

Chromatin structure and function



Genomic functions take place in chromatin, not in naked DNA. Over the last few years, we have become increasingly aware of the important contribution of chromatin to the regulation of genomic functions. Changes in chromatin structure have been found to 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. Furthermore, increasing evidence indicates that alterations in chromatin structure and function are at the root of many 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/CENP-A, H3.3, H2A.Z, macroH2A, etc.), and non-coding RNAs that modify chromatin structure and regulate gene expression. Our research focuses on the molecular basis of chromatin function and its regulation. More precisely, we seek to elucidate the contribution of chromatin to the regulation of the following: (i) centromere identity and function, (ii) gene expression, and (iii) long-distance genomic interactions.

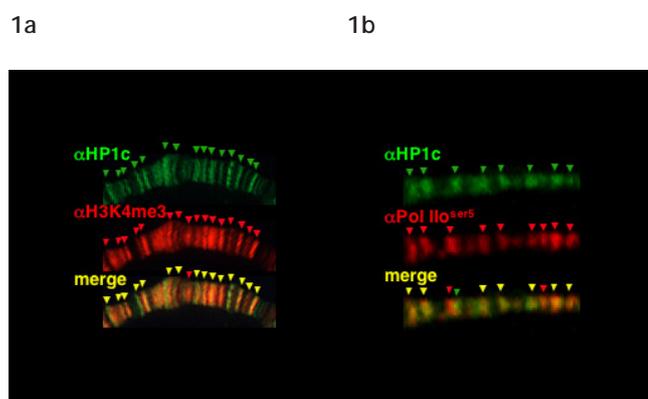


Figure 1. *Drosophila* HP1c co-localises with H3K4me3 (a) and the poised form of RNA polymerase II, Pol II^{o^{ser5}} (b).

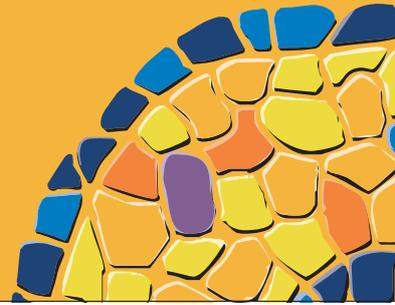
The contribution of HP1 proteins to the regulation of gene expression

The contribution of chromatin to the regulation of gene expression is well established. 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. Heterochromatin-protein 1 (HP1) is one of the best studied examples, where a regulatory non-histone protein is recruited to chromatin through the recognition of a specific histone modification, di- or tri-methylation of lysine 9 on the histone H3 tail (H3K9me_{2,3}). This interaction, which involves the N-terminal chromo-domain of HP1, is known to play a fundamental role in the formation and maintenance of heterochromatic domains.

With the exception of budding yeast, HP1 is widely conserved in eukaryotes, with most species having several isoforms. HP1 proteins are characterised by a common structural organisation consisting of two conserved domains, the N-terminal chromo-

RESEARCH GROUP MEMBERS

Principal Investigator Ferran Azorín **Research Associates** Jordi Bernués, Maria Lluïsa Espinàs **Postdoctoral Fellows** Martí Badal, Francesc Xavier Marsellach, Olga Moreno, Mònica Torras **PhD Students** Lorena Aguilar, Xavier Aran, Marta Batlle, Marta Blanch, Sergi Cuartero, Joan Font, Marta Lloret, Sonia Medina, David Piñeyro, Olivera Vujatovic **Research Assistants** Carles Bonet, Esther Fuentes, Gemma Molla, Alicia Vera **Visiting Students** Tomasso Cabaza (Italy), Nina Schuback (Germany), Pablo Sirkin (Argentina)



domain and the C-terminal chromo-shadow domain, which are spaced by a variable non-conserved hinge-domain. The existence of multiple isoforms suggests functional specialisation, with distinct isoforms exerting different functions. For instance, in *Drosophila*, three of the five HP1 isoforms (HP1a, HP1b and HP1c) are ubiquitously expressed, while the other two (HP1d/Rhino and HP1e) are predominantly expressed in the germ-line. Moreover, ubiquitously expressed HP1 isoforms show differential chromosomal distributions, as HP1a is mainly associated with heterochromatin, while HP1c is excluded from centromeric heterochromatin and HP1b is found both in euchromatic and heterochromatic domains. A similar situation is observed in mammals, where the localisation patterns of the three HP1 isoforms (HP1 α , HP1 β and HP1 γ) overlap only partially and show differential dynamics during differentiation and cell cycle progression.

The molecular mechanisms that determine the distribution of the various HP1 isoforms and their differential functional properties remain largely unknown. Most of our knowledge about the mechanisms of action of HP1 proteins derives from studies addressing the functional properties of *Drosophila* HP1a or mammalian HP1 α . From these studies, a general picture emerges by which, through the chromo-domain, HP1 proteins bind chromatin regions enriched in H3K9me_{2,3}, while through the chromo-shadow domain they recruit different factors, thereby resulting in various functional outcomes, namely heterochromatin assembly and gene silencing. It is uncertain whether this general picture applies to all HP1 proteins and scenarios. In this context, we have reported the functional characterisation of HP1c, a *Drosophila* HP1 protein of largely unknown properties (Font-Burgada *et al*, 2008). Our results show that HP1c extensively co-localises with poised RNA polymerase II and H3K4me₃, a modification that correlates with active chromatin

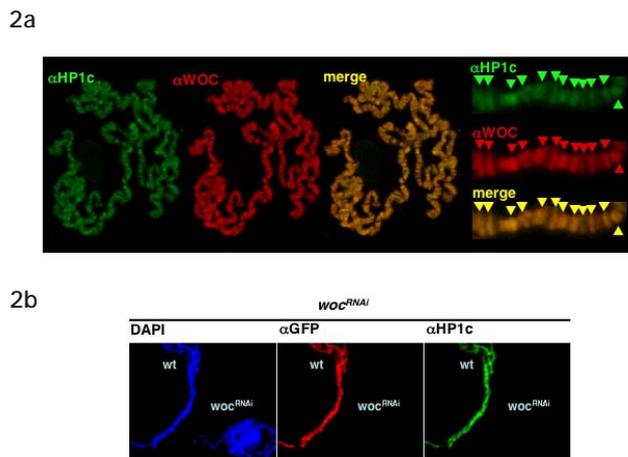


Figure 2. *Drosophila* HP1c interacts with the zinc-finger proteins WOC and ROW. (a) HP1c extensively co-localises with WOC (a similar co-localisation is observed with ROW). (b) Binding of HP1c to chromatin is abolished in the absence of ROW (*row^{RNAi}* chromosomes; similar results are obtained in *woc^{RNAi}* chromosomes).

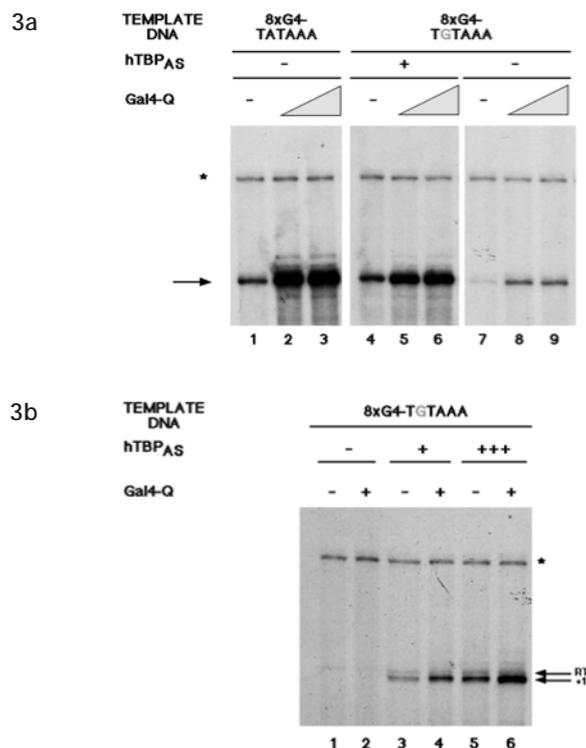


Figure 3. *In vitro* transcription experiments showing that the GAGA Q domain, when fused to GAL4, activates transcription despite the presence of a non-functional TATA-box (TGTAAG). When complemented with an altered specificity, TBP (*hTBP_{AS}*) basal transcription is restored and activation is possible even in a heat-inactivated nuclear extract.

domains, thereby indicating that HP1c contributes to the regulation of gene expression (Figure 1). In fact, HP1c forms a distinct multi-protein complex with two zinc-finger proteins, WOC (without children) and Relative-of-WOC (ROW), both putative transcription factors. HP1c efficiently binds H3K9me_{2,3} *in vitro*, but its binding to chromatin depends strictly on both WOC and ROW (Figure 2). Moreover, expression profiling indicates that HP1c, WOC and ROW extensively co-operate to regulate gene expression, especially in the context of the nervous system. From this study, which unveils the essential contribution of DNA-binding proteins to HP1c functionality and recruitment, HP1 proteins emerge as an increasingly diverse family of chromatin regulators.

The contribution of the *Drosophila* GAGA factor to the regulation of gene expression

GAGA is a *Drosophila* transcription factor that is involved in many nuclear activities. In transcription experiments performed *in vitro*, we have obtained evidence that GAGA factor enhances transcription by stabilising pre-initiation complex (PIC) and by promoting reinitiation (Vaquero *et al*, 2008). Formation of PIC prior to GAGA addition prevents activation, thereby suggesting that GAGA is required early in the formation of activated complexes. GAGA stimulation of transcription can be attributed in part to the stabilisation of PIC and can occur *in vitro* and *in vivo* even in the absence of a functional TATA box. This finding suggests that GAGA recruits and/or nucleates the general transcription factors (in particular TBP and/or TFIID) to assemble PIC in a similar way as the well-known Gal4VP16 activator acts. All these properties depend on the GAGA C-terminal glutamine-rich domain and, in addition to other roles and previous data, support a role of GAGA as a transcription factor (Figure 3).

GAGA overexpression and depletion experiments performed genome-wide in *Drosophila* S2 cells have shown that GAGA is essentially an activator of transcription (Bernués *et al*, 2007). Despite a relatively short list of genes whose expression is affected by GAGA dosage in these experiments, a large number of uncharacterised genes have been shown to be under its control (~50% of them). Among the known targets, some show a high score and are likely to be direct GAGA targets. While they are currently under assay, the first studied in more detail is a pro-apoptotic gene (*sickle*), which we have shown to be responsible for the high lethality observed in two experimental conditions in flies. Moreover, we have demonstrated this lethality to be fully rescued by co-expression of the anti-apoptotic DIAP1. Remarkably, lethality in other conditions cannot be recovered in this way, thereby strongly suggesting that the effects of GAGA are context-dependent. Gene ontology analysis of our results revealed that GAGA is most highly relevant in regulating the expression of genes involved in cell adhesion and development in S2 cells. These findings have been obtained in a cellular system and do not represent the complete picture of GAGA action. Therefore, and also because of the highly significant effect of GAGA on genes involved in development, we are currently performing similar experiments in transgenic flies (Figure 4).

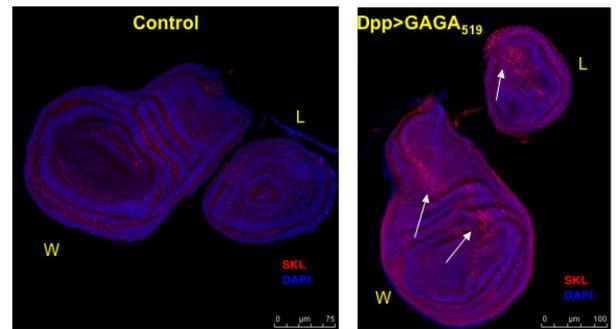
Long-distance genomic interactions: the contribution of CENP-B to the regulation of mating-type switching in fission yeast

In the fission yeast *Schizosaccharomyces pombe*, haploid cells switch mating type by means of a tightly regulated gene conversion event that involves long-distance interactions between an expressed locus (*mat1*) with either of two silent donor loci (*mat2* and *mat3*), which are located 17 kb and 29 kb away from *mat1*, respectively. Mating-type information is contained in the silent *mat2-P(plus)* and *mat3-M(minus)* loci but is expressed only after translocation to the *mat1* locus, thereby giving rise to *mat1P* or *mat1M* cells, depending on whether *mat2* or *mat3* information is expressed at *mat1*. Silencing at the *mat2* and *mat3* loci is mediated by heterochromatin, which, in the mating-type region, extends for a 20-kb long domain.

Mating-type switching initiates during DNA replication with the introduction of a strand-specific single-strand (SSB) imprint at *mat1*, which, in the next round of DNA replication, is converted into a double-strand break (DSB). This DSB is then healed by gene conversion using *mat2* or *mat3* as donors. Donor selection is, however, not random. On the contrary, *mat1P* cells preferentially use *mat3* as donor while *mat1M* cells use *mat2*. Directionality of switching therefore ensures that cells switch to the opposite mating-type with a very high frequency. Directionality of switching is determined by the cell-type specific distribution of the Swi2/Swi5-complex, which promotes recombination. In *mat1P* cells, Swi2/Swi5 localisation is restricted to a recombination-enhancer (SRE) located adjacently to *mat3* so that under these circumstances only *mat3* is efficiently used as donor. In contrast, in *mat1M* cells, Swi2/Swi5 spreads across the entire mating-type region, reaching the *mat2* locus, which becomes the preferred donor site as a result of the structural constraints imposed by heterochromatin. The spread of Swi2/Swi5 in *mat1M* cells relies on heterochromatin, as it is abolished by mutations that affect heterochromatin formation. Consequently, in the absence of heterochromatin, *mat3* is used as a donor at a much higher frequency than *mat2* (Figure 5).

However, the mechanisms that regulate the spread of Swi2/Swi5 across heterochromatin are not fully understood. In this context, we have reported on the identification of the first factor required for the heterochromatin-mediated spread of Swi2/Swi5, Abp1 (Aguilar-Arnal *et al*, 2008). We have shown that Abp1 binds at the mating-type locus and regulates directionality of switching. In *abp1Δ* cells, *mat3* is preferentially used as donor as occurs when heterochromatin is disrupted. The deletion of *abp1*, however, does not affect heterochromatin organisation at the mating-type region. This observation indicates that Abp1 is required for the efficient spread of Swi2/Swi5 across heterochromatin. ChIP-experiments have confirmed this hypothesis.

Abp1 is a member of the CENP-B family of proteins, which were first identified in humans as centromere-associated proteins. CENP-Bs are conserved sequence-specific DNA-binding proteins that derive from pogo-like transposases through independent domestication events. In fact, in *S. pombe*, Cbh1 and Cbh2 are also homologous to CENP-B and, together with Abp1, play redundant functions in the regulation of various nu-



W: wing imaginal disk; L: leg imaginal disk

Figure 4. *DppGal4*-directed overexpression of *GAGA₅₁₉* induces sickle (*SkI*) misexpression in wing and leg disks that later results in loss of legs and strong lethality. Arrows indicate *SkI* expression.

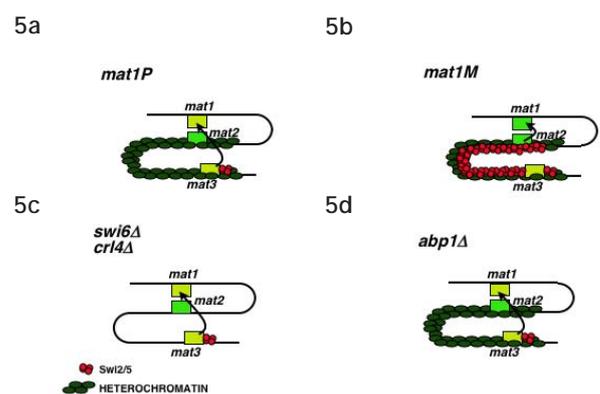


Figure 5. *Abp1* is required for heterochromatin-dependent spread of the Swi2/Swi5 complex. In *mat1P* cells, the Swi2/Swi5-complex localises to *mat3* (a) and only in *mat1M* cells does it spread to *mat2* (b), thereby allowing its use as donor during switching. The spread of Swi2/Swi5 to *mat2* is mediated by heterochromatin and is abolished by mutations in heterochromatin assembly factors (ie, *swi6Δ*, *cri4Δ*) (c), which prevent the use of *mat2* as donor. In *abp1Δ* cells, heterochromatin organisation of the mating-type locus is preserved but *mat2* is not efficiently used as donor, thereby indicating that the spread of Swi2/Swi5 to *mat2* is abolished (d).

clear processes. However, the contribution to the regulation of directionality of mating-type switching is specific to Abp1, as deletion of *cbh1* or *cbh2* shows no effects on this process (Aguilar-Arnal *et al*, 2008). Taken together, these observations unveil the diversity of functions regulated by CENP-B proteins, which emerge as general regulators of chromatin structure and function.

Publications

Aguilar-Arnal L, Marsellach F-X and Azorín F. The fission yeast homologue of CENP-B, Abp1, regulates directionality of mating-type switching. *EMBO J*, 27(7), 1029-38 (2008)

Font-Burgada J, Rossell D, Auer H and Azorín F. *Drosophila* HP1c isoform interacts with the zinc-finger proteins WOC and relative-of-WOC 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)

Pérez-Lluch S, Cuartero S, Azorín F and Espinàs ML. Characterisation of new regulatory elements within the *Drosophila* bithorax complex. *Nucleic Acids Res*, 36(21), 6926-33 (2008)

Vaquero A, Blanch M, Espinàs ML and Bernués J. Activation properties of GAGA transcription factor. *Biochim Biophys Acta*, 1779(5), 312-17 (2008)

Research networks and grants

Anàlisi estructural i funcional de la cromatina
Agency for Administration of University and Research Grants (AGAUR), 2005-SGR-00678 (2005-2008)
Principal investigator: Ferran Azorín

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 (2008-2010)
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 (2007-2009)
Principal investigator: Ferran Azorín

Characterisation of the role of histone H1 and its post-translational modifications in the functional regulation of chromatin
European Commission, Marie Curie Action, International Reintegration Grant (2007-2009)
Principal investigator: Ferran Azorín

Epigenética: Mecanismos y enfermedad, Ingenio 2010
Spanish Ministry of Science and Innovation, CSD2006-49 (2006-2011)
Principal investigator: Ferran Azorín

Episomal vectors as gene delivery systems for therapeutic application
European Commission, FP6-2003-LSH-2 (2005-2008)
Principal investigator: Ferran Azorín

Episomal vectors as gene delivery systems for therapeutic application (EPI VECTOR)
European Commission, LSHB-CT-2004-511965 (2005-2008)
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)
Principal investigator: Ferran Azorín

Collaborations

Analysis of the contribution of chromatin to the stability of episomal vectors
Hans J Lipps, University of Witten (Witten, Germany)

Analysis of the contribution of HP1c, WOC and ROW to the regulation of gene expression in Drosophila
Herbert Auer and David Rossell, IRB Barcelona (Barcelona, Spain)

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