inactive genes. Fornerod showed that several of the lamin B-associated genes are part of developmentally regulated units. Furthermore, his data suggest a fractal organization of the chromatin fibre, which is non-random and accessible at all levels of compaction.

A systems biology approach to chromatin function depends on detailed and quantitative information about the location, levels and dynamics of specific components in individual cells. A. Sigal (Rehovot, Israel) described fluorescent labelling of endogenous proteins in individual human cells and quantitatively monitored the proteins during the course of five cell generations. After correcting for cell-cycle fluctuations, autocorrelation analysis showed that expression levels were not constant and that correlation decayed at rates that ranged from one to several cell generations. Although the levels of pairs of proteins in different pathways did not correlate, reasonable correlations existed for proteins in the same cellular pathway. This indicates an 'expression memory' that can span mitosis and involves not just a single protein but an entire system. Cell-to-cell variability suggests a potential epigenetic variability, which is generally undetected as most techniques measure pools of cells. A. Visser (Leiden, The Netherlands) made a case for reading epigenetic patterns on individual DNA molecules and showed a technical development towards reading DNA methylation on individual DNA fibres. J. Langowski (Heidelberg, Germany) described the use of single-molecule fluorescent resonance energy transfer, fluorescence correlation microscopy and continuous fluorescence microphotolysis of fluorescence probes for measuring the mobility and conformation of nucleosomes, and the mobility of free and bound fractions of proteins at specific locations in the cell. Importantly, he confirmed that the diffusion rate of proteins in the nucleus is only about five times slower than in water, and is therefore similar to the diffusion rate of proteins in the cytoplasm. He also showed that the chromatin network is highly penetrable for proteins with a molecular mass of less than 200,000. Such measurements provide parameter estimates for his Monte-Carlo models of the Brownian motion of the chromatin fibre, which are being used to distinguish between stable and unstable modes of packing in chromatin.

Systems biology

D. Fell (Oxford, UK) made the case for metabolic control analysis as a paradigm for studying the systems biology of enzyme networks. This approach takes into account the kinetics of all enzymes involved in a functional system and not just the one that was previously considered to be the 'rate-limiting' component (Fell, 1997). It allows the sensitivity of a metabolic function to alterations in enzyme activity to be determined both experimentally and theoretically. It also explains why even essential enzymes often need to be almost fully

Fig 2 | Modelling pipeline that creates a systems biology network. (A) Initially, a network model is formulated based on the literature and expert knowledge. Properties of small modules (for example, the mitogen-activated protein kinase (MAPK) cascade or metabolism) are studied and quantitative experimental data are collected, such as mRNA levels, modifications and concentration changes over time. Model parameters are estimated, and individual reaction and system equations are formulated (r , number of reactions; S_i , metabolite concentration; v_j , reaction rates; n_{ij} , stoichiometric coefficients). Computer simulations are then performed to evaluate agreement between model and experiment (data from Klipp *et al.*, 2005). (B) Iterative cycles of experiment-driven modelling and model-driven experiments: a computer simulation of initially formulated models results in predictions that are compared with experimental data (yellow boxes). Disagreements and agreements between predictions and experimental data lead to refinement of the model (orange boxes). Subsequent sensitivity analysis predicts which parameters have most influence on the system. New experiments are designed to focus on and manipulate these parameters (red box). Comparing the obtained data and computer predictions leads to further refinements of the model and focused experimental designs. Fps1, fdp1 suppressor; GPD1, gene for glycerol-3-phosphate dehydrogenase; gpd1Δ, mutant lacking GPD1; Hog1, high osmolarity glycerol response; Pbs2, polymyxin B sensitivity; Sln1, synthetic lethal of N-end rule; Ssk, suppressor of sensor kinase; Ypd1, tyrosine (Y) phosphatase dependent.

inhibited before changes in function become significant, with partial inhibition leading only to mild changes. Similar principles could also be applied to models of signal transduction pathways and the cell cycle to explain the sometimes apparently paradoxical action of drugs, such as the inhibition of an enzyme resulting in an increase in the level of its active form. H. Westerhoff (Manchester, UK) showed how metabolic control analysis could be extended to dissect the relative contributions of changes in enzyme expression and intrinsic adjustments in the metabolic network to allow metabolism to adapt to meet different demands. Combined experimental and theoretical analysis by Westerhoff illustrated how the control of supercoiling of the bacterial chromosome is divided between different factors, including the activity of topoisomerases and the influence of supercoiling on the transcription of these enzymes.

E. Klipp (Berlin, Germany) provided excellent insights into the steps involved in building, defining parameters for and experimentally testing bottom-up models on the basis of her work on the stress response in yeast (Fig 2). In addition, R. van Driel (Amsterdam, The Netherlands) stressed the importance of reiterative cycles of experiment-driven modelling and model-driven experimentation to design a smaller number of more informative experiments rather than evaluating all potential hits in genome-wide assays.

Two further talks illustrated additional approaches to studying the systems biology of signal transduction systems. M. White (Liverpool, UK) described a joint experimental and theoretical study of nuclear factor-κB (NF-κB) signalling in mammalian cells. He used fluorescent imaging to study the temporal response of single cells to tumour necrosis factor α (TNFα) and could show subsequent

oscillations in the level of NF-κB in the nucleus. This would not have been possible to observe in a cell population because of a lack of synchronization. A computer model of the oscillations was developed, and further experiments at different doses of TNFα were carried out to refine the model and its parameters. This work suggests that the frequency of the oscillations is also involved in regulating gene expression. M. Kimmel (Houston, TX, USA) further analysed the variability of gene expression responses to NF-κB signalling by using stochastic modelling. He identified two main sources of variability in the responses—the amplification at the translational level of the burst of mRNA molecules produced by gene activation and, at low doses of signal, the random fluctuations in the number of activated receptors. This model exemplifies the phenomenon of stochastic robustness, in which individual cells respond differently to the same stimulus although within an individual cell the response is well-defined.

B. Novak (Budapest, Hungary) illustrated how his computer model of the eukaryotic cell cycle disproves the often-repeated claim that the irreversible transitions of the cycle are ensured by proteolytic degradation of crucial components. In fact, the levels of components such as cyclin B are set by a dynamically reversible process formed by the two irreversible processes of synthesis and degradation. In each of the irreversible transitions of the cycle—G1/S, G2/M and mitotic exit—it is the existence of positive-feedback effects that creates alternative stable states for the concentrations of components, such as cyclin B, with the transition to the low concentration state being abrupt and irreversible. In other words, the irreversibility is a systems property that cannot be ascribed to any

individual component. Interestingly, Meinhardt proposed positive feedback and alternative stable states as a model for irreversible switches of gene expression during development (Fig 1).

Conclusions

Overall, the meeting provided several examples in which theoretical and modelling approaches in conjunction with experiments have furthered our understanding of chromatin function. However, it was also apparent that the eukaryotic nucleus and the control of chromatin function pose greater challenges for a systems biology modelling approach than many of the other cellular networks described in the talks. Issues that are similar to those that are likely to arise in modelling chromatin function have been addressed before, but usually one by one. However, several different complications will have to be dealt with simultaneously in the creation of realistic models of the control of chromatin function. For example, the low concentration of some species, such as the one or two copies per cell of a specific site on the DNA, require the use of stochastic simulation techniques; however, the relatively higher concentrations of histones and some metabolites (up to millimolar) limit the ability to simulate the whole system in that way. In these cases, models based on continuous concentrations and differential equations would be preferable, but then the heterogeneity of the nuclear environment could cause difficulties. Such problems have been encountered before in simulations of signal transduction and specifically of bacterial chemotaxis (Bray *et al*, 2007), but in these cases the events catalysed by the enzymes were simple chemical transitions compared with eukaryotic gene transcription.

Nevertheless, the meeting gave grounds for believing that the systems biology of chromatin function is a worthwhile goal, and, far from being out of reach, is already emerging (Gorski & Misteli, 2005). Indeed, the large number of components involved in chromatin structure and function, and the highly structured but heterogeneous environment in the nucleus make it a necessary development, as traditional qualitative modes of biological

explanation are inadequate in the face of such complexity. It will be highly challenging, and will require the formation of multidisciplinary collaborations to ensure a productive interaction between the theoretical and experimental strands. The next step is to find a mechanism to foster such collaborations.

REFERENCES

- Blais A, Dynlacht BD (2005) Constructing transcriptional regulatory networks. *Genes Dev* **19**: 1499–1511
- Bray D, Levin MD, Lipkow K (2007) The chemotactic behavior of computer-based surrogate bacteria. *Curr Biol* **17**: 12–19
- Broers JL, Ramaekers FC, Bonne G, Yaou RB, Hutchison CJ (2006) Nuclear lamins: laminopathies and their role in premature ageing. *Physiol Rev* **86**: 967–1008
- Fell D (1997) *Understanding the Control of Metabolism*. London, UK: Portland
- Gorski S, Misteli T (2005) Systems biology in the cell nucleus. *J Cell Sci* **118**: 4083–4092
- Klipp E, Nordlander B, Krueger R, Gennemark P, Hohmann S (2005) Integrative model of the response of yeast to osmotic shock. *Nat Biotechnol* **23**: 975–982
- Lancôt C, Cheutin T, Cremer M, Cavalli G, Cremer T (2007) Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet* **8**: 104–115
- Meinhardt H (1982) *Models of Biological Pattern Formation*. London, UK: Academic Press
- Mellor J (2005) The dynamics of chromatin remodelling at promoters. *Mol Cell* **19**: 147–157

Astrid E. Visser

David A. Fell