Functional characterisation of different MLL fusion proteins by using inducible Sleeping Beauty vectors

K. Wächter 1, E. Kowarz 1, R. Marschalek *

Institute of Pharm. Biology/DCAL, Goethe-University, Frankfurt/Main, Germany

ARTICLE INFO

Article history:
Received 9 May 2014
Received in revised form 18 June 2014
Accepted 24 June 2014

Keywords:
MLL fusion proteins
NEBL
LASP1
MAML2
SMAP1

ABSTRACT

Our focus is the identification, characterisation and functional analysis of different MLL fusions. In general, MLL fusion proteins are encoded by large cDNA cassettes that are difficult to transduce into haematopoietic stem cells. This is due to the size limitations of the packaging process of those vector-encoded RNAs into retro- or lentiviral particles. Here, we present our efforts in establishing a universal vector system to analyse different MLL fusions. The universal cloning system was embedded into the backbone of the Sleeping Beauty transposable element. This transposon has no size limitation and displays no integration preference, thereby avoiding the integration into active genes or their promoter regions. We utilised this novel system to test different MLL fusion alleles (MLL-NEBL, NEBL-MLL, MLL-LASP1, LASP1-MLL, MLL-MAML2, MAML2-MLL, MLL-SMAP1 and SMAP1-MLL) in appropriate cell lines. Stable cell lines were analysed for their growth behaviour, focus formation and colony formation capacity.

Introduction

Chromosomal translocations of the human MLL gene are associated with different types of human leukemias (ALL: acute lymphoblastic leukemia; AML: acute myeloid leukemia; MLL: mixed-lineage leukemia), and so far about 80 MLL fusions have been described in the literature [1,2]. Despite the chromosomal translocation t(1;11)(q21;q23), resulting in the expression of the MLL-AF1q fusion protein and displaying a very good clinical prognosis (EFS = 0.92), all other MLL fusions give a dismal prognosis (EFS = 0.11–0.50) [3]. This indicates that these MLL fusion proteins are probably highly potent in initiating and maintaining the corresponding leukemia disease phenotype.

By contrast, our understanding about the pathological function(s) of MLL fusion proteins is based on experiments performed with a handful of these chimeric proteins. Only the MLL fusions deriving from t(4;11), t(9;11), t(10;11), t(11;19) and few others have been functionally investigated. The fusion proteins – deriving from the aforementioned MLL translocations – trigger very similar downstream events, namely the ability to bind and activate endogenous AF4/AF5 complexes (also named superelongation complexes), or the direct activation of RNA Polymerase II, as does the AF4-MLL fusion protein [4]. This causes an increase in mRNA levels and ectopic/extended histone methylation signatures (H3K4 and H3K79) [5–7]. Furthermore, these 4 MLL gene fusions account for more than 90% of ALL cases (infant, childhood and adult) and about 50% of AML cases (infant, childhood and adult). However, for the vast majority of MLL fusions (n = 76), the underlying pathological mechanism(s) are unclear. Fusion partner genes encode proteins that are barely investigated or are even of cytosolic origin. Thus, potential oncogenic functions of these proteins – when fused to MLL and translocated into the nucleus – must be deemed artificial.

One general obstacle when dealing with MLL fusions is the large size of their open reading frames, which is contrasted with the limitations for packaging these constructs into retro- or lentiviral backbones. This hampers the in vivo analysis for most known MLL fusions. On the other hand, in vitro systems allow for important conclusions to be drawn when used appropriately and in a systematic fashion. Therefore, we decided to establish a universal system that allows for rapid and functional characterisation of different MLL fusions in order to screen for interesting candidates.

* Corresponding author. Address: Institute of Pharmaceutical Biology/Diagnostic Center of Acute Leukemia DCAL, University of Frankfurt, Marie-Curie Str. 9, D-60439 Frankfurt/Main, Germany. Tel.: +49 69 798 29647; fax: +49 69 798 29662.
E-mail address: Rolf.Marschalek@em.uni-frankfurt.de (R. Marschalek).

1 These authors contributed equally to this work.
that can subsequently be used for further investigation. For these purposes, we used the Sleeping Beauty transposon system as vector backbone which can be stably integrated in low copy numbers into mammalian cell lines by co-transfection with Sleeping Beauty transposase. Stable cell lines can be rapidly selected and functional analyses performed in parallel to investigate the functions of different MLL fusions under identical conditions.

Here, we present the results of our study in which we analysed 4 different MLL translocations. The results of our study suggest that reciprocal MLL fusions should be put into the focus of scientific research.

Materials and methods

Construction of a universal vector system using the Sleeping Beauty backbone

As summarised in Fig. 1A, universal vector backbones were cloned. For direct MLL fusions, we designed a vector that contains a constitutive promoter followed by SfiI(A)-NcoI restriction endonuclease sites. Following these 2 restriction sites, a bona fide AUG start codon and a Flag-Tag, fused in frame to MLL exons 1–9, was added to the vector. The last exon, MLL exon 9, is flanked by a short portion of the germline MLL intron 9 sequence that ends with a multiple cloning site consisting of XhoI, BstBI and SfiI(B) restriction sites. Corresponding cDNA fragments of partner genes (gene of interest, GOI) were cloned with a region of their germline intron with all exons (cDNA) up to and including their natural stop codon. Both primers used to amplify these partner cDNAs contained restriction recognition sites for XhoI and BstBI, respectively. After digests with these enzymes, the hybrid gDNA/cDNA fragments were cloned into the XhoI and BstBI restriction sites of the universal vector. Since the complete cassette is flanked by two different SfiI sites (A and B), the cassette (MLL-X) can be mobilised without promoter and polyA signal sequence.

Similarly, a vector for reciprocal MLL fusions was designed. It contains also a constitutive promoter followed by SfiI(A)-NcoI restriction endonuclease sites. Behind this multiple cloning site a short portion of the germline MLL intron 13 is fused to MLL exons 14–37, with no stop codon. Instead, the MLL sequence is followed by a Flag-tag and a stop codon prior to the BstBI-SfiI(B) restriction sites. Partner cDNAs (GOI) were amplified with oligonucleotides that exhibit a short track of 5'-NTR nucleotides and the NcoI recognition site in combination with an oligonucleotide, comprising the sequence of the last exon, a short portion of the following intron and the AvrII recognition site. The amplified sequences were cloned upstream of the MLL sequence by using Ncol-AvrII digests. Here, the complete cassette (X-MLL) could also be mobilised by an SfiI digest.

The original Sleeping Beauty (SB) vector backbone [8] was modified with a TREG-promoter, two consecutive SfiI sites (A and B) followed by an poly A signal. This inducible cassette was separated from the PGK promoter which drives translation of a polycistronic mRNA cassette that encodes a fluorescent protein, the reverse Tet repressor protein (M2S rtTA) and either a hygromycin (MLL-X) or puromycin resistance protein (X-MLL). All 3 reading frames were separated by T2A peptide sequences.

These 2 different SB vector backbones were used to clone the different MLL fusion genes with SfiI(A/B) cassettes. All direct MLL fusions (MLL-LASP1, MLL-NEBL, MLL-MAML2 and MLL-SMAP1) were cloned into the GFP-expressing vector (TCZH), while all reciprocal MLL fusions (LASP1-MLL, NEBL-MLL, MAML2-MLL and SMAP1-MLL) were cloned into the RFP-expressing vector (TCTP). This allowed us to study all these constructs alone or the combination of both.

A human mutant RAS protein (G12V) served as positive control. The RAS gene was cloned into the TCZH vector. Additionally, we cloned MLL exons 14–37 without any fused partner gene sequence into the TCTP vector. The MLL protein is naturally
transcribed and translated by a transcript that starts at the MLL intron 11/MLL exon 12 borderline, and exhibits a \textit{bona fide} AUG start codon localised within MLL exon 18 [8]. The resulting MLL protein (227 KDa) has been demonstrated to be expressed at the protein level and is properly processed by Taspase1.

Transfection of cells

In order to establish stable cell lines from all aforementioned constructs (n = 8), all SB vector constructs were co-transfected with low amounts (50 ng) of the SB transposase vector SB100X [19]. PuGene\textsuperscript{-}transfections into MEF cells were carried out as recommended by the manufacturer. Twenty-four hours after transfection, cells were subjected either to Hygromycin (300 μg/ml), Puromycin (1 mg/ml) or both. Selection was carried out for 3–7 days and terminated when virtually all cells were emitting the expected green or red light derived from their corresponding reporter genes. Cells were cultivated for 4 weeks without selection medium and the stability of the transfectected vector constructs were monitored. In all cases, the transfected constructs remained stable, expressing either GFP or RFP as a surrogate marker.

\textit{RT and QRT-PCR experiments}

The transgenes were induced by applying 1 μg/ml Doxycycline to the cell culture medium. After 48 h, total RNA was isolated and cDNA syntheses were performed. The following oligonucleotides were used for RT-PCR analyses: MLLex9-F: 5'-GCTTGAGCCAGCATCTA GAG-3', MLLex14-R: 5'-ATGACACAGTCGAATATATATGAG-3', NEBleX3-F: 5'-GCCGGACAACTGAGTTTTTCTG-3', NMLleX5-F: 5'-GCCAGTGGCTGACATCTGCTG-3', LASP1ex-F: 5'-GCCAGTTGGCCCA GTCATAT-3', LASP1ex7-F: 5'-TATGACAGTGAGAAATCATATGAC-3', MAML2ex1-F: 5'-ACC CCAAGCCCCACACAGACTAT-3', MAML2ex2-R: 5'-GTCTGCTTCTCTCCATCAAT TGC-3', SMAP1ex6-F2: 5'-CCAGAAGACCGAAGACACCATCTA-3', SMAP1ex7-R: 5'-TTTTTGCCTCAGGTTCTGGTCCTTCTC-3. Resulting PCR amplicons were subjected to DNA sequencing analysis to validate all splice events were correctly performed. For murine \textit{Hoxa} gene transcript analyses, we used the following oligonucleotides: Hoxa7-F: 5'-AGCCTTTTTTATGAAATATACG-3', Hoxa7-R: 5'-GCCCAACCCCCACAGACTAT-3', Hoxa9-F: 5'-CAGGAGTGTCCGTGACGATGCT GAAG-3', Hoxa9-R: 5'-TTCAGCAGGCGACCCAAAACA-3', Hoxa10-F: 5'-CAGGAGGCTTCGGTCTTCTT-3 and Hoxa10-R: 5'-TGGTCCGCCAGCATGAGAAG-3.

\textbf{Fig. 1.} Creation of stable cell lines from all aforementioned constructs (n = 8). The resulting MLL protein (227 KDa) has been demonstrated to be expressed at the protein level and is properly processed by Taspase1.

\textbf{Results}

Transfection was carried out in MEF cells which were subsequently selected in antibiotic (Pyromycin, Hygromycin or both) to create stable cell lines.

First, integration efficiency was validated by genomic qPCR experiments, which were normalised to a single copy locus. Integration efficiency was in the range of 1–10 transgene copies per cell (Fig. 1B). Therefore, we concluded that we can compare the data of each set of MLL fusions, because we had no strong discrepancy in copy numbers, as they deviate at a maximum of ±2 copies.

Splice events of all our constructs were investigated by RT-PCR experiments followed by DNA sequencing. After induction of transgenes for 48 h, we performed RT-PCR experiments to validate correct splicing. PCR amplicons were subjected to DNA sequencing analysis in order to confirm a correct fusion of the open reading frames. As shown in Fig. 1C, the introns in all 8 constructs were spliced out and all reading frames were correctly fused, regardless whether the constructs were tested from single- or co-transfected cells.

\textit{MLL fusions are characterised by their unique ability to enhance the transcription of \textit{HOXA} genes in human cells. As shown in Fig. 2, we were not able to see any clear changes of \textit{Hoxa} clusters in MEF cells when compared to the untransfected cells. Even Q-PCR experiments could not reveal any clear changes (data not shown). From these data we concluded that an ectopic \textit{Hoxa} gene activation could not be induced, either by the direct nor by the reciprocal MLL fusions tested here. This may indicate that MLL fusion protein – not docking to the endogenous AF4/AF5 complex (superelongation complex) – may not be able to change \textit{HOXA} gene patterns.

Next, we investigated the growth behaviour of single and double transfected cell lines. For this purpose we used the CCK-8 assay,
a non-toxic variant of the well-known MTT assay. These assays are used as a surrogate marker for cell growth and viability. Fig. 3A summarises our results obtained with the stable MEF cell lines. As shown in the first panel, RAS* expressing cells displayed a slight growth advantage after 4 days. Growth/viability was also enhanced when MLL* was expressed. The next two panels summarise the data obtained with single transfected cells (MLL-X or X-MLL) and the last panel displays the effect in the presence of both fusion proteins. Except for MLL-MAML2, the direct MLL fusions (MLL-NEBL, MLL-LASP1 and MLL-SMAP1) displayed a better cell growth when compared to the control. However, the observed growth rates were lower than in RAS* or MLL* expressing cells. With the exception of SMAP1-MLL, the reciprocal MLL fusions displayed a strong growth enhancement, with higher values than the positive controls. This effect diminished, however, when both fusion proteins are expressed concomitantly in the same cell. Only the co-LASP1 and co-MAML2 cell still displayed a slightly enhanced growth behaviour.

To evaluate anchorage-independent cell growth, colony formation assays were carried out. Cell lines were seeded in softagar and dishes were stained after day 21. The pictures obtained from these dishes were subsequently analysed by the ImageJ program. Colony formation was evident in all dishes in varying sizes. Therefore, we focussed only on larger colonies (see Fig. 3B). In addition we defined a discriminator (mean between negative and positive controls): a number below that borderline was defined as negative, while a number above this borderline was defined as a positive result. Based on these assumptions, we obtained a positive score in this assay only for RAS*, MLL*, SMAP1-MLL, NEBL-MLL, LASP1-MLL and the co-LASP1 cells. All others were scored negatively or were not significantly classifiable. Cells expressing MLL-MAML2, MAML2-MLL or both constructs were consistently negative in this assay, with scores similar to the untransfected or mock-transfected cells.

Focus formation experiments characteristically give the first level of evidence for oncogenic behaviour of mutated or variant proteins. Cell growth of non-malignant cells is inhibited in these assays and enter quiescence after confluency is reached, while oncogenic cells grow into the 3rd dimension and form foci. Here, we investigated this phenomenon for our cell lines in 60 mm dishes and monitored their loss-of-contact inhibition after 8 and 15 days post-induction with Doxycycline.

We first observed foci after 8 days with cells expressing the RAS* protein, while none of the 4 tested direct MLL fusions nor the vector control showed any effect (Fig. 4A, left panel). By contrast, when we analysed the 4 different reciprocal MLL fusions, we observed foci after 8 days in NEBL-MLL cells and MAML2-MLL cells (Fig. 4B, left panel). In cells co-expressing both reciprocal MLL fusions, only the co-MLL2 cells displayed focus formation (Fig. 4C, left panel). After 15 days, the picture became more clear. Besides RAS*, only MLL-SMAP1 was able to induce foci formation, while none of the other MLL-X constructs displayed such a phenotype (Fig. 4A, right panel). Oppositely, the reciprocal NEBL-MLL, LASP1-MLL, MAML2-MLL and MLL* displayed focus formation, while SMAP1-MLL did not (Fig. 4B, right panel). In addition, all co-transfected cells displayed focus formation activity after day 15 (Fig. 4C, right panel). This is not surprising because at least one of each tested construct was scoring positive. It also indicated that the observed “loss of contact inhibition” is a dominant phenotype. To our surprise, expression of the MLL* protein also caused focus formation, indicating that the N-terminal truncated MLL protein by itself may exhibit/activate oncogenic functions.

**Discussion**

In the last decade, our laboratory has identified a large series of direct and reciprocal MLL fusions. So far, 79 direct MLL fusions...
(MLL-X) and more than 120 reciprocal MLL fusions (X-MLL) have been described by our group [2]. Most of these MLL fusions were never tested in functional experiments.

Here, we present data for 4 MLL fusions that were recurrently identified in leukemia patients. We have chosen **NEBL** and **LASP1** because both belong to the same Nbeulette gene family and both were found to be fused in a reciprocal fashion with the **MLL** gene in AML patients [11,12]. Both proteins are cytosolic LIM domain proteins with a length of 270 and 261 amino acids, respectively (the LIM domain is localised at the N-terminus, while the C-terminus exhibits a SH3 domain), and known to regulate/modulate the actin cytoskeleton via binding to SRC kinase. In addition we were interested in **MAML2** that was diagnosed as a MLL fusion partner gene in therapy-related, secondary AML and T-ALL patients [13,14]. MAML2 belongs to a family of coactivators for the NOTCH signalling pathway. The protein has a length of 1,153 amino acids.

**Fig. 4.** Sizes of all pictures are identical and indicated in the first row of pictures (scale for 200 μm). (A) Focus formation of direct MLL fusions (MLL-X) at day 8 and 15. Top: positive RAS control. Bottom: mock- and untransfected controls. (B) Focus formation of reciprocal MLL fusions (X-MLL) at day 8 and 15. Top: positive MLL control. Bottom: mock- and untransfected controls. (C) Focus formation of co-transfected MLL fusions (MLL-X and X-MLL) at day 8 and 15.
and the N-terminal portion of MAML2 interacts with the Ankyrin repeats of the -Secretase cleaved NOTCH receptor (NICD). Within the nucleus, MAML2 forms a complex with NICD and other nuclear coactivators to mediate transcription of NOTCH target genes. Of interest, fusions of MAML2 have already been described for salivary gland carcinomas, or more generally in mucous/epithelial carcinomas (MEC), where MAML2 is reciprocally fused with MECT1 (Mucous/epithelial Carcinoma Translocated 1) [15]. Finally, we investigated SMAP1 as a prototype for the many fusion partners found in AML patients [16]. SMAP1 has a length of 468 amino acids and encodes an in-frame fusion protein or in other words, a “incompatible gene fusions” (~80%), Of interest, all these “incompatible gene fusions” are able to express the MLL’ protein. This is due to the fact that the MLL gene contain a cryptic promoter at the MLL intron 11/MLL exon 12 borderline that is enhanced by MLL intron 12 sequences [9]. This results in the production of the 227 kDa MLL’ protein that exhibits about half of the BD domain, the FYRN interaction domain, the TAD domain, the FYRN interaction domain and the SET domain (H3K4 HMT).

Here, we have successfully generated a universal vector system to analyse direct and reciprocal MLL fusion genes. This vector system has several benefits over conventional systems, because expression of transgenes can be induced by Doxycycline, while a constitutive expression of fluorescent proteins and selection marker allows to trace the transfected cells. Selection of stably integrated vectors took between 3 and 7 days only. By using these stable cell lines, we attempted to dissect certain functions of fusion proteins that are derived from their transgenes: (1) cell growth/viability; (2) Hoxa gene expression; (3) colony formation in soft agar and (4) focus formation capacity. Investigating cell growth/viability is difficult due to issues in differentiating between proliferation and apoptosis. However, these experiments overcome some of these issues and reveals that the expression of direct MLL fusions seems to be dominant over growth promoting activities exerted by the reciprocal MLL fusions. This might be explained by recent findings, where the steady-state amount of direct MLL fusions is responsible for either growth promotion or inhibition [21]. The picture changed when we analysed colony formation and focus formation capacity. LASP1-MLL seems to exert a dominant colony formation phenotype that could not be suppressed by MLL-LASP1. MLL-SMAP1, NEBL-MLL and MLL’ scored all weakly positive, but the two MAML2 constructs remained negative in this assay.

Results obtained in focus formation experiments were the strongest indicator for oncogenic capacity. Here, MLL-SMAP1, NEBL-MLL, LASP1-MLL and MAML2-MLL displayed oncogenic activities. Again, the MLL’ protein was sufficient to cause focus formation, however, with a slower kinetic than the other tested constructs. Fortunately, a potent and MLL-specific SET domain inhibitor has recently been developed by the group of Yali Dou [22]. Thus, even if this mutant form of MLL displays oncogenic features, future experiments may decipher whether an already existing drug may block these oncogenic features.

In summary, we have successfully established a novel tool to rapidly screen different MLL fusions in functional experiments, allowing the prediction of oncogenic behaviour. This will assist the investigation of other MLL fusions, and aid the identification of other constructs and fusion genes in an in vivo model setting to monitor leukemogenic potential.

Conflict of Interest

The authors declare that there are no conflicts of interest.

Acknowledgements

This study was supported by grant DKS 2011.09 from the German Children Cancer Aid to R.M., and by research grants Ma1876/10-1, Ma1876/11-1 from the DFG to R.M. R.M. is PI within the CEF on Macromolecular Complexes funded by DFG grant EXC 115.

References


