Ferran Azorín

Chromatin structure and function

Changes in chromatin structure play a fundamental role in the regulation of multiple genomic processes, from gene expression to chromosome segregation and the maintenance of genome integrity and stability. Increasing evidence indicates that alterations in chromatin structure and function are at the root of several human pathologies, including some types of cancer and neurological disorders. Our current knowledge about the regulation of chromatin functions has benefited from the identification of components and mechanisms that covalently and structurally modify chromatin. These include chromatin assembly and remodelling complexes, histone modifications (eg, acetylation, methylation, phosphorylation, ubiquitination, etc) and the corresponding enzymes (eg, HATs, HDACs, HMTs, HDMs, etc), non-histone proteins that recognise specific histone modifications and contribute to the establishment of distinct functional domains (eg, HP1, PC, etc), histone variants that localise to specific chromosomal locations (eg, CenH3, H3.3, H2A.Z, macroH2A, etc), and non-coding RNAs that modify chromatin structure and regulate gene expression. Our research focuses on the study of the molecular basis of chromatin function and its regulation. More precisely, we seek to elucidate the contribution of chromatin to the control of: (i) centromere identity and function; (ii) gene expression and (iii) long-distance genomic interactions.

Centromere identity and function

Centromere function ensures accurate chromosome segregation during mitosis and meiosis, as the centromere dictates the assembly of the kinetochore. This chromosomal protein structure, in turn, regulates the spindle attachment checkpoint (SAC), which delays anaphase onset until all chromosomes are correctly attached in a bipolar fashion to the mitotic spindle. In all eukaryotes, centromeres are characterised by the presence of a specific histone H3 variant (CenH3), which is essential for centromere function and cell viability. CenH3 dictates kinetochore assembly, being required for the localisation of all other inner as well as outer kinetochore proteins tested to date. However, the precise molecular interactions underlying the contribution of CenH3 to kinetochore assembly are not understood. Likewise, mechanism(s) that ensure centromere-only deposition of CenH3 are only partially known. We are currently addressing these questions.

Regulated proteolysis restricts centromeric localisation of CenH3

Centromere identity is determined by the deposition of CenH3, which replaces canonical H3.1 in all eukaryotic centromeres. CenH3 is exclusively found at centromeres. We are addressing how centromere-specific deposition of CenH3 is achieved in *Drosophila*. Contrary to canonical histones, which are deposited during DNA replication, CenH3 incorporates into chromatin independently of DNA replication (see Figure 1). In human cells, the deposition of newly synthesised CenH3^{CENP-A} occurs during mitosis, at late telophase, and early G1. In syncytial *Drosophila* embryos, where no G1/G2-phases are observed, CenH3^{CID} deposition also takes place during mitosis, at anaphase. As a result

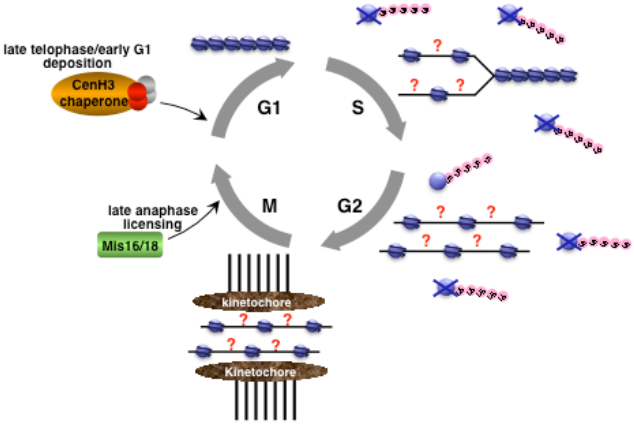

Figure 1. ***Assembly and dynamic behaviour of CenH3-chromatin during the cell cycle.*** *Like other histone variants, CenH3 incorporates into chromatin independently of DNA replication. Deposition of newly synthesised CenH3 takes place during mitosis, at late telophase, or early G1. Specific CenH3-chaperones localise to the centromere coincidentally with the deposition of new CenH3 and mediate the assembly of CenH3-nucleosomes. During assembly, CenH3 might become resistant to proteolysis, which otherwise degrades CenH3 and prevents deposition at non-centromeric sites. Prior to CenH3 deposition, at late anaphase, specific complexes (Mis16/Mis18) appear to modify centromeric chromatin to allow assembly of new CenH3 nucleosomes. During DNA-replication at S-phase, CenH3 concentration at centromeres is diluted and kinetochore assembly occurs before replenishment with new CenH3 nucleosomes. It is unclear whether 'gaps' generated during DNA replication remain nucleosome-free or are filled by replicative H3 nucleosomes. It may also be that CenH3 nucleosomes are disassembled into 'half-nucleosomes' to compensate for this deficit.*

of its loading outside S-phase, CenH3 concentration at centromeres is diluted during DNA replication, thereby generating 'gaps' that can remain nucleosome-free, ready for CenH3 deposition during mitosis, or that can be filled by replicative H3-nucleosomes, which are later replaced by CenH3-nucleosomes. During DNA replication, it is feasible that CenH3-nucleosomes are disassembled into 'half-nucleosomes'. The loading of CenH3 late in mitosis raises several intriguing questions. It implies that kinetochore assembly occurs before centromeres are fully replenished with CenH3-nucleosomes and also that loading occurs in close coincidence with chromosome segregation, thereby suggesting that signalling events occurring during segregation trigger CenH3 deposition. Finally, it raises the question as to how CenH3 assembly takes place during mitosis, when chromatin is believed to be more inaccessible and, in general, refractory to transactions. In this regard, the Mis16/Mis18 complex appears to be involved in licensing the centromeric chromatin for deposition of new CenH3.

The deposition of CenH3 nucleosomes is, however, a promiscuous process, as it may also occur during DNA replication, thereby leading to the mislocalisation of these nucleosomes throughout chromatin. Therefore, additional mechanisms must be present to either prevent the deposition of CenH3-containing nucleosomes at non-centromeric sites during DNA replication and/or to remove them afterwards. We have shown that, in cultured *Drosophila* S2 cells, proteasome-mediated degradation restricts the localisation of $CenH3^{CID}$ to centromeres by removing mislocalised CenH3CID as well as by regulating available $CenH3C^{CID}$ (Moreno-Moreno *et al*, 2006). Moreover, in the fly, proteasome mutants show increased expression and mislocalisation of $CenH3^{CID}$. Proteasome-mediated degradation appears to be an evolutionarily conserved mechanism that regulates available CenH3 to favour its preferential deposition at centromeres. This notion is supported by observation that the levels of $CenH3^{Cse4}$ in the yeast *Saccharomyces cerevisiae* are also regulated by the proteasome and proteolysis-resistant mutants mislocalise throughout chromatin. How is CenH3 proteolysis regulated? In this regard, using a yeast two-hybrid screen, we have identified the interaction of $CenH3^{CID}$ with *partner of paired* (*Ppa*), an F-box-containing protein that interacts with *Skp1*, an evolutionarily conserved component of SCF, an E3 ubiquitin ligase complex. The interaction of $CenH3^{CID}$ with SCF was confirmed *in vitro* by GST-pull down assays, as well as *in vivo*, where SCF mutant conditions result in the overexpression and mislocalisation of $CenH3^{CID}$.

Research Group Members

Group Leader:
Ferran Azorín

Research Associates:
Jordi Bernués, Maria Lluisa Espinas

Postdoctoral Fellows:
Martí Badal, Anne Daulny, Olga Moreno, Mónica Torras

PhD Students:
Marta Batlle, Marta Blanch, Sergio Cuartero, Joan Font, Roman Kessler, Marta Lloret, Sonia Medina, Olivera Vujatovic

Research Assistants:
Carles Bonet, Gemma Molla, Alicia Vera

Lab Technician:
Esther Fuentes

Visiting Student:
Marcia Sofia Ribeiro Lamy (Portugal)

The contribution of CenH3 to kinetochore assembly

CenH3 is essential for centromere function. Cells deficient in CenH3 fail to localise kinetochore proteins and show strong chromosome segregation defects. These observations indicate that CenH3 is required for kinetochore formation. Kinetochore assembly at the centromere involves complex pathways of hierarchical, sometimes reciprocal, interactions. CenH3 is at the bottom of this network of interactions. However, at present, we are only just beginning to understand its true contribution to kinetochore assembly. Kinetochores are large macromolecular entities. Depending on the organism, kinetochores are composed by dozens to more than a hundred different protein components. Biochemical studies have identified a number of protein complexes/networks that act at several stages of assembly, some being constitutively associated with the centromere throughout the cell cycle while others localise to the kinetochore only transiently during mitosis (Figure 2). The complexes identified include the KNM-network (KNL1, NDC80 and MIND), which is involved in microtubule binding, and, in particular, the NAC/CAD-network, which directly associates with centromeric chromatin. In addition, a third protein-network that regulates chromosome movement, Dam1/DASH, has been identified both in *S. cerevisiae* and *S. pombe*. However, the precise molecular interactions that link CenH3 to kinetochore assembly have not been fully elucidated. We currently seek to identify the components that directly bind CenH3 nucleosomes and the features of CenH3 nucleosomes that these components recognise. In this regard, we have reported that deletion of the N-domain of $CenH3^{Cse4}$ in *S. cerevisiae* causes lethality, and mutation analyses have identified a 33 aa motif (END) within the N-domain of $CenH3^{Cse4}$. This motif is essential for viability and interacts with COMA, a kinetochore complex that is functionally related to NAC/CAD. We have also found that, in *Drosophila*, ectopic targeting of $CenH3^{CID}$ results in the recruitment of BubR1, an evolutionarily conserved outer kinetochore protein that is a core component of SAC, and that this interaction is mediated by the N-terminal domain of $CenH3^{CID}$. Despite its remarkable low degree of conservation, CenH3s from *S. cerevisiae*, *S. pombe* and humans show similar properties. Interestingly, in *S. cerevisiae*, END corresponds to an R-rich motif and the presence of R-rich domains at the N-terminal domain is conserved in CenH3s of distant species.

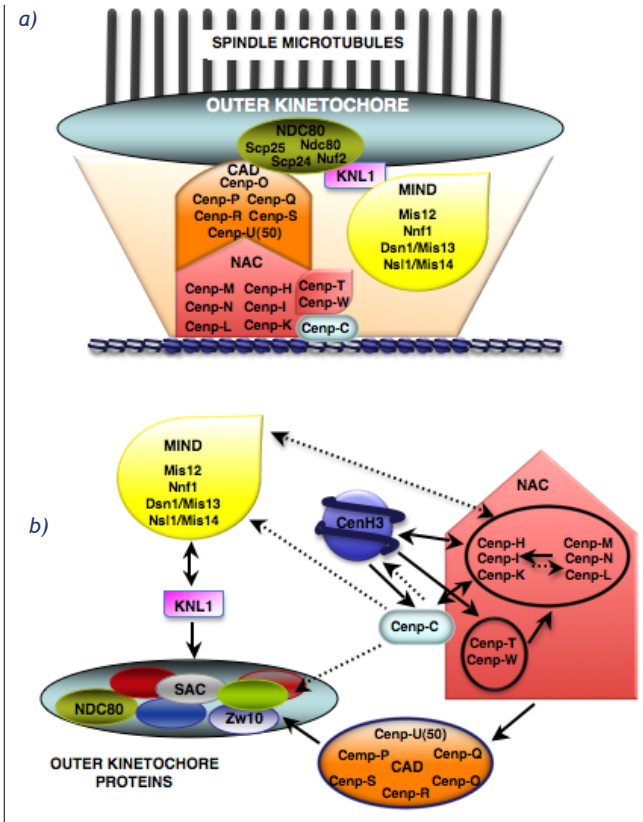

Figure 2. ***Kinetochores are large macromolecular entities.*** *(a) Various protein complexes/networks are known to act at different stages of kinetochore assembly. These include the KNM-network (KNL1, NDC80 and MIND), which is involved in microtubule binding, and the NAC/CAD-network, which directly associates with centromeric chromatin. (b) CenH3 is essential for kinetochore assembly. CenH3 is at the bottom of a complex network of interactions that ultimately leads to assembly of a fully functional kinetochore. Dependencies for centromeric/kinetochore localisation are indicated by solid arrows. Possible interactions, observed only in some species or not fully confirmed, are indicated by dotted arrows.*

Regulation of gene expression: The role of histone lysine (K)-methylation

Most frequently, regulation by chromatin involves the establishment of specific patterns of post-translational histone modifications, which result in the recruitment of regulatory non-histone proteins that bind chromatin and modify its functional state. In this context, methylation at K-residues plays a central role, as it is involved in regulating a wide range of genomic functions, including gene expression.

Characterisation of histone K-demethylases

K-specific histone methyl-transferases (KMTs) were first reported some years ago. However, enzymes capable of antagonising K-methylation (KDMs) have been identified only recently. In the last two years, proteins containing the JumonjiC (JmjC) domain were found to have the capacity to act as KDMs. In this context, we have characterised the demethylase activity of the four JmjC+N-proteins of *Drosophila*: dJARID1/Lid (small imaginal discs), dJMJD2(1), dJMJD2(2) and dJARID2 (Lloret-Llinares *et al*, 2008).

dJARID1/Lid demethylates H3K4me3 and, contrary to what would be inferred from its demethylase activity, contributes to the establishment of transcriptional-competent chromatin states. What are the molecular mechanism(s) underlying the contribution of dJARID1/Lid to transcription activation? The pattern of H3K4me3 depends strongly on the gene-expression status. Thus in active genes H3K4me3 is constrained to a relatively short region spanning the transcription start-site, while inactive genes show no significant enrichment in H3K4me3 throughout the *locus*. How does dJARID1/Lid contribute to the pattern of H3K4me3? Does it bind to coding regions to demethylate H3K4me3, or does the binding take place at promoter regions to regulate dynamic H3K4me3? How is dJARID1/Lid recruited

to chromatin? dJARID1/Lid contains three PHD-fingers (PHD1, PHD2, PHD3) that may mediate binding to methylated-K residues in chromatin. Are they required for binding to chromatin? What methylated-K residues in histones do they bind to? These are some of the questions we are currently addressing.

dJMJD2 demethylates both H3K9me3 and H3K36me3, and dJMJD2(1) influences heterochromatin organisation, as its overexpression in polytene chromosomes induces the spread of HP1 from heterochromatin to euchromatin. The contribution of dJMJD2(1) to heterochromatin organisation is, however, more complex than anticipated, as endogenous dJMJD2(1) is expressed only early in development. In fact, no significant dJMJD2(1) is detected in polytene chromosomes. Moreover, overexpressed dJMJD2(1) is excluded from heterochromatin, localising to multiple euchromatic sites, where it demethylates H3K36me3. In contrast, little is known about the expression, localisation or functional properties of dJMJD2(2). dJMJD2s demethylases may act early in development to establish heterochromatin/euchromatin boundaries. We are currently testing this hypothesis.

The contribution of HP1 proteins to regulation

HP1 is one of the best examples of a regulatory non-histone protein that is recruited to chromatin through the recognition of a specific histone modification, namely di- or tri-methylation of H3K9. HP1 is widely conserved in eukaryotes, with most species containing several isoforms. For instance, *Drosophila* has five HP1 isoforms, three of which are ubiquitously expressed (HP1a, HP1b and HP1c), while the other two (HP1d/Rhino and HP1e) are expressed in the germline. Ubiquitously expressed isoforms show differential chromosomal distributions, as HP1a associates with heterochromatin, while HP1c is excluded from heterochromatin and HP1b is found both at euchromatic and heterochromatic domains. All these observations indicate that current models suggesting that HP1s bind H3K9me2,3 and recruit factors involved in heterochromatin assembly are not likely to apply to all isoforms and functional contexts. In this context, we have shown that, although *Drosophila* HP1c efficiently binds H3K9me2,3 *in vitro*, its binding to chromatin strictly depends on two sequence-specific DNA binding proteins, WOC and ROW, which are putative transcription factors (Font-Burgada *et al*, 2008). HP1c regulates transcription since it extensively co-localises with active RNA polymerase II and H3K4me3, a modification that correlates with active chromatin domains. Moreover, expression-profiling analyses show that HP1c, WOC and ROW extensively co-operate to regulate gene expression during development. To a large extent, the binding of HP1b to euchromatin also depends both on WOC and ROW, which physically interact with HP1b. These findings suggest that HP1c and HP1b co-operate in gene expression regulation (Figure 3). The mechanism(s) behind the contribution of HP1b/c to transcription regulation are currently being studied.

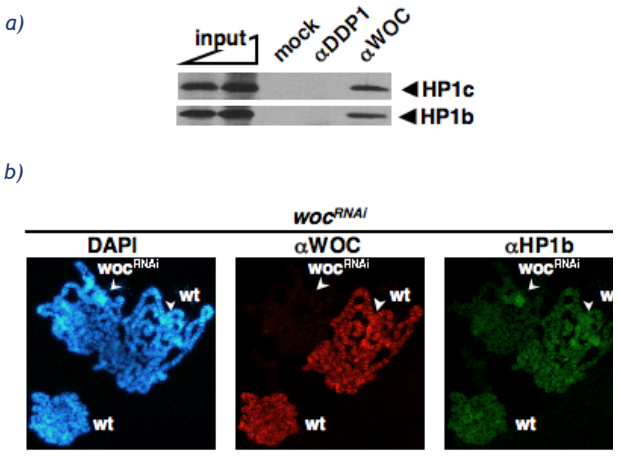

Figure 3. **Drosophila *HP1b interacts with the zinc-finger proteins WOC and ROW.*** *(a) HP1b co-immunoprecipitates with WOC. (b) Binding of HP1b to euchromatin is strongly impaired in the absence of WOC (wocRNAi chromosomes).*

Transcriptional regulation of genes under the control of GAGA factor

GAGA is a *Drosophila* transcription factor involved in many nuclear activities. Expression microarray analyses of GAGA knockdown and overexpression in cells and flies (carried out at the IRB Barcelona Functional Genomics Core Facility) have identified a list of genes that appear to be bona-fide GAGA targets as they are also positive in ChIP-on-Chip experiments. Although we are still analysing the results in flies, with the help of the IRB Barcelona Bioinformatics/Biostatistics Unit, several phenotypical defects have been noted. Among others, growth defects in GAGA-depleted flies that result in smaller wings have been observed in a variety of experimental conditions. We are still analysing other morphological parameters to evaluate whether these statistically significant differences in size (ranging from 10% up to 40%) are general or organ-dependent, and also to assign this and other defects to the genes differentially expressed in the microarray analysis.

Because of the multiplicity of processes in which GAGA factor is involved, we are also studying its molecular dynamics by means of Fluorescence Recovery After Photobleaching (FRAP) and related techniques (with the help of Julien Colombelli at the IRB-PCB Advanced Digital Microscopy Facility). Results obtained in stably transfected cell lines constitutively expressing a GAGA-PAGFP (photoactivatable GFP) fusion protein show that GAGA factor behaves much like a transcription factor. It displays high mobility, lower that free GFP but much higher than core histones, and similar to that shown by linker histone H1. GAGA mobility can be attributed to its binding to DNA because a GAGA mutant at the DBD that abolishes stable interaction with DNA ($GAGA_{DBDKO}$) behaves similarly to GFP (Figure 4). While wt GAGA factor shows a mobile fraction of ~60%, in the case of mutant $GAGA_{DBDKO}$ this fraction increases to ~95%. A ~10-fold difference in their exchange time has also been noted. These results indicate that GAGA factor mobility is restricted by its DNA binding capacity, protein-protein interactions representing a minor contribution, if any. Our results also indicate a half-life estimation of ~13h for wt GAGA factor in cells. This is remarkably long for a transcription factor, thereby suggesting the existence of additional regulatory mechanisms that modulate GAGA activities. A highly dynamic mechanism that may account for this modulation is acetylation. We have shown that, in addition to being phos-

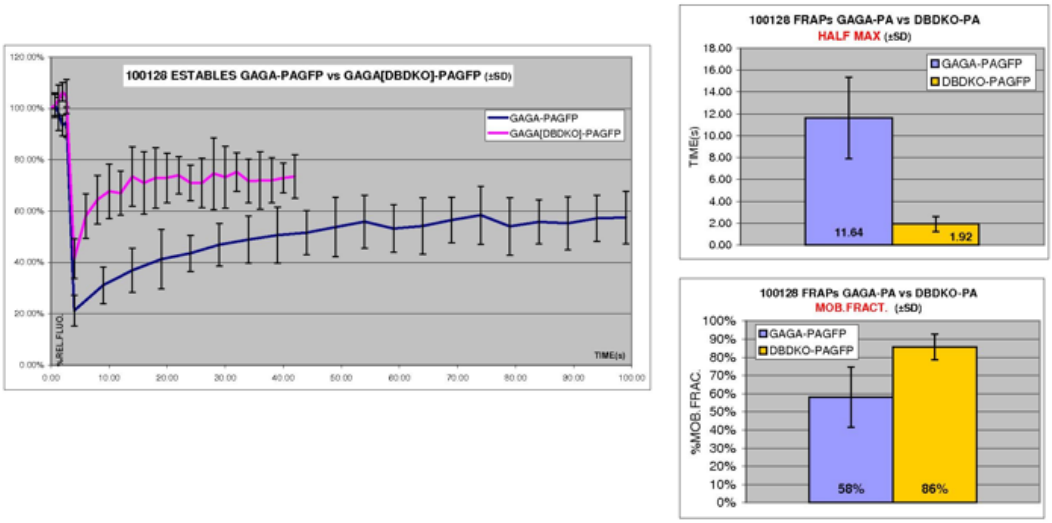

Figure 4. FRAP results of wtGAGA-PAGFP (blue line) and $GAGA_{DBDKO}$-PAGFP (pink line; left panel). Quantification of the exchange half-time and of the mobile fractions for wtGAGA-PAGFP (in blue) and $GAGA_{DBDKO}$-PAGFP (in yellow) are shown in the right upper and lower panels, respectively.

phorylated, GAGA factor is also acetylated *in vivo*. Two lysines at the DNA binding domain that are acetylated by PCAF have been mapped. This acetylation confers a notable reduction in affinity to DNA. Remarkably, while these two lysines are highly conserved in 11 *Drosophila* species, we found that their mutation to glutamines affected neither DNA affinity nor transcriptional activity, thereby suggesting that their conservation is for regulatory purposes.

Scientific output

Publications

Torras-Llort M, Moreno-Moreno O and Azorín F. Focus on the centre: the role of chromatin on the regulation of centromere identity and function. *EMBO J*, **28**(16), 2337-48 (2009)

Other references

Font-Burgada J, Rossell D, Auer H and Azorín F. *Drosophila* HP1c isoform interacts with the zinc-finger proteins WOC and Relative-of-WOC (ROW) to regulate gene expression. *Genes Dev*, **22**(21), 3007-23 (2008)

Lloret-Llinares M, Carré C, Vaquero A, de Olano N and Azorín F. Characterisation of *Drosophila melanogaster* JmjC+N histone demethylases. *Nucleic Acids Res*, **36**(9), 2852-63 (2008)

Moreno-Moreno O, Torras-Llort M and Azorín F. Proteolysis restricts localization of CID, the centromere-specific histone H3 variant of *Drosophila*, to centromeres. *Nucleic Acids Res*, **34**(21), 6247-55 (2006)

Research networks and grants

Ayuda complementaria al proyecto europeo 'Vectores episomales como sistemas de modificación genética para aplicaciones terapéuticas'
Spanish Ministry of Science and Innovation, BIO2006-26123-E (2007-2009)
Principal investigator: Ferran Azorín

Caracterización biológica de inhibidores de metil transferasas
Spanish Ministry of Science and Innovation, PET2007-0319-02 (2009-2011)
Principal investigator: Ferran Azorín

Cromatina silenciada: Análisis de los factores y mecanismos implicados en su formación y mantenimiento
Spanish Ministry of Science and Innovation, BFU2006-01627/BMC (2006-2009)
Principal investigator: Ferran Azorín

Epigenética: Mecanismos y enfermedad
Spanish Ministry of Science and Innovation, CSD2006-49 (2006-2010)
Principal investigator: Ferran Azorín

Regulación transcripcional de genes controlados por el factor GAGA: Identificación de nuevos genes diana y de los mecanismos de activación/represión que operan in vivo
Spanish Ministry of Science and Innovation, BFU2007-64395/BMC (2007-2010)
Researcher: Jordi Bernués

Collaborations

Analysis of the contribution of multi-KH-domain proteins to RNA editing and heterochromatin organisation
Sergio Pimpinelli, University of Rome (Rome, Italy)

ChIP-seq analyses of histone modifications and chromatin binding proteins in Drosophila
Herbert Auer and David Rossell, IRB Barcelona (Barcelona, Spain)