Chromatin modifying activity of leukaemia associated fusion proteins

Luciano Di Croce*

ICREA and Centre de Regulació Genòmica (CRG), Passeig Marítim 37-49, 08003 Barcelona, Spain

Received January 12, 2005; Revised and Accepted February 24, 2005

The leukaemias, which are divided into chronic and acute forms, are malignant diseases of haematopoietic cells in which the proper balance between proliferation, differentiation and apoptosis is no longer operative. Genes, such as those of mixed-lineage leukaemia, AML1 and retinoic acid receptor alpha, have been found to be aberrantly fused to different partners, which often encode transcription factors or other chromatin modifying enzymes, in numerous types of acute lymphoid and myeloid leukaemias. These chimeric fusion oncoproteins, generated by reciprocal chromosomal translocations, are responsible for chromatin alterations on target genes whose expression is critical to stem cell development or lineage specification in haematopoiesis. Alterations in the ‘histone code’ or in the DNA methylation content occur as consequence of aberrant targeting of the corresponding enzymatic activities. Here, the author will review the most recent progress in the field, focusing on how fusion proteins generated by chromosomal translocation are responsible for chromatin alterations, gene deregulation and haematopoietic differentiation block and their implications for clinical treatment.

INTRODUCTION

The packaging of DNA into chromatin fibres provides the cell with the means to compact and to store its 2–3 billion base pairs of DNA, but it also regulates the accessibility of transcription factors and general transcription machinery to gene promoters (1). The chromatin fibre is composed of repetitive units, known as nucleosomes, which are comprised of an octamer of core histones (two of each H2A, H2B, H3 and H4), around which ~146 bp of DNA are wrapped. In addition, higher eukaryotes possess linker histones (e.g. H1), which are thought to bind to the exit and entry points of DNA as it winds around the nucleosome, thereby facilitating higher-order packaging of chromatin (2,3). Within chromatin, the N-terminal of core histones protrude out of the core structure and can contact adjacent nucleosomes. These ‘tail’ interactions are crucial for the regulation of chromatin structure and are subjected to multiple post-translational modifications (Fig. 1). These modifications form the basis of the ‘histone code’ hypothesis, where combination of modifications forms the recognition surfaces for chromatin regulatory factors (4).

Lysine acetylation

Acetylation of lysine residues within the histone tails is associated with a relaxed chromatin configuration, which is believed to decrease the histone–DNA interaction and facilitate transcription factor access to DNA (5). Consistent with this, many transcriptional co-activators possess histone acetyltransferase (HAT) activity. Conversely, many transcriptional co-repressors are enzymes with histone deacetylase (HDAC) activity, which contribute to maintain chromatin in the condensed state (6).

Lysine and arginine methylation

Histone lysines can be mono-, di- and trimethylated by a group of enzymes called histone methyltransferases (HMTase) (7). Members of this enzyme family contain a conserved SET domain, which is flanked by cysteine-rich regions. Methylated lysines are recognized by the conserved chromodomain modules found in chromatin-associated proteins related to heterochromatin proteins HP1 and Pc (8,9). Although methylation of histone H3 lysine 9 (H3-K9), H3-K27 and H4-K20 functions as a repressive mark, not all lysine methylation appears to be a signal for repression of transcription. It has been recently demonstrated that methylation of histones H3-K4, H3-K36 and H3-K79 are associated with transcriptionally competent euchromatin. Similarly, histone arginine methylation correlates with gene activation (10). An enzyme with histone demethylase activity has been recently identified (11), suggesting that dynamic regulation of histone
methylation by both histone methylases and demethylases occurs, in a mechanism similar to acetylation/deacetylation.

Serine phosphorylation
Phosphorylation at histone 3 serine 10 (H3-S10), H3-S28 and H4-S1 has been documented. Phosphorylation at H3-S10 and H3-S28 coincides with the induction of immediate-early gene expression and with the onset of mitosis. Phosphorylation is reversed by the protein phosphatase 1 family. The kinases that can phosphorylate H3 are Aurora-B/Ipl1, PKA, Rsk-2 and Msk1, which tend to target Ser/Thr sites that are surrounded by basic residues (12). Phosphorylation of a site adjacent to (or nearby) a methyl mark that engages an effector module could lead to consecutive loss of binding to that factor, a mechanism that has been proposed as ‘methyl/phos switching’ (13).

DNA methylation
Modification of the DNA itself can likewise lead to remodelling of chromatin and gene regulation. DNA methyltransferase (DNMT) enzymes catalyze the addition of a methyl group to cytosine residues at CpG dinucleotides, which if located within a gene’s regulatory regions can lead to transcriptional silencing (14). The process of DNA methylation in mammals is carried out by at least three catalytically active DNMT enzymes (15).

DNA methylation represses gene transcription by creating docking site for methyl-CpG binding domain (MBD) proteins, which selectively recognize methylated CpG dinucleotides (16). The presence of MBD proteins within methylated promoters could prevent gene activation by precluding binding of positive transcription factors. Recent data demonstrated that in addition to this ‘passive’ mechanism, MBD proteins act also by recruiting other repressive enzymes, such as HDACs and HMTs, to hypermethylated promoters (17). Aberrant methylation underlies susceptibilities to several forms of cancer and is likely to be involved in numerous other human diseases (18,19).

BLOOD CELL DIFFERENTIATION
All blood cell types are derived from haematopoietic stem cells (HSCs) (20). Two properties define these cells. First, they can generate more HSCs (self-renewal), and second, they have the potential to differentiate in a stepwise process of binary decisions into various progenitor cells (common lymphoid and myeloid progenitors), which in turn differentiate into additional intermediate progenitors. Self-renewal is required for HSCs to maintain haematopoiesis over the lifetime of the host. Differentiation defines the sequence of events by which cells undergo orderly changes into mature cells, which is regulated by several lineage-restricted transcription factors (Fig. 2). Mis-expression, disruption or, more often, chromosomal translocation of one of these ‘master’ genes required for haematopoiesis is usually a key event in leukaemias (21). In leukaemia, chimeric proteins created through chromosomal translocations are believed in several instances to function in a dominant-negative fashion on gene regulation, thus preventing proper cell differentiation. The fusion proteins exert their action via direct or indirect interactions with transcription co-factors or chromatin modifying enzymes. In contrast, tyrosine kinases, rather than transcription factors, are targeted by chromosomal translocation...
in chronic myeloid leukaemias (CML). CML are characterized by leukocytosis with normal differentiation and cell function. This suggests that alterations in tyrosine kinases provide mainly a proliferative and survival advantage, while chimeric transcription factors impair cellular differentiation (22). Why are solid tumours so rarely characterized by the presence of chimeric fusion proteins? The answer could be linked to the high frequency of recombination events that occurs in lymphocytes and to the requirement of a ‘permissive’ cellular context, which may be present in haematopoietic cells.

MIXED-LINEAGE LEUKAEMIA FUSION PROTEINS

Mixed-lineage leukaemia (MLL) protein (also known as ALL1 or HRX), the mammalian orthologue of the epigenetic transcriptional regulator trithorax in Drosophila, has been implicated in fusions with more than 60 other partners in both acute lymphoblastic leukaemia (ALL) and acute myeloid leukaemia (AML) (23). The MLL protein (Fig. 3) was shown to possess HMTase activity toward H3-K4 via its SET domain, leading to transcriptional activation at Hox gene promoters (24,25), the expression of which is high in progenitor cells and downregulated during differentiation and maturation. MLL, additionally, functions in a large super-complex of at least 29 proteins, which is involved in the remodelling, acetylation, deacetylation and methylation of nucleosomes and histones through association with SWI/SNF, TFIID, NuRD, MBDs, HDACs and HAT CBP (24). MLL fusion proteins are predicted to disrupt critical patterns of HOX gene expression in haematopoietic progenitor cells. Indeed, the Hoxa9 gene is upregulated in MLL-dependent leukaemia and together with Hoxa7 is required for MLL-ENV-dependent transformation (26). It has been recently shown that the CXXC domain of MLL, which is also found in MBD1 and DNMT1 proteins, is essential for myeloid transformation (27).

Because MLL fusion proteins lack the SET domain, the CBP-binding domain, and the PHD domain (involved in HDAC binding), the proper balance between activation and repression is likely compromised in MLL fusions. Fusion products of MLL with CBP t(11;16) or with p300 t(11;22) retain the bromodomain and HAT domain and might lead to leukaemia by increasing histone acetylation of genomic regions targeted by MLL (23). Another fusion partner of MLL is ENL, a subunit of the SWI/SNF complex. MLL–ENL fusion protein, t(11;19), associate and co-operate with SWI/SNF complexes to activate transcription of the Hoxa7 promoter (28). Moreover, partial tandem duplications of exons 2–6 or 2–8 are present in ~10% of AML.

Although, at present, it is unclear whether a loss or gain of MLL function is responsible for oncogenesis, these data suggest that MLL can activate genes at inappropriate times by mis-targeting enzymes involved in epigenetic decisions, consequently modifying the chromatin to allow gene activation (Fig. 4).

A similar scenario has been postulated for the fusion proteins generated by the t(5;11) translocation, NUP98–NSD1 (29), and by the t(8;11) translocation, NUP98–NSD3 (30), which are specifically associated with AML. The NUP98 gene encodes...
for the 98 kDa subunit of the nuclear pore complex and has been already identified as partner of recurrent translocations in leukaemia with several homeobox genes (e.g. HOXA9, HOXD13, PMX1). NSD1 and NSD3, which interact with nuclear receptors, are members of a family of proteins with a conserved architecture: multiple PHD fingers, a PWWP domain and a SAC/SET domain. The SAC/SET domain possesses HMTase activity towards H3-K36 and H4-K20, thus suggesting a potential involvement of these fusion proteins in chromatin remodelling and regulation of transcription.

MECHANISM OF REPRESSION OF AML1-FUSION ONCOPROTEINS

CBF is a heterodimeric transcription factor composed of a DNA-binding component, AML1 (RUNX1/CBFA2) and the CBFβ subunit. CBF heterodimer transactivates expression of a broad spectrum of genes that are critical for normal haematopoietic development. Thus, chromosomal translocations that resulted in loss-of-function of CBF would be predicted to impair haematopoietic development.

The AML1 protein binds to the co-repressors Groucho/transducin-like enhancer and Sin3 and also to CBP. Therefore, depending on the specific target gene, AML1 is capable of functioning as both a transcriptional activator and a repressor (Fig. 4).

Molecular mechanisms of AML1–ETO translocation in AML

ETO, the partner of AML1 in t(8; 21), is the mammalian homologue of Drosophila nervy and is expressed in haematopoietic progenitors (31). ETO interacts with HDACs, Sin3, N-CoR and SMRT co-repressors (32). The AML1–ETO fusion protein localises to AML1 target genes, where, in contrast to the wild-type AML1 protein, it actively suppresses transcription via the co-repressors N-CoR/Sin3/HDAC1. Because AML1 is required for differentiation of haematopoietic cells, AML1–ETO may be directly responsible for the block in myeloid development and for the leukaemic transformation, in part, by dominantly interfering with the function of the residual normal allele, resulting in a complete loss-of-function phenotype for CBF (33). However, murine model demonstrates that AML1–ETO expression alone is not sufficient to cause leukaemia, suggesting that additional event(s) may be necessary for AML1–ETO-associated leukaemogenesis, at least in murine cancer model. Roeder and co-workers (34) have recently suggested that the ability of AML1–ETO to interfere with E protein transcriptional activation could provide the ‘second hit’ necessary for leukaemia progression.

TEL-AML1 and childhood ALL

TEL–AML1, the most prevalent fusion gene in pediatric cancer, acts as a transcriptional repressor in ALL (35). TEL is a member of the ETS family transcription of factors, which mediates the interaction with several co-repressors including mSin3A, N-CoR and HDAC3. Recruitment of mSin3A and/or N-CoR through the N-terminal part of the TEL moiety possibly converts the fusion protein into an HDAC-dependent constitutive repressor, thus contributing to leukaemogenesis. Recent evidence suggests that TEL–AML1 usually arises prenatally as an early or initiating
mutation (36) and been the ‘second hit’ in most cases deletion of the non-rearranged TEL allele. Because recent data have identified SUV39H1 as a co-factor for AML1 (37), a therapeutic intervention direct against HMTase should be considered.

**Role of lysine acetyltransferases in leukaemia**

As discussed earlier, fusion products of MLL with CBP t(11;16) or with p300 t(11;22) are associated with the pathogenesis of leukaemia. The involvement of alteration of HAT activity in haematological malignancies is further supported by the discovery of chromosomal translocations where CBP (38) or p300 (39) is fused to monocytic leukaemia zinc finger (MOZ). Interestingly, MOZ is also a lysine acetyltransferase enzyme. One likely mechanism by which MOZ–CBP and MOZ–p300 contribute to malignant transformation is by chromatin decondensation and constitutive activation of MOZ-regulated genes. Besides p300 and CBP, other partners are involved in the fusion with the MOZ gene including TIF2, a member of the p160 family of nuclear receptor co-activators known to interact with p300 and CBP (40).

Finally, MOZ-related factor (MORF) is rearranged in a manner similar to the MOZ gene. Indeed, the MORF gene was recently found to be rearranged and fused to the CBP gene in t(10;16) translocation (41). The resulting MORF–CBP fusion protein is structurally similar to the MOZ–CBP/ MOZ–p300 as described earlier and may cause the development of leukaemia by mis-targeting HAT activity and/or deregulating AML1-dependent gene expression.

**EPIGENETIC TRANSCRIPTIONAL SILENCING IN ACUTE PROMYELOCYTIC LEUKAEMIAS**

Acute promyelocytic leukaemias (APLs) are phenotypically characterized by the accumulation of clonal haematopoietic precursors blocked at the stage of promyelocytic cells (42).
Genetically, they are consistently associated with chromosomal translocations that involve the retinoic acid (RA) receptor alpha (RARα) locus on chromosome 17 and one of five different partner genes (PML, PLZF, NUMA, NPM or STAT5b). In the absence of ligand, RAR behaves as a transcription repressor of target genes through binding to specific DNA sequences (so-called RA responsive elements or RARE) and recruitment of co-repressors such as the NCoR–HDAC complex. RA dissociates the NCoR–HDAC complex and leads to recruitment of HATs and other co-activators, thus resulting in chromatin decondensation and transcriptional activation. The fusion proteins cause a block in differentiation by interfering with the RAR signalling pathway, which is important for cellular proliferation and differentiation of the haematopoietic myeloid cellular compartment (43). The most extensively studied, and the most common fusion gene in APL (>95%) is PML–RARα, t(15;17), which involves the promyelocytic leukaemia (PML) gene (16). When compared with RARα, PML–RARα has increased binding efficiency to transcriptional co-repressors, in a manner similar to the AML1–ETO and TEL–AML1 fusions. Furthermore, at physiological concentration of RA (10⁻⁹ to 10⁻⁸ M), PML–RARα remains associated with co-repressors; pharmacological doses of RA (10⁻⁶ M) are needed to dissociate the NCoR–HDAC complex and to recruit HAT enzymes. The stronger affinity for co-repressors could also be responsible for the reduced intranuclear mobility of the fusion protein when compared with that of wild-type RAR (44).

Our group has recently demonstrated that the transcriptional repression of PML–RARα target genes is further reinforced by recruitment of DNMTs, leading to methylation of CpG islands of key promoters (e.g. RARβ2) (45). Once established, the PML–RARα-induced epigenetic chromatin modifications and the resulting gene repression are stable and maintained throughout cell divisions. Importantly, these epigenetic modifications were found to contribute to the leukemogenic potential of PML–RARα, prompting similar analyses also in other leukaemias bearing oncogenic transcription factors. Although pharmacological doses of RA dissociate the NCoR–HDAC complex and restore terminal differentiation of AML blasts in vitro, clinical evidence indicates that RA per se is unable to cure this disease. Similarly, combination of RA and chemotherapeutic agents induce clinical remission, but often patients experiencing relapse. Our hypothesis is that epigenetic modifications persist under these conditions and that combinatorial treatment of RA with HDAC inhibitors and/or DNA demethylating drugs might be required for removal of epigenetic repressor markers.

Several genes have been identified as PML–RARα targets, including members of the C/EBP family of transcription factors (46), Hox genes (47), CL2 and TNFR2 (48) and genes involved in stem cell maintenance and DNA repair (49). Not all of the identified PML–RARα target genes contain a RARE. As recently demonstrated, PML–RARα homodimers are also able to bind to non-consensus RAREs, thus deregulating a much wider network of target genes than previously believed (50).

Like AML1–ETO, expression of PML–RARα is not sufficient to induce leukaemia (51) and additional genetic events might provide ‘second-hits’ for progression to leukaemia. A potential ‘second-hit’ could be provided by the reciprocal fusion protein (RARα–PML) (52). Alternatively, inhibition of the p53 pathway (53,54), activating mutations of FLT3 receptor tyrosine kinase (55) or additional cytogenetic abnormalities could contribute to APL pathogenesis (56).

The nature of the fusion partner has an important impact on the disease characteristics. In fact, PLZF–RARα fusion protein binds to co-repressors not only through the RAR moiety, but also through the PLZF moiety. Thus, RA cannot induce the release of the co-repressor complex bound to the PLZF moiety. Conversely, PLZF–RARα behaves as a constitutively transcriptional repressor even in presence of pharmacological doses of RA (43). It has been recently demonstrated that Bmi-1 member of the Polycomb group (PcG) of proteins repress the HOXd genes via direct interaction with PLZF (57). Thus, similar to MLL, the HOX genes may be the major targets of the PLZF–RARα fusion protein. Polycomb group (PcG) proteins form multimeric chromatin-associated protein complexes (58), which are not only involved in repression of gene activity, but also contribute to the cellular memory system responsible for maintaining the epigenetic status of target genes throughout cell divisions. Therefore, the PLZF–RARα fusion protein could repress gene activity in a hereditable manner. In fact, RA-induced degradation of the fusion protein in leukaemic blast does not restore cell differentiation, suggesting that PLZF–RARα may leave a permanent epigenetic mark on the leukaemic cell and cause heterochromatinization of otherwise transcriptionally active region of the chromatin. As a consequence of the initial repression, the spreading of inactivation into flanking regions may occur, unless boundary elements maintain and protect active domains from this process.

**PERSPECTIVES**

Much of the pathological gene silencing that occurs in cancer is a consequence of the mis-targeting of enzymes involved in chromatin regulation (59,60). This is often accompanied by an enforced oligomerization of the fusion proteins that may alter the binding capacity to DNA or their affinity to co-factors (61), usually co-repressors. Moreover, sequestration and/or mis-localization of co-repressors, such as HDACs and DNMTs, may result in unwanted gene activation of otherwise silenced genes. Although much has been learned, there are still several unanswered questions: how DNA methylation and histone modification influence each other, how transcriptional memory is propagated through cell division, how non-coding RNAs (including micro-RNAs) affect gene regulation and, most importantly, how to re-set the epigenetic equilibrium in tumour cells. Pathologically, silenced genes are uniquely susceptible to drugs that inhibit these enzymes. Indeed, experiments performed using cell lines demonstrated that HDAC inhibitors and demethylating agents synergize in upregulation of silenced genes (62). Unfortunately, the use of HDAC inhibitors (alone or in combination with RA) has been somewhat disappointing (63,64) in APL clinical trials, because of the appearance of blasts resistant to differentiation. Because epigenetic modifications are so intricately interconnected, novel approaches that include demethylating agent and
HMTase inhibitors should be considered to completely re-set the epigenetic alterations of both undifferentiated cells and in those rare cells that posses a stem-like neoplastic phenotype. Such combinatorial therapies may also more effectively target the pre-malignant clone, thus reducing the likelihood of a relapsed disease.

ACKNOWLEDGEMENTS

The work from Di Croce laboratory described in this review was supported by grants from the Ministry of Science and Technology (BFU2004-03862/BMC) and the Fundació ‘La Caixa’ (04/054-00). We thank Veronica Raker and all members of the Di Croce laboratory for stimulating discussions.

REFERENCES

Oncogene, 20, 7186–7205.


