Leukemia (2011) 25, 663–670
© 2011 Macmillan Publishers Limited All rights reserved 0887-6924/11
www.nature.com/leu

ORIGINAL ARTICLE

The heterodimerization domains of MLL—FYRN and FYRC—are potential target structures in t(4;11) leukemia

B Pless[1], C Oehm[1], S Knauer[2], RH Stauber[3], T Dingermann[1] and R Marschalek[1]

[1]Institute of Pharmaceutical Biology/ZAFES, Goethe-University of Frankfurt, Biocenter, Frankfurt/Main, Germany; [2]Institute for Molecular Biology, Centre for Medical Biotechnology (ZMB), University Duisburg-Essen, Universitätsstraße, Essen, Germany and [3]Molecular and Cellular Oncology/Mainzer Screening Center (MSC), University Hospital of Mainz, Mainz, Germany

The chromosomal translocation t(4;11)(q21;q23) is a frequent genetic aberration of the mixed lineage leukemia (*MLL)* gene, predominantly associated with high-risk acute lymphoblastic leukemia (ALL) in pediatric patients. Previous studies demonstrated that mice transplanted with hematopoietic cells expressing the AF4–MLL fusion protein develop proB ALL. The AF4–MLL oncoprotein becomes activated by Taspase1-mediated hydrolysis, which subsequently leads to a heterodimer of the cleavage products AF4–MLL·N and MLL·C. This protein–protein interaction is due to the FYRN and FYRC interaction domains present in both protein fragments. Heterodimerization subsequently induces high-molecular-weight protein complex formation that is protected against SIAH1/2-mediated polyubiquitinylation. Here, we attempted to selectively block this initial heterodimerization step, aiming to prevent the oncogenic activation of the AF4–MLL multiprotein complex. The minimal interaction interface was experimentally defined first in a bacterial two-hybrid system, and then in mammalian cells by using a biosensor assay. Expression of the FYRC domain, or smaller portions thereof, resulted in the inhibition of heterodimer formation, and blocked AF4–MLL multiprotein complex formation with subsequent destruction of the AF4–MLL oncoprotein. Thus, it is in principle possible to specifically target the AF4–MLL protein. This knowledge can now be exploited to design inhibitory decoys in order to destroy the AF4–MLL oncoprotein.

Leukemia (2011) **25,** 663–670; doi:10.1038/leu.2010.308; published online 14 January 2011
Keywords: acute leukemia; MLL; AF4; fusion genes; fusion proteins

Introduction

Chromosomal rearrangements of the mixed lineage leukemia (*MLL*) gene with more than 60 partner genes are frequently found in acute myeloid leukemia and acute lymphoblastic leukemia (ALL).[1] The fusion partners AF4, LAF4 and AF5q31 belong to the ALF protein family,[2] and are associated with the onset and development of an ALL disease phenotype. *AF4* is the most common translocation partner in *MLL*-mediated leukemia, and leads to the creation of the two fusion genes *MLL–AF4* and *AF4–MLL*. However, about 20% of t(4;11) leukemia patients display complex genomic rearrangements with three or more fusion alleles. In these patients, the *AF4–MLL* fusion allele is absent because the 3′-*MLL* gene fragment is fused to other fusion partners.[3] However, it is important to mention that t(4;11)-leukemia patients always express a reciprocal *MLL* fusion allele, because even patients with rare out-of-frame fusions, such as *FXYD6-MLL* or *TCF12-MLL*,[1] are still able to transcribe *MLL* exons 12–37 because of the presence of a gene internal promoter element.[4] This C-terminal portion of MLL comprises 2389 amino acids and encodes a portion of the MLL protein encompassing the PHD domain until the SET domain. This portion is still able to confer H3K4 signatures, however, in an MENIN1-independent fashion.

The pathomolecular mechanism of t(4;11)-mediated ALL is still controversial. Two transgenic mouse knock-in model systems, expressing the MLL–AF4 fusion protein, developed B-cell lymphomas (50 or 70%, respectively) after a long latency (540 or 720 days, respectively).[5,6] The lack of the ALL phenotype in these animal models could be explained by the finding that the MLL–AF4 fusion protein binds to the promoter of the cell cycle inhibitors CDKN1B ($p27^{kip1}$) and CDKN4C (p18INK4c),[7,8] and thus interferes with cell cycle progression. Recently, an *Mll–AF4* knock-in mouse model system has been established demonstrating that the MLL–AF4 allele is able to cause predominantly acute myeloid leukemia, but also ALL and MLL.[9] By contrast, we have demonstrated that the reciprocal AF4–MLL fusion protein causes growth transformation in murine embryonal fibroblasts.[10,11] Additionally, infection of murine hematopoietic progenitor cells (Lin^- $Sca1^+$) with a low titer retrovirus coding for the *AF4–MLL* fusion gene led to the development of predominantly proB ALL in the transplanted mice.[12]

Hence, we investigated the molecular actions of the AF4–MLL oncoprotein in a subsequent study.[13] The AF4–MLL fusion protein is a *bona fide* substrate for Taspase1,[14] similar to the prototypic MLL protein.[15,16] The resulting protein fragments (AF4–MLL·N and MLL·C) heterodimerize through their FYRN and FYRC interaction domains, which have been roughly mapped by two other groups. Yokoyama *et al.*[15] identified two regions within the MLL protein, encompassing the amino acids 1256–2257 (N-terminal domain) and 3610–3745 (C-terminal domain), while Hsieh *et al.*[16] defined regions spanning from amino acids 1978–2161 (N-terminal domain) and 3659–3972 (C-terminal domain). The defined interaction domains included the 40–90-amino-acids-long FYR regions rich in phenylalanine (F) and tyrosine (Y) residues. Such domains are often found in chromatin-associated proteins,[17] and define a novel interface for protein–protein interactions that has been structurally analyzed very recently.[18]

On the basis of our experiments, the heterodimer of AF4–MLL·N and MLL·C forms a high-molecular-weight complex including the positive elongation factor b (P-TEFb; heterodimer of CDK9 and CCNT1/2) and several histone-modifying enzymes.[13] P-TEFb is usually stored in 'large complexes' consisting of HEXIM1, 7SK-RNA, MEPCE and

Correspondence: Professor Dr R Marschalek, Institute of Pharmaceutical Biology/ZAFES, University of Frankfurt, Max-von-Laue-Str. 9, 60438 Frankfurt/Main, Germany.
E-mail: Rolf.Marschalek@em.uni-frankfurt.de
Received 9 September 2010; revised 2 November 2010; accepted 25 November 2010; published online 14 January 2011

LARP7,[19] and is liberated in the presence of AF4 family members or the AF4–MLL·N/MLL·C heterodimer. P-TEFb then activates RNA polymerase II for transcriptional elongation.[20] For this key event, the hepta-peptide repeats of RNA polymerase II (CTD domain) are phosphorylated at serine-2 (P-TEFb) and serine-5 residues (heterodimer of CDK7 and CCNH). In addition, the associated repressor proteins DSIF and the NELF complex are phosphorylated by P-TEFb. Consequently, DSIF is converted into a positive transcription elongation factor while the NELF complex is destroyed. All these events convert promoter-proximal arrested RNA polymerase II into elongating polymerase.[21] The SET-domain complex and the associated histone methyltransferases DOT1L and CARM1 subsequently modify specific arginine and lysine residues of histone H3. This leads to activated chromatin and enhanced transcription. Over time, enhanced transcription and ectopic chromatin signatures result in a genome-wide reprogramming of affected cells. Finally, these processes lead to cells with tumorigenic properties, and presumably to the establishment of tumor stem cells.

Of importance, both the AF4 protein and the AF4–MLL fusion protein are substrates for the E3-ubiquitin ligases SIAH1 and SIAH2.[10] Therefore, their steady-state level is normally very low. However, after the AF4–MLL fusion protein is processed by Taspase1, the two resulting protein fragments form a high-molecular-weight protein complex of about 2 MDa, which is resistant to SIAH-mediated degradation.[13] Thus, we hypothesized that either blocking the hydrolysis of the AF4–MLL fusion protein by inhibiting Taspase1 (Bier *et al*.[22]) or blocking the process of complex formation may prevent leukemogenesis.

We were able to define the minimal FYRN and FYRC domains, still able to confer specific binding, by using a commercially available bacterial two-hybrid system and a recently established *in vivo* biosensor system.[23] Expression of FYRC-derived peptides in HEK293T cells led to a selective inhibition of the heterodimerization process. We also investigated the fate of the Taspase1-hydrolyzed AF4–MLL fusion protein in the absence or presence of co-expressed FYRC peptides. Upon expression of the FYRC-derived peptides B1 and B3, both the AF4–MLL·N and MLL·C portion were rapidly destructed through the proteasomal pathway, thus eliminating the AF4–MLL oncoprotein in these cells. This proof-of-concept study demonstrates that the AF4–MLL oncoprotein can be specifically targeted, and potential implications of this finding will be discussed.

Materials and methods

Expression plasmids, B-2-H experiments and cell culture

For initial testing of protein–protein interactions, we used a commercially available bacterial two-hybrid system (Bacterio-Match II Two-Hybrid System Vector Kit, B2H-System; Stratagene, Amsterdam, The Netherlands). Briefly, the two plasmids pBT (expresses a modified lambda C1 repressor in fusion with domain of interest) and pTRG (expresses a λ RNA polymerase in fusion with the domain of interest) were used to clone the potential FYRN (MLL amino acids 1867–2164) and two different FYRC domains (FYRC1: MLL amino acids 3659–3815; FYRC2: MLL amino acids 3659–3972). Both plasmids were co-transformed into *Escherichia coli* encoding a reporter gene (λ operator, and λ minimal promoter fused to the *HIS3/aada* gene). The presence of both plasmids was selected by using the antibiotics chloramphenicol (20 μg/ml) and tetracycline (12.5 μg/ml). Potential protein interactions were monitored by bacterial cell growth on His-deficient media and in the presence of the inhibitor 3-amino-1,2,4-triazol (5 mM), which can be overcome by bacteria only by a strong transcriptional activation of the *HIS3* gene. A protein of interest (the bait) is fused to the bacteriophage lambda repressor protein (lambda cI), containing the DNA-binding domain. The target protein is fused to the N-terminal domain of the α-subunit of RNA polymerase. Thus, the bait is bound to the lambda operator sequence upstream of the *HIS3* gene, and an interaction of bait and target leads to transcription of *HIS3* by recruiting RNA polymerase. As the *aada* gene codes for streptomycin resistance, all positive interactions were validated by growth on Petri dishes containing 12 μg/ml streptomycin. As positive control, we used two plasmids that express two interacting protein domains (pBT-LGF2 and pTRG-Gal11p). All tested domains were co-transformed together with the corresponding empty vector (negative controls).

For protein expression and analysis of purified protein complexes, complementary DNAs (cDNAs) coding for parts of the MLL interaction domain were cloned either into the p3-NLS-GFP-GST-NES vector (bait; FYRC fragments) or into the pc3-REV-BFP vector (prey; FYRN fragments). The cDNA coding for the AF4–MLL fusion protein was cloned into the backbone of either the pEXPR-IBA105 or the pEXPR-IBA103 vector (IBA GmbH, Göttingen, Germany), producing an AF4–MLL fusion protein that is fused either with an N- or with a C-terminal Strep-tag. HEK293T and HeLa cells were maintained in Dulbecco's modified Eagle's medium. All media were supplemented with 10% fetal calf serum, 1% L-glutamine and 1% Pen/Strep.

Transfection of adherent and suspension cells

For microscopic analyses, 1×10^6 HeLa cells were seeded out in Petri dishes with sterile cover glasses 1 day before transfection. The cells were then transfected with 10 μg of each plasmid with lipofectamine 2000 (Invitrogen, Darmstadt, Germany). For protein purifications, HEK293T cells were transfected using the PEI method (Polyethylenimine; Sigma, Munich, Germany). Generally, 5×10^6 cells were transfected with 25 μg of each plasmid.

Fixation of HeLa cells for microscopy

For microscopic analyses, 48 h after transfection, the medium was removed from Petri dishes and the cover glasses were washed with phosphate-buffered saline. The cells were then fixed for 10 min with 4% paraformaldehyde in phosphate-buffered saline and washed first with phosphate-buffered saline, second with 0.1% Triton X-100, third with 0.1% $NaBH_4$ and 3 times with 0.1% Triton X-100 in phosphate-buffered saline (each washing step for 5 min). After the final washing step, the cover glasses were placed on microscope slides, dowsed with 20 μl Mowiol solution and dried for 12 h at 4 °C in the dark.

Strep-Tactin affinity chromatography

A total of 2.5×10^7 transiently transfected HEK293T cells were lysed 48 h after transfection in immunoprecipitation buffer (150 mM NaCl, 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.5, 1% Triton X-100, 1 mM Na_3VO_4, 10 mM NaF, 1 mM PMSF, 1 × Protease Inhibitor Cocktail V, EDTA free from Calbiochem, Darmstadt, Germany). Where necessary, proteasomal degradation was prevented by adding 5 μM MG132 (Calbiochem) 16 h before lysis. Proteins were affinity purified using Strep-Tactin Superflow (IBA GmbH) and chromatography columns. After protein lysis (1 h, 4 °C), cell debris was pelleted and the supernatant was blocked with Avidin (15 min, 4 °C). The lysate was then incubated with Strep-Tactin Superflow in chromatography columns for 1 h at

25 °C. Washing and elution were performed according to the manufacturer's instructions.

Immunoblot analyses

Samples were loaded on an SDS-PAGE and transferred onto a Hybond P membrane (GE Healthcare, Munich, Germany). AF4–MLL was detected with the monoclonal mouse antibody 6E95 (anti-AF4-N). Endogenous actin was used to normalize protein concentrations and was detected by a monoclonal rabbit β-actin antibody (Sigma). Transferred proteins were visualized with the ECL detection system (GE Healthcare). The following antisera were used: anti-AF4N,[10] -MLL, -WDR5, -RBBP5 antisera obtained from Abcam (Cambridge, UK); anti-EGFP antiserum obtained from Santa Cruz Biotechnologies (Heidelberg, Germany).

Structural comparison

All structural comparisons were made with the 2WZO.pdb datafile, available from the RCSB database (http://www.rcsb.org/pdb/home/home.do), in conjunction with PyMOL (http://www.pymol.org).

Results

Putative interaction interface of the MLL interaction domains FYRN and FYRC

Specific endoproteolysis of the AF4–MLL oncoprotein by Taspase1, followed by heterodimerization of AF4–MLL·N and MLL·C, protects the AF4–MLL fusion protein against SIAH1/2-mediated proteasomal degradation.[10] Thus, hydrolysis and subsequent heterodimerization is a key event for activating the process of deriving oncogenic functions from the AF4–MLL fusion protein. The protein interaction domains necessary for this interaction, FYRN and FYRC, have been initially identified by Yokoyama *et al*.[15] and Hsieh *et al*.[16] However, both the groups mapped these interaction domains at different positions within the MLL protein sequence. As depicted in Figure 1a, the FYRN interaction domain was mapped to amino-acid positions 1256–2257 or 1978–2161, while the FYRC interaction domain was mapped to 3610–3745 or 3659–3872.[15,16] In order to refine these domains, the corresponding cDNA fragments were amplified from the *MLL* cDNA and cloned into adequate plasmids (pBT and pTRG) to perform interaction studies using a bacterial two-hybrid system. These experiments revealed that the predicted FYRN (amino acids 1867–2164) and FYRC1 domain (amino acids 3659–3815) are able to interact with each other (see Figure 1b). Less interaction capacity was observed when we tested the FYRN domain in combination with the longer FYRC2 peptide (containing FYRC1 and the C-terminal SET-domain; amino acids 3659–3972). Subsequently, we aimed to define the minimal interaction interface by concentrating on the regions that contain a high content of phenylalanines and tyrosines (FY-rich). Two regions were identified that fulfilled these prerequisites. These newly defined shorter FYRN (amino acids 1991–2104) and FYRC domains (amino acids 3651–3752) are displayed in Figure 1c. Besides their high content of F and Y, the FYRN domain could be separated by proline residues into four short subregions that exhibit a distinct pattern of charged amino acids that are somehow mirrored in the four cognate subregions of the FYRC domain. As depicted in Figure 1c, these subregions are between 20 and 34 amino acids long and are mostly separated by proline residues. On the basis of this hypothetical structure, the FYRN domain consists of a total of 110 amino acids, while the FYRC domain consists of only 103 amino acids.

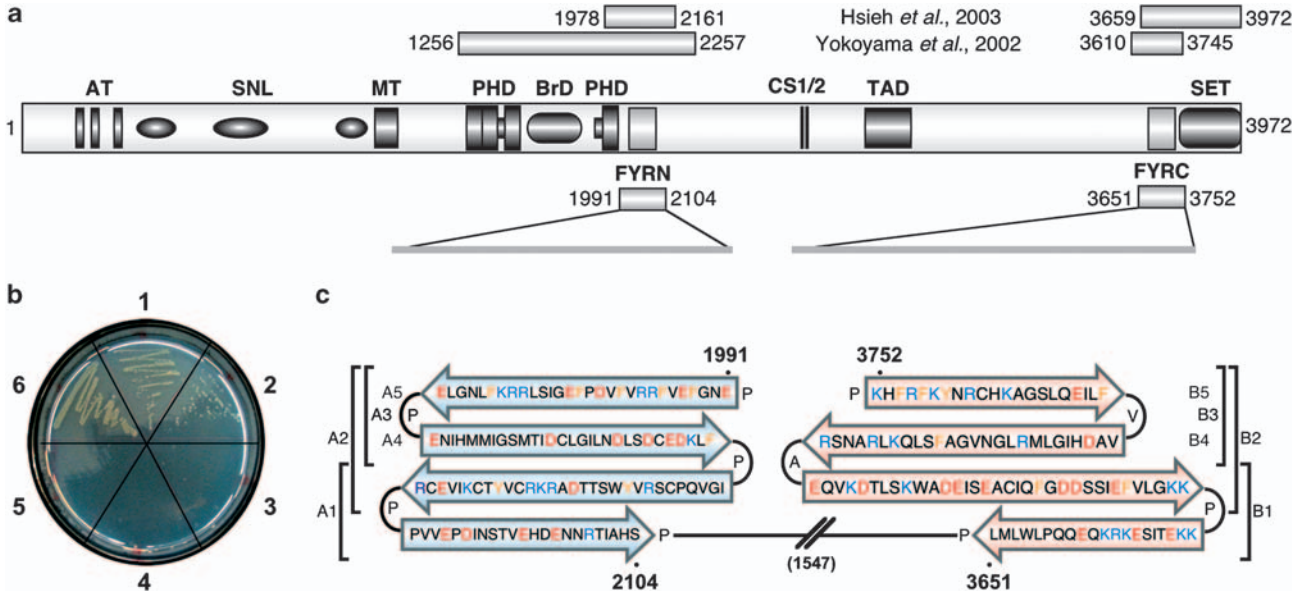

Figure 1 Hypothetical design of the interaction interface of MLL. (**a**) Structure of the MLL protein with all known domains, including the N-terminal and C-terminal F/Y-rich domains of MLL (FYRN and FYRC) that have been mapped by two different groups to the indicated amino-acid positions.[15,16] Below the MLL protein: identified FYRN and FYRC interaction domains. All amino-acid positions are relative to the MLL sequence with 3.972 amino acids as described in Meyer *et al*.[24] Numbers represent amino-acid positions. (**b**) Results of the B-2-H experiments. Selection was carried out on His^- agar containing tetracycline, chloramphenicol and 3-AT. Plasmids used in co-transformation experiments were (1) pBT-FYRN and pTRG-FYRC1; (2) pBT-FYRN and pTRG-FYRC2; (3) pBT and pTRG-FYRC1; (4) pBT and pTRG-FYRC2; (5) pBT-FYRN and pTRG; (6) pBT-LGF2 and pTRG-Gal11p (positive control). (**c**) Newly defined structures of the MLL FYRN and FYRC domains**.** All tested subfragments (A1–A5 and B1–B5) are indicated. Numbers represent amino-acid positions. The FYRN- (amino acids 1991–2104) and FYRC-domain (amino acids 3651–3752) are divided into four subregions separated by proline- (or valine- and alanine-) residues, respectively. Negatively charged residues (E and D) are depicted in red, positively charged residues (K and R) are depicted in blue and aromatic residues (F and Y) are depicted in orange.

Interaction of the fragments based on the hypothetical interaction interface

On the basis of the hypothetical structure of this newly defined interaction interface, the shorter cDNA fragments were cloned into a bait and a prey vector comprising a mammalian biosensor system (see Figure 2).[23] Basically, the bait vector encodes a GFP–GST fusion protein that exhibits a weaker NLS and a stronger NES at the N- and C-terminal end, respectively. The prey vector encodes a REV–BFP fusion protein that anchors to the nucleolus. Both the hypothetical FYRN and FYRC domains were cloned into the interaction domain cloning sites (ID1 and ID2) of both vectors, and co-transfected into HEK293T cells. The expressed NLS–GFP–GST–FYRC–NES fusion protein shuttles between the nucleus and cytosol, while the FYRN–REV–BFP fusion protein localizes to the nucleolus. Intracellular localization was then monitored by fluorescence microscopy.

As shown in Figure 3a, the co-expression of both fusion proteins resulted in their nucleolar accumulation; which indicated a direct interaction between the FYRN and FYRC domains. Appropriate control experiments with corresponding mock-bait or mock-prey vectors always resulted in the cytosolic localization of the bait protein (FYRC) or the nucleolar localization of the prey protein (FYRN; see Figure 3a: controls). From these initial experiments, we concluded that the proposed structures of FYRN and FYRC indeed represent an interaction interface for intramolecular heterodimerization of Taspase1-hydrolyzed MLL or AF4–MLL.

In order to refine the interaction interface, both the FYRN and FYRC domains were further subdivided into smaller protein fragments (see Figure 1c). Each of these peptides (A1–A3; B1–B3) resembles particular subregions of the proposed FYRN and FYRC domains. These smaller fragments were again cloned into the bait and prey vector system. As shown in Figure 3b, the three FYRC-derived peptides B1, B2 and B3 again accumulated in the nucleus when co-expressed with the corresponding FYRN-derived peptides A1, A2 and A3, respectively. This demonstrated that the predefined substructures also retained their ability to bind to each other. Cross-testing these peptides (for example, B1 with A2 or A3) revealed no such interactions (see Supplementary Figure S1). Again, when single peptides were co-transfected with the corresponding mock-bait or mock-prey controls, no nuclear accumulation of the prey proteins was observed (Figure 3, controls). Thus, translocation into the nucleus and nucleolar accumulation could be attributed to specific interactions between the corresponding FYRC- and FYRN-derived peptides.

To test the binding specificity of even smaller fragments, the peptides FYRN-A3 (57 amino acids) and FYRC-B3 (56 amino acids) were further separated into the smaller fragments A4 (28 amino acids), A5 (30 amino acids), B4 (27 amino acids) and B5 (29 amino acids; see Figure 1c). When A4 or A5 were co-expressed with B3, no retention of B3 at the nucleolus was observed. This indicated that the interaction of these smallest subfragments was not strong enough to cause nucleolar localization of the bait protein. B4 exhibited a non-specific interaction with the mock-prey (REV-BFP). Co-expression of A5 and B5 revealed a weak retention of B5 in the nucleus (data not shown). On the basis of these data, we decided to concentrate on the peptides A1–A3 and B1–B3, respectively.

Inhibition of the heterodimerization of AF4–MLL · N and MLL · C is associated with a higher turn-over of the AF4–MLL fusion protein and inhibits multiprotein complex formation

To verify that GFP-B1 and GFP-B3 are able to bind not only to their cognate recombinant counterparts (A1 and A3) but also to the FRYN domain of the AF4–MLL protein, HEK293T cells were co-transfected with an N-terminally Strep-tagged AF4–MLL expression plasmid together with either GFP-B1 or GFP-B3 expression constructs. Cells were treated for 16 h with MG132 to block any proteasomal degradation. After cell lysis, the

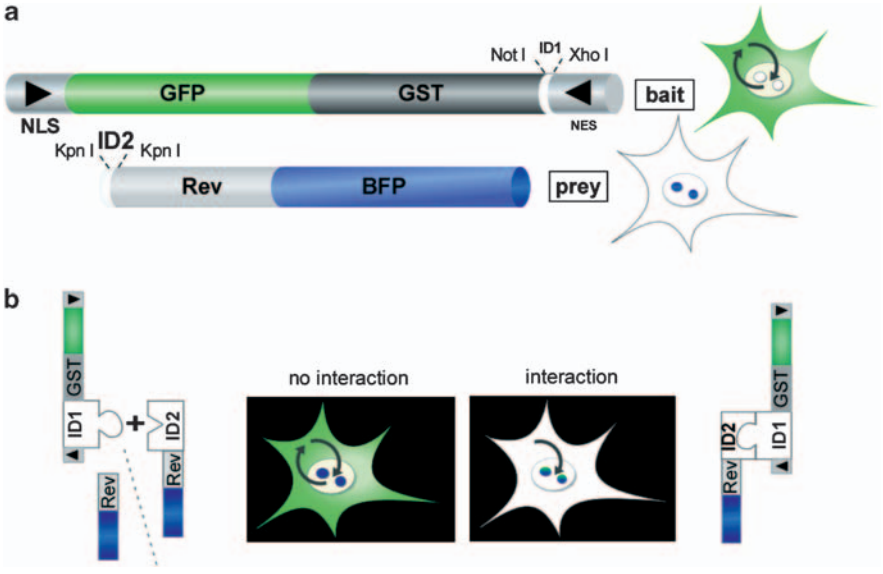

Figure 2 Principles of the protein-interaction biosensor assay. (**a**) Schematic illustration of the domain organization of the BFP–prey and GFP–bait proteins. The GFP–bait plasmid encodes the SV40 NLS, GST, GFP and the HIV-1 RevNES. The BFP–prey plasmid codes for a fusion of the HIV-1 RevM10BL protein fused to BFP, anchoring the recombinant expressed protein to the nucleolus. Unique restriction sites (*Not*1, *Xho*1 and *Kpn*1) for the cloning of the respective interaction domains (FYRN or FYRC) are indicated. (**b**) Principle of the translocation-based biosensor system. GFP–bait fusion proteins continuously shuttle between the cytoplasm and the nucleus. Owing to a stronger export signal, the protein is predominantly found in the cytoplasm. In case of a protein interaction, the prey protein co-localizes with the bait that is anchored to the nucleolus. Filled forward triangle: NLS; filled reverse triangle: NES.

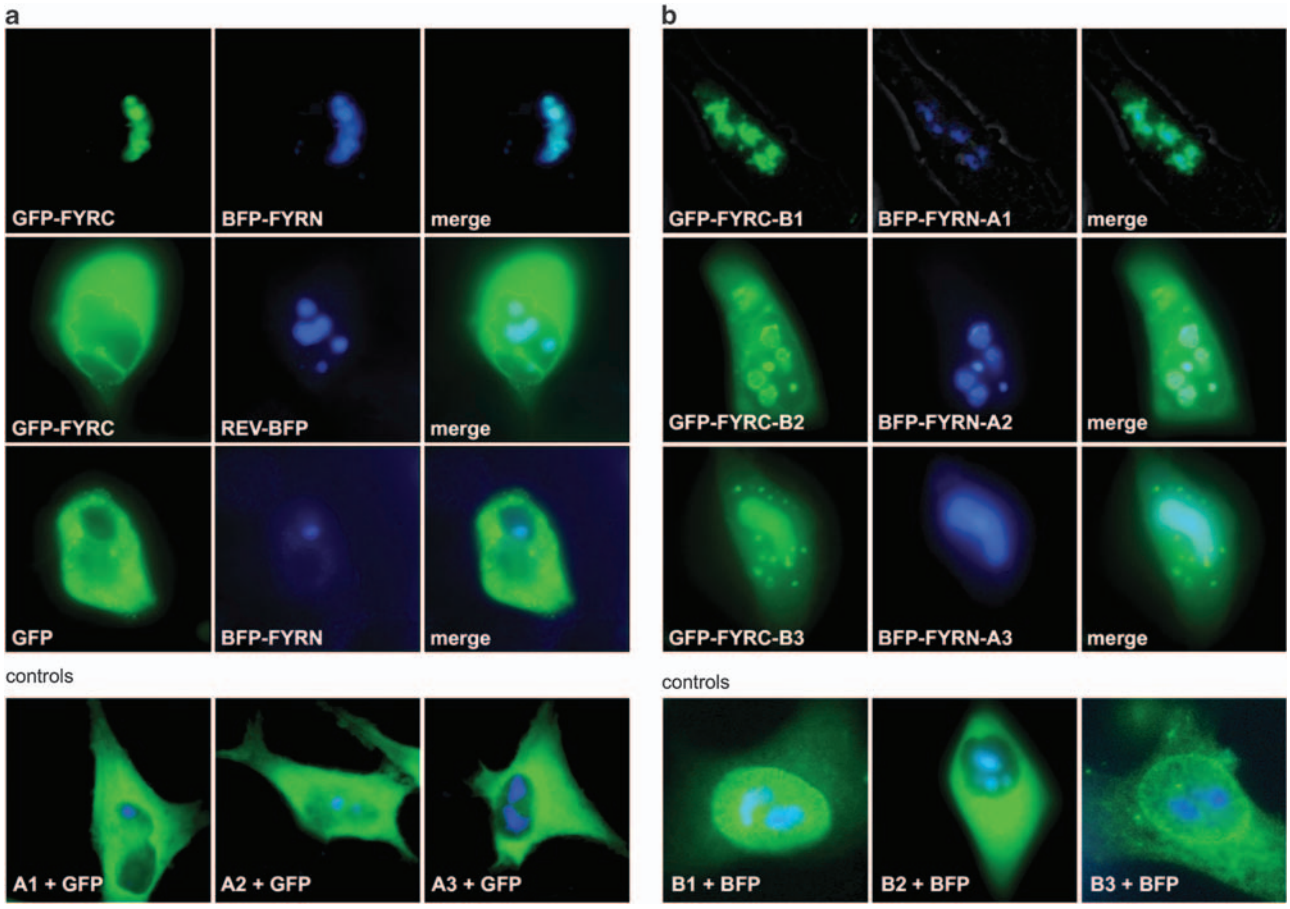

Figure 3 Interaction of FYRC-peptide fragments with the interaction interface of MLL. Peptides representing 2–4 subregions of the hypothetical interaction interface as shown in Figure 1 were cloned into the prey and bait vectors of a biosensor system. The prey protein (fused to BFP) is associated with the nucleolus, while the bait protein (fused to GFP) shuttles between the cytosol and nucleus. (**a**) Upper panel: the expression of GFP-FYRN and BFP-FYRC leads to the translocation of FYRC-GFP to the nucleolus, indicating an interaction between FYRC and FYRN. Middle and lower panels display various combinations of control experiments (GFP-FYRC with the empty mock–prey Rev–BFP; empty mock–GFP–bait with BFP-FYRN). Both control experiments do not show any interactions. Controls: additional control experiments were performed for the tested A1–A3 proteins (fused to BFP; prey vector) and the mock–GFP–bait protein; in all these experiments, no co-localization between prey and control proteins was observed. (**b**) The expression of the peptides GFP-B1 to -B3 and BFP-A1 to -A3. All tested combinations co-localized at the nucleoli, indicating their ability for specific interaction. Controls: additional control experiments were performed for the tested B1–B3 proteins (fused to the GFP bait) and the mock-BFP-prey control protein; in all these experiments, no co-localization between bait and control prey-proteins was observed.

N-terminally Strep-tagged AF4–MLL and all directly bound proteins were subsequently purified with the Strep-Tactin affinity system. As shown in Figure 4a, the amount of the affinity-purified Strep-tagged AF4–MLL·N fragment (p178 + 5 kDa Strep-tag = 183 kDa, upper panel) was equal in all transfected cells under these conditions. GFP-B1 and GFP-B3 were specifically bound to the affinity-purified AF4–MLL·N protein (second panel). The association to the MLL·C fragment was completely blocked in the presence of GFP-B1 or GFP-B3 (third panel). We also tested for the presence of WDR5 and RBBP5, two components of the SET-domain complex, and reduced binding was observed when compared with the GFP control. This indicated that the presence of B1- and B3-peptides not only inhibits heterodimer formation but also interferes with the establishment of the oncogenic AF4–MLL complex. Control blots performed with cell lysates of these transfected cells proved the presence of GFP-B1, GFB-B3 and GFP, as well as equal amounts of actin in all transfected cells.

As shown in Figure 4b, co-expression of AF4–MLL with GFP-B1 or GFP-B3 in the absence of proteasomal inhibitors led to dramatically reduced amounts of the Taspase1-hydrolyzed AF4–MLL·N protein fragment (p183), whereas the co-expression of GFP revealed no reduction. This is in line with data showing that the MLL·N protein fragment becomes resistant to proteasomal degradation when complexed with the MLL·C protein fragment.[16] Thus, expression of B1- or B3-peptides seems to result in a higher turnover of the AF4–MLL·N protein fragment, most likely due to the interaction with the two E3-ligases SIAH1 and SIAH2, both of which bind specifically to the PxAxVxP motif located at amino acids 290–296 of the fused AF4 protein sequences.[10] Actin control blots demonstrated again that equal amounts of protein were loaded in all four lanes.

Next, we used a similar experiment; however, this time we purified a C-terminal Strep-tagged AF4–MLL fusion protein in the absence or presence of the proteasomal inhibitor MG132 (see Figure 4c). Again, we were able to demonstrate that MLL·C is also degraded when GFP-B1 or GFP-B3 were co-expressed (panel 1), whereas GFP alone had no such effect on the stability of MLL·C (panel 4). In the absence or presence of MG132, GFP-B1, GFP-B3, GFP and actin were equally expressed (panels 2, 3, 5 and 6). This implicated that the AF4–MLL·N and the MLL·C protein fragments are both subject to proteasomal degradation when the heterodimerization process is blocked in the presence of overexpressed B1- or B3-peptides, respectively.

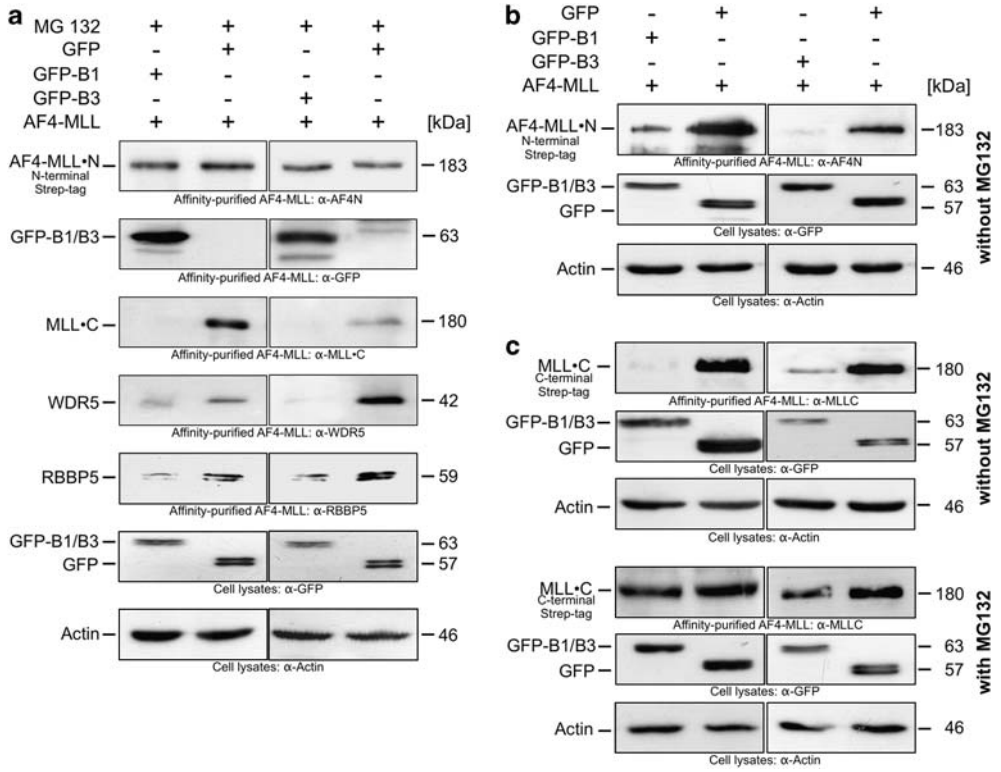

Figure 4 The effect of GFP-B1 and GFP-B3 on the complex formation of AF4–MLL. HEK293T cells were co-transfected with Strep-tagged AF4–MLL and GFP, GFP-B1 or GFP-B3. After transfection, Strep-tagged AF4–MLL•N was chromatography-purified and immunoblots were performed with indicated antibodies. As control, an aliquot of the cell lysates was also subjected to western blot analyses. (**a**) Affinity purification of AF4–MLL (N-terminal Strep-Tag) in the presence of MG132. Panel 1: the amount of AF4–MLL•N is now similar after affinity purification. Panel 2: only GFP-B1 or GFP-B3 is bound to the affinity-purified AF4–MLL protein, while GFP is not. Panel 3: MLL•C is only bound in GFP co-transfected cells, while expression of GFP-B1 or GFP-B3 abolishes nearly all the binding of MLL•C to AF4–MLL•N. Panels 4 and 5: bound WDR5 or RBBP5 protein is reduced in GFP-B1 or GFP-B3 transfected cells. Panel 6: amount of recombinant GFP, GFP-B1 or GFP-B3 in cell lysates. Panel 7: actin control blots. (**b**) Affinity purification of AF4–MLL (N-terminal Strep-Tag) in the absence of MG132. Panel 1: the amount of the Strep-tagged AF4–MLL•N is reduced in the presence of GFP-B1 or GFP-B3 compared with the GFP control. Panel 2: GFP-B1, GFP-B3 and GFP are present in equal amounts in the cell lysates. Panel 3: actin control blots. (**c**) Affinity purification of AF4–MLL (C-terminal Strep-Tag) in the absence and presence of MG132. Panel 1: the amount of the Strep-tagged MLL•C is reduced in the presence of GFP-B1 or GFP-B3 compared with the GFP control. Panel 2: GFP-B1, GFP-B3 and GFP are present in equal amounts in the cell lysates. Panel 3: actin control blots. Panel 4: the amount of the Strep-tagged MLL•C is equal in the presence of MG132. Panel 5: GFP-B1, GFP-B3 and GFP are present in equal amounts in the cell lysates. Panel 6: actin control blots.

Discussion

Overexpression of AF4 or AF4–MLL has been associated with a loss-of-contact inhibition in mammalian cells.[10,11] From these data it was concluded that AF4 behaves like a proto-oncoprotein. Recent efforts have unravelled the function of the AF4 protein. The AF4 protein forms a high-molecular-weight protein complex containing P-TEFb, ENL/AF10/DOT1L and several other proteins.[13] P-TEFb kinase activity is essential to convert arrested RNA polymerase II into its elongation state. Thus, the assembly of the AF4 multiprotein complex seems to represent a crucial step for all transcriptional processes. Several of these activating steps have recently been identified and investigated in the murine system.[25] Thus, overexpression of AF4 leads to ectopic transcription, while the absence of AF4 causes a dramatic decrease of transcriptional processes (our own unpublished observations). The steady-state amount of the AF4 protein and its complex is strictly controlled by the two E3-ligases SIAH1 and SIAH2 that confer polyubiquitinylation, and subsequently proteasomal degradation.[10] Likewise, the full-length AF4–MLL fusion protein is also targeted for SIAH-mediated degradation. However, once processed by Taspase1, the two resulting protein fragments AF4–MLL•N and MLL•C heterodimerize through their FYRN/FYRC interaction domains, and form a high-molecular-weight protein complex that is resistant to SIAH1/2-mediated degradation. Thus, the high-molecular-weight protein complex can act like a dominant-positive variant of AF4 that activates transcriptional processes (about 50-fold) and causes ectopic chromatin signatures in a genome-wide fashion ($H3K4_{me3}$, $H3K79_{me3}$, $H3R2_{me}$, $H3R17_{me}$, $H3R26_{me}$, $H2A_{Ac}$, $H2B_{Ac}$ and $H3_{Ac}$).[13] Owing to its resistance to proteasomal degradation, the AF4–MLL complex accumulates in cells expressing the *AF4–MLL* allele. The activities exerted by the AF4–MLL protein complex seem to be oncogenic, as transduced hematopoietic stem/precursor cells expressing the AF4–MLL fusion protein caused proB ALL after transplantation into isogenic mice.[12]

Therefore, we attempted here to investigate the process of heterodimerization—mediated by the FYRN and FYRC interaction domains—and to use this knowledge to interfere with one of the initial steps that subsequently lead to the formation of the oncogenic AF4–MLL multiprotein complex. For the purpose of our studies, we first confirmed that these two regions of the MLL protein (FYRN and FYRC1, or FYRN and FYRC2) are indeed able to interact (see Figure 1b), similar to that reported in earlier studies.[15,16] Next, we analyzed the amino-acid composition of these experimentally defined regions. Both regions display a high content of phenylalanine and tyrosine residues, and the

pattern of charged amino acids in combination with several proline residues provoked us to postulate that the MLL regions 1991–2104 (FYRN) and 3651–3752 (FYRC) represent an interaction interface (see Figure 1c). Subsequently, these smaller regions were cloned and tested in a mammalian biosensor system.[23] This system is based on a bait protein that is able to shuttle between the cytosol and nucleus, whereas the prey protein is anchored to the nucleolus. If a protein interaction occurs, both bait and prey localize to the nucleolus. In the case that no protein interaction occurs, the bait protein is predominantly found in the cytosol, while the prey protein stains the nucleolus (see Figure 2).

We first confirmed in this *in vivo* system that our hypothesized structures (FYRN: 110 amino acids; FYRC: 103 amino acids) were indeed able to confer heterodimerization. These domains were then further dissected by using combinations of two or three out of the four substructures (between 48 and 88 amino acids long, see Figure 1c). These experiments again confirmed that both domains interact in a mirrored fashion, as expected from the hypothetical structures (see Figure 3). However, fragments shorter than 30 amino acids (only one substructure) were not able to bind to their cognate counterparts.

In order to demonstrate efficient binding of these shorter protein fragments to the full-length protein, we co-expressed a Strep-tagged AF4–MLL fusion protein together with the GFP-B1-(48 amino acids) or GFP-B3-peptide (56 amino acids), both derived from the C-terminal FYRC domain. We expected that both peptides should be able to bind to their cognate structures within the FYRN domain of AF4–MLL, and thus should prevent or interfere with heterodimerization of the AF4–MLL·N and MLL·C protein fragments. In these experiments we were able to demonstrate that heterodimerization between the AF4–MLL·N and MLL·C protein fragments could be completely blocked. Moreover, the abundance of either the AF4–MLL·N or the MLL·C fragment was significantly reduced, indicating that the heterodimerization step is important for the steady-state abundance of the AF4–MLL oncoprotein. Moreover, reduced amounts of bound WDR5 and RBBP5 were observed. Both proteins have been recently found to be part of the affinity-purified AF4–MLL complex,[13] and are also part of the SET-domain complex (ASH2L, WDR5, RBBP5 and SPY-30). This led to the conclusion that the heterodimerization domains of AF4–MLL (and MLL) can potentially be targeted by small peptides.

Fortunately, a recent study has investigated the crystal structure of the FYRN/FYRC domains derived from the transforming growth factor beta regulator 1 protein, which inhibits cell growth and has a role in maintaining chromosomal stability.[18] This study has aligned the crucial portions to other proteins that display FYRN/FYRC domains, including the MLL protein. As shown in Figure 5a (left panel), the core interaction domains of transforming growth factor beta regulator 1 are based on specific stacking interactions between specific tyrosine (FYRN: Y_{212}, Y_{222}) and specific phenylalanine residues (FYRC: F_{247}, F_{293}). Moreover, several side-chain interactions between positively (K) and negatively charged amino acids (D, E) seem to stabilize the structure of (R_{202}::D_{240}) and between both domains (K_{239}::E_{248}; K_{226}::D_{254}). When comparing the established structure at these residues with that of the AF4–MLL/MLL protein, it becomes obvious that most of these key positions are not conserved, and thus are not able to confer the predicted interactions (see Figure 5a, right panel). Basically, García-Alai *et al.*[18] have analyzed an FYRN/FYRC structure that corresponds to our FYRN subdomains 2 and 3 and to our FYRC substructures

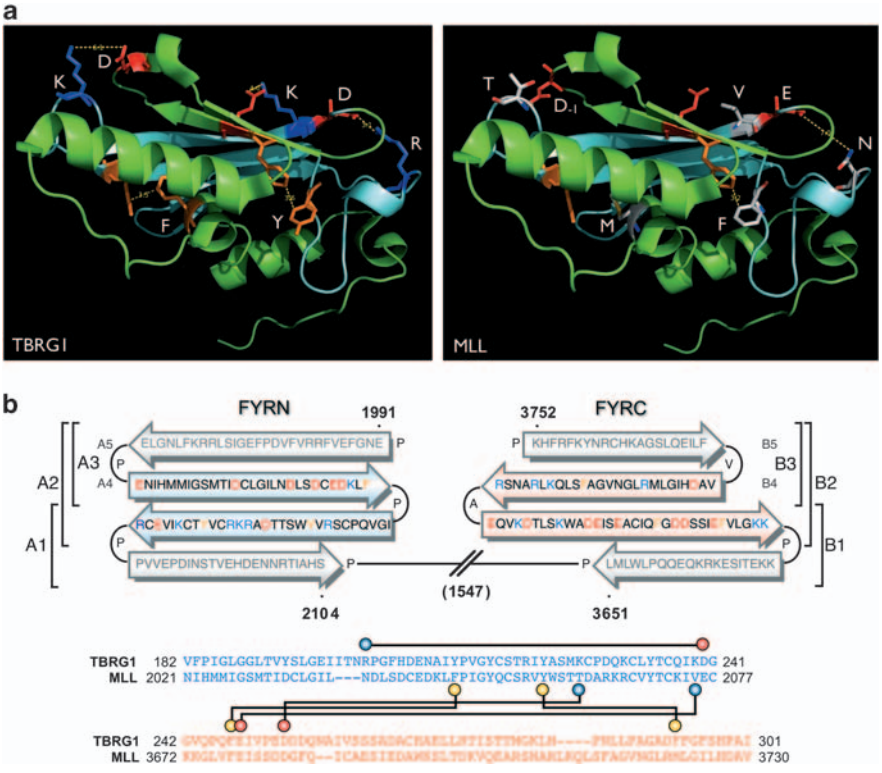

Figure 5 Comparison of the FYRN/FYRC domain of transforming growth factor beta regulator 1 (TBRG1) and MLL. (**a**) Left: the structure of the FYRN (blue) and FYRC (green) interaction interface of TBRG1. Crucial amino-acid positions (orange: aromatic amino acid; blue: positively charged amino acid; red: negatively charged amino acid) are indicated. Right: same structure mutated into the amino acids of the MLL protein sequence: only one interaction remains. (**b**) The hypothetical structure of the MLL interaction interface used in this study. Blue and red areas are shown with their sequence below, while the gray structures are not shown. These were the regions defined as the interaction interface of the TBRG1 protein. All crucial amino-acid positions are marked by colored dots and interactions are shown by the corresponding lines.

2–4. The important interactions are located in substructures 2 and 3 of both interaction domains (see Figure 5b, upper panel). The crucial interactions are depicted in the lower panel. Only a single Phe_{2045}::Phe_{3671} stacking interaction remains within the MLL-derived FYRN and FYRC structures. Taking into account the data of this recent study and our own data, we conclude that the FYRN/FYRC interaction interface in the MLL protein should be extended to the four substructures that we have analyzed to confer a stable interaction. Two or three out of these four substructures are still able to confer stable protein interactions within the AF4–MLL·N/MLL·N and the MLL·C protein fragment, whereas single substructures are unable to do so.

The results presented here demonstrate for the first time that the interaction interface of AF4–MLL fusion protein can be specifically targeted. On the basis of our knowledge about the oncogenic functions derived from the AF4–MLL fusion protein, it may also indicate that these domains are potential therapeutic targets in t(4;11) leukemia. The precise definition of FYRN and FYRC in combination with their relatively small length will allow to characterize their three-dimensional structure. On the basis of such a structure, the design of small molecules that potentially interfere with the heterodimerization process can be performed. We are also planning to stably express B1 or B3 in t(4;11) cell lines. However, based on our own experiences with AF4–MLL, a protein with very long half-life, and its epigenetic reprogramming activity, we expect that only long-term experiments using these novel protein inhibitors will unravel their therapeutic potential. However, this may open an interesting venue to inhibit the process of deriving oncogenic properties from the AF4–MLL oncoprotein.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

This work has been supported by grant 102363 from the Deutsche Krebshilfe e.V. to RM and RS. The biosensor assay has been established with the support of Boehringer Ingelheim and Wilhelm-Sander Foundation, Funds of the Chemical Industry, and the Stiftung Rheinland-Pfalz für Innovationen.

References

1 Meyer C, Kowarz E, Hofmann J, Renneville A, Zuna J, Trka J *et al.* New insights to the MLL recombinome of acute leukemias. *Leukemia* 2009; **23**: 1490–1499.
2 Nilson I, Reichel M, Ennas MG, Greim R, Knörr C, Siegler G *et al.* The exon/intron structure of the human AF-4 gene, a member of the AF-4/LAF-4/FMR-2 gene family coding for a nuclear protein with structural alterations in acute leukemia. *Brit J Haematol* 1987; **98**: 157–169.
3 Kowarz E, Burmeister T, Nigro LL, Jansen MWJC, Delabesse E, Klingebiel T *et al.* Complex MLL rearrangements in MLL·AF4+/AF4·MLL- leukemia patients conceal the presence of reciprocal MLL fusion genes. *Leukemia* 2007; **21**: 1232–1238.
4 Scharf S, Zech J, Bursen A, Schraets D, Oliver PL, Kliem S *et al.* Transcription links to recombination: a gene-internal promoter coincides with the recombination hotspot II of the human MLL gene. *Oncogene* 2007; **26**: 1361–1371.
5 Chen W, Li Q, Hudson WA, Kumar A, Kirchhof N, Kersey JH. A murine Mll-AF4 knock-in model results in lymphoid and myeloid deregulation and hematologic malignancy. *Blood* 2006; **108**: 669–677.
6 Metzler M, Forster A, Pannell R, Arends MJ, Daser A, Lobato MN *et al.* A conditional model of MLL-AF4 B-cell tumourigenesis using invertor technology. *Oncogene* 2006; **25**: 3093–3103.
7 Xia ZB, Popovic R, Chen J, Theisler C, Stuart T, Santillan DA *et al.* The MLL fusion gene, MLL-AF4, regulates cyclin-dependent kinase inhibitor CDKN1B (p27kip1) expression. *Proc Natl Acad Sci USA* 2005; **102**: 14028–14033.
8 Milne TA, Hughes CM, Lloyd R, Yang Z, Rozenblatt-Rosen O, Dou Y *et al.* Menin and MLL cooperatively regulate expression of cyclin-dependent kinase inhibitors. *Proc Natl Acad Sci USA* 2005; **102**: 749–754.
9 Krivtsov AV, Feng Z, Lemieux ME, Faber J, Vempati S, Sinha AU *et al.* H3K79 methylation profiles define murine and human MLL-AF4 leukemias. *Cancer Cell* 2008; **14**: 355–368.
10 Bursen A, Moritz S, Gaussmann A, Moritz S, Dingermann T, Marschalek R. Interaction of AF4 wild-type and AF4·MLL fusion protein with SIAH proteins: indication for t(4;11) pathobiology? *Oncogene* 2004; **23**: 6237–6249.
11 Gaussmann A, Wenger T, Eberle I, Bursen A, Bracharz S, Herr I *et al.* Combined effects of the two reciprocal t(4;11) fusion proteins MLL·AF4 and AF4-MLL confer resistance to apoptosis, cell cycling capacity and growth transformation. *Oncogene* 2007; **26**: 3352–3363.
12 Bursen A, Schwabe K, Rüster B, Henschler R, Ruthardt M, Dingermann T *et al.* The AF4-MLL fusion protein is capable of inducing ALL in mice without requirement of MLL-AF4. *Blood* 2010; **115**: 3570–3579.
13 Benedikt A, Baltruschat S, Scholz B, Bursen A, Arrey TN, Meyer B *et al.* The leukemogenic AF4-MLL fusion protein causes P-TEFb kinase activation and altered epigenetic signatures. *Leukemia* 2010; **25**: 135–144.
14 Hsieh JJD, Cheng EHY, Korsmeyer SJ. Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell* 2003a; **115**: 293–303.
15 Yokoyama A, Kitabayashi I, Ayton PM, Cleary ML, Ohki M. Leukemia proto-oncoprotein MLL is proteolytically processed into 2 fragments with opposite transcriptional properties. *Blood* 2002; **100**: 3710–3718.
16 Hsieh JJ, Ernst P, Erdjument-Bromage H, Tempst P, Korsmeyer SJ. Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. *Mol Cell Biol* 2003b; **23**: 186–194.
17 Doerks T, Copley RR, Schultz J, Ponting C, Bork P. Systematic identification of novel protein domain families associated with nuclear functions. *Genome Res* 2002; **12**: 47–56.
18 García-Alai M, Allen D, Joerger A, Bycroft M. The structure of the FYR domains of transforming growth factor beta regulator 1 (TBRG1). *Protein Science* 2010; **17**: 1432–1438.
19 Kohoutek J. P-TEFb- the final frontier. *Cell Div* 2009; **4**: 19.
20 Peterlin BM, Price DH. Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* 2006; **23**: 297–305.
21 Marshall NF, Price DH. Purification of P-TEFb, a transcription factor required for the transition into productive elongation. *J Biol Chem* 1995; **270**: 12335–12338.
22 Bier C, Knauer SK, Klapthor A, Schweitzer A, Rekik A, Krämer OH *et al.* Cell-based analysis of structure-function activity of Threonine Aspartase 1. *J Biol Chem* 2010, e-pub ahead of print.
23 Knauer SK, Stauber RH. Development of an autofluorescent translocation biosensor system to investigate protein-protein interactions in living cells. *Anal Chem* 2005; **77**: 4815–4820.
24 Meyer C, Kowarz E, Schneider B, Oehm C, Klingebiel T, Dingermann T *et al.* Genomic DNA of leukemic patients: target for clinical diagnosis of MLL rearrangements. *Biotechnol J* 2006; **1**: 656–663.
25 Bitoun E, Oliver PL, Davies KE. The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. *Hum Mol Genet* 2007; **16**: 92–106.

Supplementary Information accompanies the paper on the Leukemia website (http://www.nature.com/leu)