

Oncometabolite tinkers with genome folding, boosting oncogene expression

Elizabeth Ing-Simmons and Matthias Merkschlager

MRC Clinical Sciences Centre, Imperial College London, Du Cane Road, London W12 0NN, UK

Address correspondence to EI-S (liz.ing-simmons@csc.mrc.ac.uk) and MM (matthias.merkschlager@csc.mrc.ac.uk)

Abstract. A recent paper makes a compelling case for a new mechanism by which heterozygous mutations in isocitrate dehydrogenases (IDH1/2) undermine gene regulatory mechanisms: 2-hydroxyglutarate (2HG) produced by mutant IDH alters the binding of the chromosomal organiser protein CTCF, disrupting the spatial and regulatory organisation of the genome.

IDH mutations occur with high frequency in gliomas and other human cancers. Mutant IDH enzymes produce the oncometabolite 2HG, which accumulates to high concentrations. 2HG not only changes the metabolic state of IDH mutant cells, but - crucially - inhibits the activity of enzymes that are critical for chromatin structure and gene regulation. Tet enzymes remove DNA methylation [1] and their inhibition by 2HG increases DNA methylation in IDH mutant cells. 2HG also inhibits the Jumanji family of histone demethylases, increasing histone methylation. These and other effects of 2HG [1] perturb gene expression in IDH mutant cells and contribute to tumor formation [1].

A recent paper [2] adds a new twist to the story by demonstrating that IDH mutations affect the binding of the chromatin organiser protein CTCF, which demarcates domains of genome folding [3]. To illustrate the relationship between 3D chromatin domains and CTCF binding, Figure 1 shows high-resolution Hi-C data (Figure 1A) and CTCF ChIP-sequencing data (Figure 1B) for the genomic region home to the known glioma oncogene, *PDGFRA*. The *PDGFRA* promoter is adjacent to CTCF sites in reverse orientation (green in Figure 1C). Illustrating an emerging principle of genome folding [3],

these reverse sites interact with CTCF motifs in convergent (forward) orientation to form a 'loop domain', visible in Hi-C data as a triangle connecting the CTCF sites (Figure 1A).

IDH mutant glioma cells lose ~600 CTCF binding events genome-wide and gain ~300, presumably perturbing 3D genome organisation. Loss of CTCF binding correlates with increased DNA methylation, and preferentially occurs at CTCF binding sites with motifs that contain CG dinucleotides as targets of DNA methylation [2]. DNA methylation may not be the only cause of altered CTCF binding in IDH mutant cells: At some sites at least, the reported increase in DNA methylation is smaller than the observed drop in CTCF binding [2], and increased DNA methylation does not explain the new CTCF binding sites that appear in IDH mutant cells. In IDH wild type cells, CTCF binding is sensitive to DNA methylation only at a subset of sites, presumably because CTCF co-localises with Tet proteins that remove DNA methylation [e.g.4,5]. Tet inhibition in IDH mutant cells may increase the sensitivity of CTCF to DNA methylation. Changes in histone methylation and other chromatin properties may also contribute to altered CTCF binding [1,6].

In normal cells, genes in the same domain show more similar expression than genes located in separate domains, in part because they share access to gene regulatory elements [7]. In this way, spatial segmentation of the genome facilitates the regulation of gene expression. Does altered CTCF binding perturb the 3D organisation of the genome and deregulate gene expression in IDH mutant cells? In support of this idea, the authors demonstrate that genes located in separate domains show more similar expression in IDH mutant cells than in IDH wild type cells [2]. Until high-resolution Hi-C data for IDH mutant cells become available, the precise relationship between altered CTCF binding, genome folding, and gene regulation in IDH mutant cells remains speculative. Nevertheless, the available data show that in IDH mutant cells the *PDGFRA* promoter interacts with an enhancer near the *FIPL1L* gene, almost a million base pairs away (Figure1), and that *PDGFRA* expression is upregulated [2].

The drug 5-azacytidine partially restores CTCF binding and downregulates *PDGFRA* expression in IDH mutant cells [3]. Mechanistically, 5-azacytidine reduces DNA methylation by inhibiting the methyl transferase DNMT1, but has additional effects on chromatin and gene regulation [8]. To further test the hypothesis that increased

PDGFRA expression in IDH mutant cells is due to reduced CTCF binding, the authors mutate a CTCF motif ('x' in Figure 1A) at the base of the CTCF loop domain adjacent to the *PDGFRA* promoter (Figure 1). Although the deletion of a single motif is distinct from the redistribution of CTCF binding in IDH mutant cells (Figure 1), it also results in increased interactions of the *PDGFRA* promoter with the *FIPL1L1* enhancer and elevated *PDGFRA* expression. This experiment links CTCF binding to long-range interactions and the deregulated expression of a known oncogene in IDH mutant glioma [3].

Closer inspection of the Hi-C data (Figure 1) suggests to us an explanation for the sensitivity of *PDGFRA* to perturbations in CTCF binding. The CTCF sites at the *PDGFRA* promoter interact not just with the convergent CTCF sites at the left boundary of the domain, but in addition with other CTCF sites in the region. Such interactions may initiate far upstream (far left of *PDGFRA*), where most CTCF sites are in a forward orientation. In the absence of strong CTCF sites in reverse orientation, we imagine that loop formation continues towards *PDGFRA*. It may either be aborted at the forward CTCF site ('insulator', Figure 1C) or engage the reverse CTCF site near the *PDGFRA* promoter [9,10] and connect the *PDGFRA* promoter with CTCF sites in the vicinity of the *FIPL1L1* enhancer. Such interactions are visible as diagonal lines spanning several contact domains in the Hi-C heat map, and also in the interaction track below (Figure 1A). The *FIPL1L1* enhancer itself is bound by CTCF, but does not appear to contact the *PDGFRA* promoter in IDH wild type cells (Figure 1A). Since loops that form in opposite directions are thought to compete with each other [10], weakened CTCF binding in IDH mutant cells, or the CRISPR-mediated deletion of a CTCF site at the left boundary of the loop domain next to *PDGFRA*, may increase contacts between the *PDGFRA* promoter and *FIPL1L* enhancer.

Clinically, high levels of 2HG may sensitise *IDH*-mutant tumour cells to other therapeutic approaches [1]. Alternatively, pharmacological inhibition of IDH, or selective inhibition of the abnormal enzymatic activity could be used to specifically target tumour cells [1]. IDH inhibitors can revert the impact of IDH mutations on chromatin state [6]. Whether this includes correction of CTCF binding, genome folding and gene expression is an interesting question that remains to be explored.

Acknowledgements. We thank Drs Kim Nasmyth and Elphege Nora for discussion. Supported by the Wellcome Trust (009276) and the Medical Research Council, UK.

Abbreviations. IDH: isocitrate dehydrogenase; 2HG: 2-hydroxyglutarate

References

- [1] Cairns RA, Mak TW. Oncogenic isocitrate dehydrogenase mutations: mechanisms, models, and clinical opportunities. *Cancer Discov.* 2013;3: 730-41.
- [2] Flavahan WA, et al. 2015. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* doi:10.1038/nature16490 23.Dec.2015
- [3] Rao SS, et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159:1665–80
- [4] Sun Z, Terragni J, Borgaro JG, Liu Y, Yu L, Guan S, Wang H, Sun D, Cheng X, Zhu Z, Pradhan S, Zheng Y. High-resolution enzymatic mapping of genomic 5-hydroxymethylcytosine in mouse embryonic stem cells. *Cell Rep.* 2013 Feb 21;3(2):567-76.
- [5] Teif VB, Beshnova DA, Vainshtein Y, Marth C, Mallm JP, Höfer T, Rippe K. Nucleosome repositioning links DNA (de)methylation and differential CTCF binding during stem cell development. *Genome Res.* 2014 Aug;24(8):1285-95.
- [6] Kernytsky A, Wang F, Hansen E, Schalm S, Straley K, Gliser C, Yang H, Travins J, Murray S, Dorsch M, Agresta S, Schenkein DP, Biller SA, Su SM, Liu W, Yen KE. IDH2 mutation-induced histone and DNA hypermethylation is progressively reversed by small-molecule inhibition. *Blood.* 2015;125:296-303.
- [7] Symmons O, et al. 2014. Functional and topological characteristics of mammalian regulatory domains. *Genome Research* 24:390-400
- [8] Komashko VM, Farnham PJ. 5-azacytidine treatment reorganizes genomic histone modification patterns. *Epigenetics* 2010; 5: 229–240
- [9] Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. Formation of chromosomal domains by loop extrusion. bioRxiv doi: <http://dx.doi.org/10.1101/024620>
- [10] Sanborn AL, et al. 2015. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci U S A.* 112:E6456-65 doi:10.1073/pnas.1518552112

Figure legend

Figure 1. Chromosome folding across the *PDGFRA* region: mechanisms and models.

A. Hi-C heat map of the 3D chromatin organisation around *PDGFRA*. Red indicates increased interactions. Triangular shapes are contact domains. Diagonal lines emanating from triangles indicate interactions that span several domains. The line below the heat map shows interactions from the viewpoint of the *PDGFRA* promoter. The genomic positions of the *PDGFRA* promoter, a putative insulator, and the *FIPL1L* enhancer are shown underneath.

B. CTCF binding data from ChIP-seq experiments. Peak height reflects occupancy. Glioma cells GSC4 and GBM2w are IDH wt, GSC119 and AA15m are IDH mutant.

C. The direction of CTCF motif sequence (<https://github.com/liz-is/ctcf-motif-imr90>) is indicated as forward (red) or reverse (green). Motifs with low CTCF occupancy are omitted for clarity. Note that CTCF binding marks domain boundaries. Strong interactions (horizontal lines) form between the forward CTCF sites at the insulator and the reverse CTCF sites at the *PDGFRA* promoter. A cartoon illustrates how current models envisage the formation of this domain by the extrusion of a chromatin loop through a ring-shaped complex, possibly cohesin [9,10]. Extrusion stops when it encounters occupied convergent CTCF sites. Interactions between forward CTCF sites near the *FIPL1L* enhancer with reverse CTCF sites at the *PDGFRA* promoter may form as illustrated by the cartoon below: Loop extrusion that starts between CTCF sites in the same orientation, it is thought to skip CTCF sites until a reverse site is reached [9,10]. Hi-C data suggests that loop extrusion skips the weak reverse site in pale green and continue towards the *PDGFRA* promoter.

