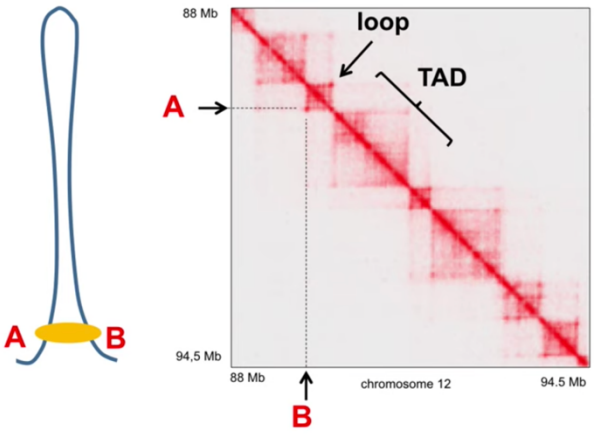
How do Cohesin and CTCF Fold DNA in Mammalian Genomes?

Notes from this lecture: https://www.imp.ac.at/news/detail/article/ibiology-talk-on-role-of-cohesin-and-ctcf-in-folding-dna/

If all the DNA contained in a single human nucleus was to be stretched out linearly, then it's length would be 2 metres. Or alternatively, if the size of a human nucleus was equal to a football, then 50 miles of DNA would be packed inside it. This means that DNA must be extremely thin and highly organised within each nucleus in each cell of the human body.



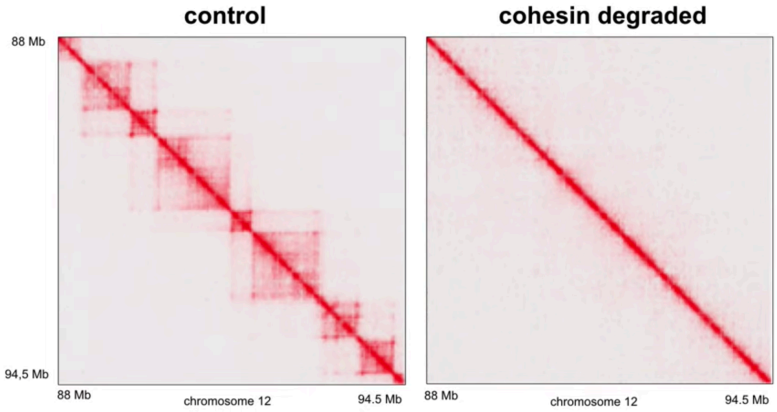
Genomic DNA is organised at different levels. At the most basic level, DNA is wrapped around 8 proteins, called histones, to form nucleosomes. These nucleosomes are coiled to form a chromatin fibre, and these are further condensed. At the highest level, genomic DNA consists of euchromatin (mostly active "A" compartments) and heterochromatin (mostly inactive "B" compartments).

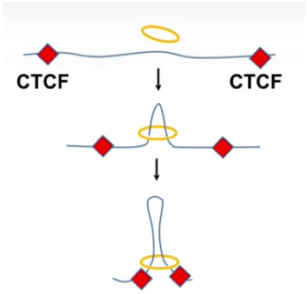
DNA is known to form loop-like structures, some of which are believed to have biological purpose, such as regulating gene regulation by bringing promoters in physical contact with distal enhancers, whilst some are thought to be purely structural. These loops are formed by cohesion complexes, and a protein complex called CTCF helps with sequence specificity, so that loops form between the correct bases (Wendt et al., 2008).

The genes required for cohesion were first identified in yeast. Seven genes were identified, four of which encode the four subunits of the cohesion complex. The cohesion protein complex is a member of the "structural maintenance of chromosomes" (SMC) family of complexes (Losado and Hirano, 2005) and looks similar to a ring that is "threaded" onto the DNA (Peters and Nishiyama, 2012). Moreover, if any of the 4 subunits of cohesion are cleaved (e.g. by WAPL), then the complex disassociates from the DNA. WAPL was therefore used extensively to explore the mechanisms of cohesion.

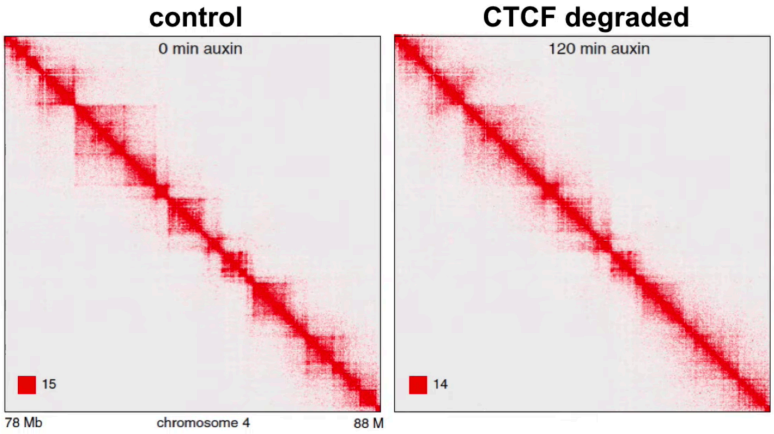
It was originally shown that cohesion is required for metaphase in mitosis, where sister chromatids align at the centre of the cell during DNA replication. If cohesion is removed from mitotic chromosomes by WAPL (a protein the binds to and cleaves the cohesion complex) and separase, then the sister chromatids do not align or pair up correctly (Huid in't Veld et al., 2014). However, observations including that cohesion is loaded onto DNA before it is needed for cohesion of sister chromatids (it is loaded in telophase) (Peters and Nishiyama, 2012), that much more cohesion is loaded than is required for this function (Hauf S, 2001) and that cohesion is also present in cells that do not need cohesion (e.g. mouse neuron cells that do not divide again) (Wendt et al., 2008) implied that cohesion may have additional roles.

Immunoprecipitation experiments showed that cohesion and CTCF are enriched at discrete sites on chromosome arms and it is known that CTCF functions in allele specific DNA loop formation (Wendt et al., 2008). It was therefore hypothesised that cohesion has an additional role in the cohesion of DNA at the base of a loop, for example to bring promoter sequences in contact with distal enhancer elements. Experiments which degraded cohesion showed that indeed cohesion is required for loops and TADs (Wutz et al., 2017).

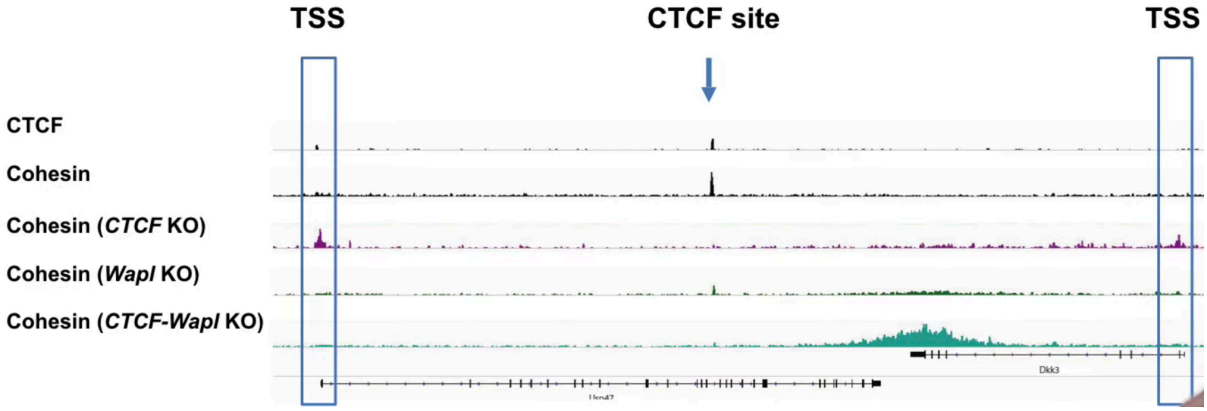




Loops form by a process called extrusion, whereby cohesion complexes bind to the DNA and the DNA is pulled through the complex in opposite directions to form a loop (Sanborn et al., 2015). The loop extrusion model asserts that the DNA is pulled through the complex until the DNA locates a CTCF motif that is pointing inward to the loop (meaning that CTCF motifs facing the "wrong way" will not be recognised). A nice animation of this is available here: https://www.youtube.com/watch?v=Tn5qgEqWgW8. Indeed, when CTCF is depleted, the chromatin loop boundaries are fuzzier, but interestingly, chromatin interactions are not abolished in the absence of CTCF.



CTCF also has a role in the positioning of cohesion in the genome (Busslinger et al., 2017). If it is knocked out, then cohesion accumulates at TSS. Moreover, if *CTCF* and *Wapl* is depleted, then cohesion accumulates in very large regions (not looking like a peak) at the 3' end of active genes that are converging to each other (one is read left to right and one is read right to left). These results imply that cohesion is mobile within the genome and the CTCF is not essential for cohesion-mediated chromatin interactions.



Summary

- Structural loops form on the basis of convergently oriented CTCF-bound motifs that are stabilised by cohesion rings (~ few hundred kbp)
- Multi-loop domains characterised by higher intradomain interactions (than interdomain interactions), known as TADs (~1 Mbp)
- Chromosomal territories are partitioned into A- and B- compartments (multi-Mbp)

- Loops connect two anchor points which are far apart on the chromosome and form by extrusion.

- CTCF binds to the DNA at a non-palindrome sequence and is recognised by the cohesion complex that is extending the loop. The convergence rule shows that the two CTCF sequences must be facing each other along the DNA.

https://www.ncbi.nlm.nih.gov/pubmed/18235444

https://www.ncbi.nlm.nih.gov/pubmed/15937217

https://www.ncbi.nlm.nih.gov/pubmed/23043155

https://science.sciencemag.org/content/346/6212/968/tab-pdf

https://www.ncbi.nlm.nih.gov/pubmed/11509732

https://www.embopress.org/doi/full/10.15252/embj.201798004

Nice video: https://www.youtube.com/watch?v=Tn5qgEqWgW8

https://www.pnas.org/content/early/2015/10/22/1518552112

https://www.ncbi.nlm.nih.gov/pubmed/28424523