($N–CF_3$; Fig. 1). The authors cleverly overcome the reluctance of fluoride ions to take part in useful reactions by using a relatively reactive salt, silver fluoride, to promote bond formation between two reagents that will be familiar to organic chemists but are rarely combined. The first of these is a sulfur-containing compound called an isothiocyanate (Fig. 1) — more specifically, the compound that promotes the classical Edman degradation reaction, used in early methods for determining the amino-acid sequence of proteins. Researchers from the same group as Scattolin *et al.* previously used[3,4] silver ions to promote the replacement of sulfur atoms by fluorine. Silver fluoride has the same role in the present work: it rips out sulfur from the Edman reagent and replaces it with three fluorine atoms, forming an intermediate compound that contains a trifluoromethyl (CF_3) group.

The second reagent is bis(trichloromethyl) carbonate (Fig. 1), which is often used to make amides and amide-like derivatives[5]. The authors report that, when used in combination with silver fluoride, this carbonate traps the typically highly unreactive CF_3-containing intermediate to form a compound called a carbamoyl fluoride. It is this compound that forms the conceptual breakthrough of Scattolin and colleagues' work.

The carbamoyl fluoride already contains the carbon–nitrogen bond of the target amide, thereby sidestepping the usual difficulties associated with making fluorinated amides. Moreover, it is stable enough to be isolated, but reactive enough to act as a building block for the synthesis of a range of *N*-trifluoromethylamides — which Scattolin *et al.* prepare by reacting the carbamoyl fluoride with various magnesium-containing compounds known as Grignard reagents (Fig. 1). The authors show that chemical groups called ureas and carbamates can also be made in this final step, widening the applicability of the chemistry beyond fluorinated amides.

In this work, Scattolin and colleagues have impressively choreographed the reactions of several seemingly incompatible, highly reactive reagents, perfectly controlling which meets which, and when. The authors demonstrate that their method can be used to make not only simple *N*-trifluoromethylamide molecules, but also chemically sensitive ones such as those based on amino acids, drug scaffolds and the monomers used to make polymeric materials.

The reported reactions will enable medicinal chemists to prepare previously unavailable compounds for testing in drug-discovery programmes — some of these compounds might well display new biological activities. The addition of a methyl group to the nitrogen atom of an amide in drug candidates has long been used to alter the conformations adopted by those molecules[6], thereby altering their biological activities; the ability to make *N*-trifluoromethylamides will broaden the scope of that strategy.

There are, however, drawbacks to the new chemistry that will need to be addressed before it can be used on the industrial scales needed to manufacture a drug. The main issue is that the reactions require substantial quantities of the necessary reagents: every fluorine atom introduced into a molecule, plus two more that are expelled during the reaction, is accompanied by an atom of silver, meaning that five atoms of silver are needed for every molecule of product. This is acceptable for the initial stages of drug discovery (lead development and optimization), which typically involve only milligram to gram quantities of drug candidates. However, it would not be sustainable for synthesizing compounds on the kilogram scales needed for testing in clinical trials, let alone for industrial manufacturing processes, because of the cost and large amount of waste produced. Another breakthrough will be needed to find a way of synthesizing *N*-trifluoromethylamides sustainably on such large scales.

Looking beyond medicinal chemistry, Scattolin and co-workers' findings reveal that isothiocyanate groups can act as precursors of trifluoromethyl groups bonded to nitrogen atoms. Their reactions might find much wider use in chemical synthesis, where the ability to make trifluoromethyl-substituted nitrogen compounds could facilitate the design and synthesis of catalysts or materials with new properties. ■

Jonathan Clayden *is at the School of Chemistry, University of Bristol, Bristol BS8 1TS, UK.*
e-mail: j.clayden@bristol.ac.uk

1. Scattolin, T., Bouayad-Gervais, S. & Schoenebeck, F. *Nature* **573**, 102–107 (2019).
2. Purser, S., Moore, P. R., Swallow, S. & Gouverneur, V. *Chem. Soc. Rev.* **37**, 320–330 (2008).
3. Tyrra, W. *J. Fluor. Chem.* **109**, 189–194 (2001).
4. Scattolin, T., Deckers, K. & Schoenebeck, F. *Angew. Chem. Int. Edn* **56**, 221–224 (2017).
5. Eckert, H. & Forster, B. *Angew. Chem. Int. Edn* **26**, 894–895 (1987).
6. Yamaguchi, K. *et al. J. Am. Chem. Soc.* **113**, 5474–5475 (1991).

CANCER EPIGENETICS

A lid for the marker

Two reports examine how a protein complex that adds transcription-repressing marks to histone proteins is potently inhibited by the protein EZHIP during the maturation of sperm and egg cells and in a type of brain cancer.

SERGI ARANDA & LUCIANO DI CROCE

How can a single genome generate specialized cells at the correct time and place in multicellular organisms? Gene-expression programs that are specific to different cell types are regulated by the addition and removal of chemical groups, known as epigenetic marks, to and from chromatin — the compact protein–DNA complex in which the genome is organized. Two papers in *Nature Communications* (one by Ragazzini *et al.*[1] and the other by Jain *et al.*[2]), together with recent studies[3–5], show how a protein called EZHIP can regulate and disrupt the addition of a major epigenetic mark by the Polycomb repressive complex 2 (PRC2) during certain biological events and in a type of cancer.

The Polycomb protein complexes are key epigenetic modulators that act to repress gene expression[6]. Mammalian PRC2 contains a core complex consisting of several subunits, including either EZH1 or EZH2 as an enzyme subunit[7]. The core complex can interact with various accessory subunits that influence its enzymatic activity and direct it to specific regions of the genome.

The major substrate of PRC2 is a nucleosome: a complex of eight histone proteins, around which DNA is wrapped; nucleosomes are the basic building blocks of chromatin. PRC2 adds a trimethyl group (me3) to a specific amino-acid residue of the H3 histone (lysine 27; abbreviated as K27), generating an epigenetic mark termed H3K27me3 (Fig. 1a). This modification occurs initially at certain sites and promotes the passage of PRC2 across adjacent regions of the chromatin[8].

The distribution of the H3K27me3 mark on nucleosomes throughout the genome changes drastically during certain biological events, including in sperm and egg cells when they mature, and in fertilized egg cells[9]. However, the molecular mechanisms that regulate this redistribution, and its functional consequences, are unknown. EZHIP is expressed only in the testes, ovaries and embryo-surrounding tissues of placental mammals[10], and was previously shown to interact with PRC2 in cultured mouse stem cells[11,12]. Ragazzini and colleagues demonstrate that, in the testes and ovaries of mice, EZHIP binds to and reduces the enzymatic activity of PRC2.

The authors also show that mice lacking EZHIP can live and develop normally, and are apparently fertile, but that female mice that lack EZHIP show progressive infertility with age. This finding calls for a re-evaluation of genomic studies in women with infertility, to clarify whether certain *EZHIP* gene variants could increase the risk of this condition. EZHIP protein is found

exclusively in placental mammals, so one might speculate that its appearance could have contributed to the evolution of the placenta.

Other reports[2,4,5] have found that, when EZHIP is expressed, it associates with PRC2 and inhibits its function — putting a 'lid' on the PRC2 marker. EZHIP expression is inversely correlated with the level of the H3K27me3 mark. Indeed, Jain *et al.* show that reintroducing EZHIP into cultured human cells that normally do not express the protein inhibits PRC2 activity and markedly reduces H3K27me3 levels. Conversely, removing the *EZHIP* gene from cultured cells that normally express it increases levels of H3K27me3 (refs 1,2,4,5). These results indicate that there is a causal relationship, rather than just a correlation, between EZHIP expression and decreases in H3K27me3 and changes to the overall epigenetic landscape.

Dysregulation of PRC2 function can be highly detrimental. For example, a mutation in one of the H3-encoding genes that leads to lysine 27 being replaced by a methionine residue greatly reduces the effectiveness of the complex (Fig. 1b). The mutant histone, which is termed H3K27M, has been implicated in aggressive childhood brain cancer[13], including one such cancer called diffuse intrinsic pontine glioma (DIPG)[14] .

Comparative analyses[2,3] have shown that the EZHIP-encoding gene has changed rapidly over evolution and, as a result, the amino-acid sequence of the protein shows little similarity between different species. However, a stretch of 12 amino acids at one end of the protein is identical in many mammalian species. Notably, this sequence is highly similar to the amino-acid sequence that contains the methionine 27 residue in the DIPG-associated histone variant H3K27M[2,3,5]. It is therefore known as the K27M-like peptide (KLP) and might mimic the effect of H3K27M on PRC2 function.

Jain *et al.* determine the functionality of EZHIP and, particularly, the function of its KLP. They propose a model in which EZHIP forms a complex with PRC2 that is bound to H3K27me3, and, in a similar way to H3K27M, prevents PRC2 from depositing further H3K27me3 marks across the chromatin (Fig. 1c). Using a computerized model of the co-crystal structure of PRC2 bound to H3K27M, Jain *et al.* predict that the KLP of EZHIP makes stable contacts with the PRC2 enzyme subunit EZH2 similarly to how H3K27M binds to EZH2.

Moreover, Jain and colleagues show that EZHIP inhibition of PRC2 activity relies on the methionine 406 residue in KLP, and that the ability of EZHIP to inhibit PRC2 is enhanced when PRC2 is activated by H3K27me3 at a different site on the complex from where EZHIP binds. Mouse embryo fibroblast cells that were engineered to express H3K27M or EZHIP show similar reductions in chromatin H3K27me3 and similar changes in gene expression relative to genetically normal mouse embryo fibroblasts. However, Jain *et al.* also find that some 'resilient' parts of the chromatin in the EZHIP-expressing cells retain their H3K27me3 marks; the mechanisms that enable the retention of PRC2 function selectively at these sites are not clear.

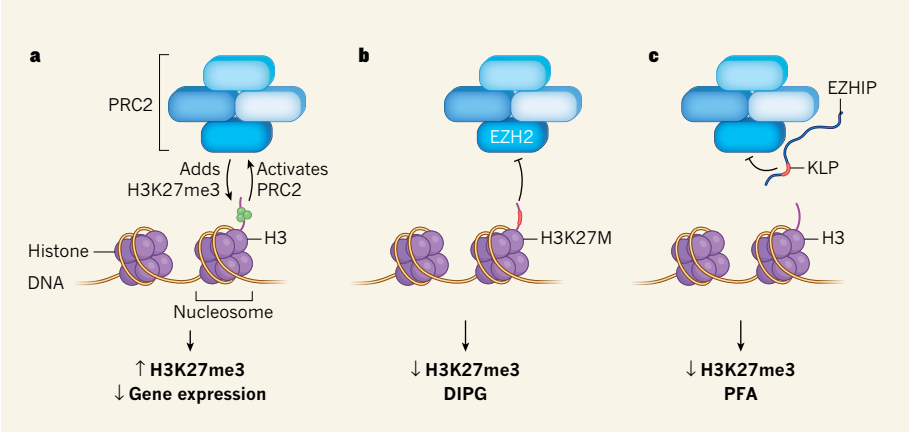

Figure 1 | The protein EZHIP inhibits the enzymatic activity of the Polycomb repressive complex 2 (PRC2). a, A nucleosome comprises DNA wrapped around a histone protein complex. PRC2 is a group of associated proteins that adds a trimethyl (me3) group to a lysine amino-acid residue (K27) in the histone protein H3, generating the H3K27me3 mark (green). H3K27me3 represses the expression of nearby genes, and further activates PRC2, leading to an increase in H3K27me3 throughout the genomic region. **b**, A mutant form of H3 known as H3K27M, in which K27 is substituted by a methionine residue, inhibits EZH2, the enzyme subunit of PRC2, leading to reductions in H3K27me3. H3K27M drives a form of brain cancer called diffuse intrinsic pontine glioma (DIPG). **c**, The protein EZHIP contains an amino-acid sequence that is highly similar to that containing the substituted methionine in H3K27M (the K27M-like peptide; KLP), and Ragazzini *et al.*[1] and Jain *et al.*[2] show that EZHIP's KLP inhibits EZH2. EZHIP expression is upregulated in a brain cancer called posterior fossa ependymoma group A (PFA)[2], but is also needed for germ-cell maturation and fertility in female mice[1].

Disrupted PRC2 function is a feature of another type of childhood brain cancer called posterior fossa ependymoma group A (PFA)[4]. However, unlike in DIPG (in which PRC2 function is reduced by H3K27M), PRC2 function in PFA is reduced by an aberrant upregulation of EZHIP[2,3]. Whether other proteins that have KLP activity might inhibit PRC2 in other human tumours, and whether other chromatin complexes besides PRC2 can similarly be inhibited by other proteins, are questions that should be explored.

Together, the results of the two papers increase our understanding not only of how PRC2 affects gene expression and cell function in different settings, but also of mammalian physiology and human disease. By advancing our understanding of EZHIP functionality, the studies lay the foundation for exploring the role of this protein in driving cancer, and indicate that targeting PRC2 or EZHIP could have therapeutic potential for children with PFA and DIPG. ■

Sergi Aranda *and* **Luciano Di Croce** *are at the Center for Genomic Regulation and the Barcelona Institute of Science and Technology, Barcelona 08003, Spain.* **L.D.C.** *is also at the Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, and the Pompeu Fabra University, Barcelona.*
e-mails: sergi.aranda@crg.eu; luciano.dicroce@crg.eu

1. Ragazzini, R. *Nature Commun.* https://doi.org/10.1038/s41467-019-11800-x (2019).
2. Jain, S. U. *et al. Nature Commun.* https://doi.org/10.1038/s41467-019-09981-6 (2019).
3. Hübner, J.-M. *et al. Neuro-Oncology* **21**, 878–889 (2019).
4. Pajtler, K. W. *et al. Acta Neuropathol.* **136**, 211–226 (2018).
5. Piunti, A. *et al. Sci. Adv.* **5**, eaax2887 (2019).
6. Schuettengruber, B., Bourbon, H. M., Di Croce, L. & Cavalli, G. *Cell* **171**, 34–57 (2017).
7. Aranda, S., Mas, G. & Di Croce, L. *Sci. Adv.* **1**, e1500737 (2015).
8. Yu, J.-R., Lee, C.-H., Oksuz, O., Stafford, J. M. & Reinberg, D. *Genes Dev.* **33**, 903–935 (2019).
9. Vaquerizas, J. M. & Torres-Padilla, M.-E. *Nature* **537**, 494–496 (2016).
10. Uhlén, M. *et al. Science* **347**, 1260419 (2015).
11. Kloet, S. L. *et al. Nature Struct. Mol. Biol.* **23**, 682–690 (2016).
12. Maier, V. K. *et al. Mol. Cell. Proteom.* **14**, 1435–1446 (2015).
13. Schwartzentruber, J. *et al. Nature* **482**, 226–231 (2012).
14. St. Jude Children's Research Hospital–Washington University Pediatric Cancer Genome Project. *Nature Genet.* **44**, 251–253 (2012).
15. Mohammad, F. *et al. Nature Med.* **23**, 483–492 (2017).
16. Piunti, A. *et al. Nature Med.* **23**, 493–500 (2017).

This article was published online on 26 August 2019.