Function and regulation of the histone methyltransferase Dot1
Frederiks, F.

Citation for published version (APA):
CHAPTER 6

Two Dot1 isoforms in *Saccharomyces cerevisiae* as a result of leaky scanning by the ribosome

Floor Frederiks¹, Guus J.J.E. Heynen¹, Sjoerd J. van Deventer², Hans Janssen², and Fred van Leeuwen¹

¹ Division of Gene Regulation, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam
² Division of Cell Biology II, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam

ABSTRACT

Dot1 is a conserved histone methyltransferase that methylates histone H3 on lysine 79. We previously observed that in *Saccharomyces cerevisiae*, a single *DOT1* gene encodes two Dot1 protein species. Here we show that the relative abundance of the two isoforms changed under nutrient-limiting conditions. A mutagenesis approach showed that the two Dot1 isoforms are produced from two alternative translation start sites as a result of leaky scanning by the ribosome. The leaky scanning was not affected by the 5’ or 3’ untranslated regions of *DOT1*, indicating that translation initiation is determined by the *DOT1* coding sequence. Construction of yeast strains expressing either one of the isoforms showed that both were sufficient for Dot1’s role in global H3K79 methylation and telomeric gene silencing. However, the absence of the long isoform of Dot1 altered the resistance of yeast cells to the chitin-binding drug Calcofluor White, suggesting that the two Dot1 isoforms have a differential function in cell wall biogenesis.
Chapter 6

INTRODUCTION

Dot1 is a histone methyltransferase that was originally identified in *Saccharomyces cerevisiae*. The enzyme, which is conserved from yeast to humans, catalyzes mono-, di and trimethylation of lysine 79 on histone H3 (H3K79), a residue located on the nucleosome core (2-5). In yeast Dot1 methylates ~90% of histone H3 and is involved in gene silencing, activation of the DNA damage response, and the pachytene checkpoint in meiosis (2;6-8). In addition, genetic interactions have been reported between *DOT1* and different genes involved in cell wall biogenesis, suggesting a function for Dot1 in this process (9;10). In mammals Dot1 plays a role in Ras-induced gene silencing and aldosterone-induced gene repression as well as in inappropriate gene activation in certain types of leukemia (11-14).

Our antibodies directed against yeast Dot1 detect two protein species on immunoblots, even though the protein is encoded by a single gene (2). Several mechanisms can result in the formation of two different protein species from one gene. First, full-length proteins can be processed by proteases, resulting in the generation of N- or C-terminally truncated forms. Second, several post-translational modifications such as phosphorylation or ubiquitination can result in an altered mobility of proteins on immunoblots. Finally, different protein isoforms can be generated by alternative translation, which can be caused by variations in the mRNA sequence (for example due to alternative mRNA splicing or alternative transcription start site selection) or by the usage of multiple translation start sites. Translation from multiple start codons can occur when the mRNA contains an internal ribosomal entry site (IRES), or when the ribosome scanning over the mRNA molecule skips the first start codon and subsequently initiates translation from a downstream alternative start codon, a phenomenon called 'leaky scanning' (15). Alternatively, the ribosome can bind to the mRNA on the 5’ side of the start codon via the normal mechanism, but then jumps over the first start codon to initiate transcription from a downstream start codon, a process called ribosome shunting (16). Translation start site selection is a highly regulated process that is critical for differential gene regulation in eukaryotes (17). Mutations that affect translation initiation have been associated with a range of genetic diseases in humans (18). However, there are still many open questions regarding the molecular mechanisms of start codon selection (19).

In this study we determined the mechanism by which the two Dot1 isoforms are generated. We show by protein tagging, protein truncation, and mutational analysis that they are the result of leaky scanning by the ribosome.
Two Dot1 isoforms as a result of leaky scanning

This enabled the construction of yeast strains expressing either one of the isoforms. The two Dot1 isoforms were found to have indistinguishable functions in global methylation and gene silencing. However, they showed distinct functions in resistance to the chitin-binding molecule Calcofluor White, suggesting that they play different roles in cell wall biogenesis. These results suggest that leaky scanning of the DOT1 mRNA is affected by sequences downstream of the start codon, whereas the efficiency of translation initiation is usually thought to depend mainly on the 5’ context of the start codon (15). The discovery of unconventional leaky scanning events like the one reported here will be instrumental to unravel the complex mechanisms that regulate translation initiation in yeast and higher eukaryotes.

RESULTS

Two isoforms of Dot1 in S. cerevisiae

An antibody directed against the C-terminal part of the yeast Dot1 protein (Fig. 1A) recognizes two discrete proteins on immunoblots (Fig. 1B). The slow migrating protein is usually more abundant than the fast migrating protein. The goal of this study was to determine the identity, regulation, and the function of the two protein isoforms of Dot1. Since Dot1 is involved in activation of the DNA damage checkpoint upon exposure to UV light (7), we tested whether the relative abundance of the two proteins was affected by DNA damage. Thirty minutes after UV irradiation, there was no detectable difference in relative abundance of the two isoforms, while phosphorylation of the central checkpoint kinase Rad53 confirmed activation of the DNA damage checkpoint (Fig. 1C). To investigate whether the production of the two isoforms was regulated by growth conditions, exponentially growing cells were compared to cells arrested by nutrient depletion. After starving the cells for two days, the abundance of the slow-migrating Dot1 protein was reduced while the fast migrating protein was unaffected (Fig. 1D). This change in the ratio of the Dot1 isoforms probably does not reflect a difference in protein stability, since the two isoforms showed approximately the same rate of degradation when protein synthesis was blocked by addition of cycloheximide (Fig. 1E). These results show that the relative abundance of the two Dot1 isoforms is dependent on growth conditions.
A long and a short Dot1 protein

Our aim was to express the two Dot1 isoforms separately in yeast, to determine whether the existence of two isoforms and the differential expression is relevant for the function of Dot1. Therefore, first the mechanism of production of the two isoforms was determined. Previous results had shown that the addition of a single N-terminal FLAG tag to the Dot1 protein resulted in a shift in mobility as well as expression of only one Dot1 isoform (Fig. 1F, [20]). In contrast, introduction of a
Two Dot1 isoforms as a result of leaky scanning

TAP tag at the C-terminus of Dot1 (and the concomitant replacement of the DOT1 3’UTR with that of the ADH1 gene) did not affect the expression of the two isoforms (Fig. 1G). However, an internal deletion of amino acids 4-20 (Dot1Δ4-20) resulted in the expression of only one Dot1 isoform (Fig. 1G). These results indicated that the protein sequence or the coding sequence of the first twenty amino acids was involved in the expression of the two isoforms. These results also suggested that the production of two isoforms was not caused by a post-translational modification of Dot1 because the double band was disrupted by two non-overlapping mutations: N-terminal tagging (which affects the N-terminal residue of Dot1; Fig. 1F) as well as an internal deletion (which leaves the first three N-terminal residues of Dot1 intact; Fig. 1G).

The Dot1Δ4-20 protein migrated at approximately the same size as the smaller isoform of Dot1 (Fig. 1G), suggesting that the slower migrating Dot1 band might represent the full length protein and the fast migrating band a Dot1 protein that is missing part of the N-terminus. A smaller protein can be generated by degradation or protein processing of a larger protein or by alternative translation. Inspection of the coding sequence of Dot1 showed that it contains a second in-frame AUG codon (AUG17), sixteen codons downstream of the first AUG (AUG1) (Fig. 2A). Alternative translation initiation at this second start codon would give rise to a Dot1 protein sixteen amino acids shorter than the full length protein. To distinguish between protein processing and alternative translation as the mechanisms responsible for Dot1 isoform production, point mutations were introduced in the DOT1 gene expressed from a single-copy plasmid. The wild-type and mutant plasmids were then introduced in dot1Δ strains. First two frameshift mutants were generated by introduction of an extra base between the first and the second AUG (FSM1 and FSM2, Fig. 2A). Both of these insertions result in a stop codon at amino acid position 10, which prevents the generation of full length Dot1. We reasoned that if the smaller protein is generated by processing of the full length protein, the frameshift mutations should destroy production of both full length and the smaller Dot1 protein. In contrast, if it is generated by alternative translation frameshift mutations should not affect production of the smaller Dot1 isoform. As shown in Fig. 2B, when the frameshift mutants were expressed in yeast, the short isoform was still present whereas the larger isoform was absent. These experiments rule out the possibility that the smaller Dot1 protein is a processed form of the full length protein.
Figure 2. The two Dot1 isoforms are generated by leaky scanning by the ribosome, which is not dependent on the 5' context of the start codon. (A) Nucleotide and amino acid sequence of the N-terminal part of Dot1. Black boxes show the two potential start codons. The different mutants used in this study are shown underlined. Briefly, in the M1T mutant AUG1 is changed to ACU, in the I7M mutant AU7 is changed to AUG and in the M17L mutant AUG17 is changed to UUG. Frameshift mutant 1 (FSM1) has an insertion of an A in codon 9 and FSM2 has an insertion of a C in codon 7. Sequence details of the frameshift mutants are shown in Table 1. (B) Immunoblot analysis of two frameshift mutants (FSM) of Dot in a dot1Δ strain (UCC7183) transformed with an empty plasmid (pRS315), a single copy plasmid expressing DOT1 (pDOT1), or a single copy plasmid expressing one of the two frameshift mutants of DOT1 (pGH18 or pGH21). Pgk1 was used as a loading control. (C) Immunoblot analysis of Dot1 expression in dot1Δ cells (UCC7183) containing an empty plasmid, wild type Dot1, Dot1-M1T or Dot1-M17L (pRS315, pDOT1, pGH01 and pGH02 respectively), showing that mutation of either of the putative start codons leads to expression of a single Dot1 isoform. Pgk1 was used as a loading control. (D) Comparison of Dot1 expression in a wild type strain (PDOT1–DOT1, UCC7164) and a dot1Δ strain (UCC7183) containing an empty vector (pTCG) or a plasmid overexpressing DOT1 (and 54 bp of its 5' UTR) from the GAL1 promoter (PGAL1–DOT1, pFvL18). A short and a long exposure of the same blot are shown. (E) Nucleotide sequence around the DOT1 start codon in strains NKI3049 (endogenous DOT1 5' UTR) and NKI3066 (non-yeast vector 5' UTR). Thirty nucleotides of 5' UTR sequence are shown. In strain NKI3049 (DOT1) 62 bp of endogenous DOT1 5' UTR remains upstream of the first start codon; in strain NKI3066 (vector) the DOT1 5' UTR has been removed by insertion of the GAL1 promoter and 45 bp of vector sequence of the integration construct pYM-N22. (F) Immunoblot comparison of Dot1 expression in the two strains described in panel E. (G) Immunoblot analysis of Dot1 expression in a dot1Δ strain (UCC7183) transformed with wild-type Dot1 (pDOT1) or Dot1-I7M (pFf018).
**Two Dot1 isoforms as a result of alternative translation**

To confirm that the two Dot1 isoforms are generated as different translation products from the two start codons present in the 5’ part of the DOT1 coding sequence, the first or the second AUG was mutated. Mutation of AUG1 to ACU, changing the initiator Met codon to a non-initiator Thr codon (Dot1-M1T) resulted in the expression of a single protein that had the same size as the small isoform (Fig. 2C), confirming that translation can initiate at a downstream start codon. Mutation of AUG17 to UUA, changing the initiator Met codon to a Leu codon (Dot1-M17L) generated a single protein with the same size as the large isoform (Fig. 2C), showing that translation starts at AUG1 as well as AUG17. These results suggest that the two isoforms are indeed two translation products starting at M1 or M17, respectively.

Next, the cause of alternative translation was investigated. Two translation products can be generated from two different transcript molecules: one initiated upstream of the first AUG, and one initiated in between the first and second AUG. Alternatively, they can be generated from a single long transcript by alternative start codon usage. To distinguish between these two possibilities, the endogenous DOT1 promoter was replaced by the GAL1 promoter, which is induced when cells are grown in media containing galactose. In this case 54 bp of the DOT1 5’ UTR was retained to avoid changes in the context of the start codon. When Dot1 was expressed from the GAL1 promoter, the two isoforms were generated in the same ratio as found when the gene was expressed from its endogenous promoter (Fig. 2D). Since it is unlikely that two unrelated promoters use identical transcription-start sites (i.e. upstream of the first AUG and in between the first and second AUG) and with identical relative frequencies, this result suggests that the two isoforms are not generated by alternative transcription start site selection.

Alternative translation can be caused by ribosome shunting, by leaky scanning of the ribosome, or by an internal ribosomal entry site (IRES). In ribosome shunting, which has mainly been described for viruses, the ribosome binds to the mRNA on the 5’ side of the start codon via the normal mechanism, but then “jumps” over the first start codon to initiate transcription from a downstream start codon (16). In leaky scanning, the ribosome does scan over the first start codon but does not recognize it in a significant fraction of initiation events (15). An IRES is an mRNA sequence that has special structural properties allowing the recruitment of ribosomes without using the standard mechanism of translation initiation (21).
To address the mechanism of alternative translation, the DOT1 promoter and the complete 5’ UTR were replaced by an exogenous GAL1 promoter and 45 bp of non-yeast vector sequence immediately upstream of the open reading frame of DOT1 (Fig. 2E). As a control, the DOT1 promoter was replaced by the GAL1 promoter cassette such that 62 bp of the endogenous DOT1 5’ UTR was retained (Fig. 2E). When transcription from the GAL1 promoter was induced, the two isoforms were generated in approximately the same ratio in both strains (Fig. 2F), suggesting that the 5’ context of the start codon did not affect the initiation of translation. Therefore, ribosome shunting was not a likely cause of the alternative translation. We note that the overall Dot1 abundance was higher in the presence of the heterologous 5’UTR (Fig. 2F), indicating that the 5’UTR might affect mRNA stability or transcription.

Most IRES elements are over 100 base pairs long, even though the minimal IRES modules may be shorter (21). The few bases between the first and second start codon of Dot1 seem insufficient to constitute an IRES. Leaky scanning by the ribosome thus seemed to be the most plausible explanation for the generation of the two isoforms. To verify this, a new start codon was inserted in between AUG1 and AUG17 at position 7 by changing the Ile codon to an initiator Met codon (I7M, Fig. 2A). If an IRES element was the cause of the alternative translation, the Dot1-I7M mutation could have two different effects. This mutation might interfere with the activity of the IRES element, resulting in loss of expression of the shorter isoform. Alternatively, the mutation could leave initiation at AUG17 unaffected, which would lead to normal expression of the two isoforms. In contrast, if leaky scanning was the cause of the alternative translation, skipping of AUG1 would lead to translation initiation at the new start codon AUG7 and thereby prevent translation start at AUG17. When the Dot1-I7M mutant was expressed two isoforms were produced, but the smaller one had become slightly larger (Fig. 2G). Although the effect of the 10 amino acid size difference was small, we consistently detected a slower migration of the short isoform in the Dot1-I7M mutant in multiple independent experiments (data not shown). This result shows that translation initiated at the new start codon, indicating that the two isoforms of Dot1 result from leaky scanning by the ribosome.

More support for this model comes from the Dot1-M1T and Dot1-M17L mutants in which only one of the isoforms is expressed. In the wild-type DOT1 transcript skipping of the first AUG occurs in a fraction of the translation events. However, if the first AUG is mutated, every ribosome will start translation from AUG17, which is expected to lead to an increase in abundance of the short Dot1 protein. This is indeed what was observed in the Dot1-M1T mutant (Fig. 2C). This
Two Dot1 isoforms as a result of leaky scanning

was not the case in the frameshift mutants, in which AUG1 was still present (Fig. 2B). As expected, in the Dot1-M17L mutant the larger isoform was expressed at the same level as in wild type cells, indicating that mutation of the downstream AUG had no major effects on translation initiation at AUG1 (Fig. 2C). We note that no smaller isoforms of Dot1 are detected in this mutant suggesting that no proteins were produced by translation initiation at a downstream start codon such as AUG40.

Mechanism of leaky scanning

Little is known about factors that affect leaky scanning. The 5’ UTR of the DOT1 gene is unlikely to be involved in determining the efficiency of translation initiation because replacement of the complete DOT1 5’ UTR by an exogenous sequence did not significantly affect the rate of leaky scanning (Fig. 2F). The efficiency of translation initiation is thought to be mainly determined by the nucleotide sequence upstream of the start codon, as a consequence of the linear scanning of the ribosome along the mRNA molecule (15). The Kozak consensus sequence defines the optimal context for translation initiation in mammals (22;23), but the requirements in yeast seem less strict (24;25). Recently, a genome-wide computational study showed that there is a strong bias in the yeast genome for an A at position -3 and a C at position +5. Presence of these residues was also correlated with the strength of expression (26). According to these observations AUG1 of the DOT1 gene is not located in an optimal context, with a G at position -3 and a G at position +5 (Fig. 2E). In contrast, AUG17 does have an optimal sequence context. To test if the -3 and +5 positions determine the efficiency of translation initiation of AUG1 of DOT1, they were mutated to the optimal A-3 and C+5, both separately and in combination. However, in these mutants the two Dot1 isoforms were still produced with approximately the same ratio as in wild type cells (Fig. 3A), showing that the suboptimal proximal context of AUG1 is not the cause of leaky scanning.

Another possibility is that the structure of the Dot1 RNA molecule affects the efficiency of translation initiation. Although this structure has not been experimentally determined, the RNAfold algorithm (1) predicts that the region of the mRNA around AUG1 contains a stem loop, which might somehow impair the binding of initiation factors or the polymerase complex to this sequence (Fig. 3B). To test this idea, mutations were made that were predicted to disrupt the overall structure of this part of the DOT1 mRNA (Fig. 3B-C). Although these mutations slightly increased the expression of Dot1, none of the mutations affected the rate
of leaky scanning (Fig. 3A and 3D). This mutagenesis study suggests that inefficient initiation of AUG1 is not caused by sequences or RNA structures proximal to the initiator codon.

**Functions of the two Dot1 isoforms**

Having addressed the origin of the two Dot1 isoforms, we designed yeast strains that express both or either one of the two isoforms to determine what the function of the Dot1 N-terminus might be. It has been shown that alternative translation in yeast can lead to production of two protein isoforms with or without N-terminal targeting sequences, leading to differences in sub-cellular localization (27-29). To visualize the localization of the two isoforms, plasmids were constructed expressing wild type Dot1, Dot1-M1T or Dot1-M17L fused to a citrine fluorescent tag (a variant of YFP, 30). Since endogenous Dot1 expression is very low (8;31),
Two Dot1 isoforms as a result of leaky scanning

Figure 4. The N-terminus of Dot1 does not affect localization, H3K79 methylation, or telomeric silencing. (A) Analysis of the localization of wild type Dot1, Dot1-M1T and Dot1-M17L. Log-phase cells overexpressing one of these proteins fused to a fluorescent citrine tag were stained with ConA-Cy5 (cell wall) and Hoechst (DNA) and then imaged using a confocal microscope. For this experiment dot1Δ strain UCC7183 was transformed with pFvL221, pTW055, pGH024 and pGH025. (B) Whole-cell extracts from log-phase cultures were prepared from dot1Δ strain UCC7183 transformed with an empty plasmid or a single copy plasmid expressing Dot1, Dot1-M1T or Dot1-M17L (pRS315, pDOT1, pGH01 and pGH02, respectively). H3K79 methylation and total histone H3 were analyzed by immunoblot. The asterisk indicates a non-specific band recognized by the H3K79me2 antibody. (C) Telomeric silencing of a URA3 reporter gene close to telomere VIII was measured in the strains described in (B). Cells were plated in tenfold dilution series on selective media in the presence or absence of 1 g/l 5-fluoroorotic acid (FOA). Cells that silence URA3 are resistant to FOA; cells with loss of telomeric silencing express URA3 and are sensitive to FOA. (D) Analysis of telomeric silencing in a wild type strain (UCC7164) transformed with pTCG, pTCG-DOT1 (pFvL18), pTCG-dot1-M1T (pGH03) or pTCG-dot1-M17L (pGH04). The GAL1 promoter on the pTCG plasmid was induced by growth on selective media containing galactose (GAL) to overexpress (OE) the Dot1 proteins.
the fusion proteins were expressed from a \textit{GAL1} promoter. In addition, the yeast cell wall was visualized using ConA-Cy5 and DNA was stained with Hoechst dye. The wild type Dot1 protein as well as the two mutants showed a diffuse nuclear staining and seemed to be absent from the cytoplasm (Fig. 4A), suggesting that the N-terminus of Dot1 does not influence its nuclear localization.

Dot1 methylates H3K79 throughout the entire yeast genome and the global levels of this modification determine the strength of telomeric silencing \((2;3;20)\). To investigate whether the N-terminus affects the catalytic activity of Dot1 the levels of methylation of H3K79 were analyzed in cells expressing wild-type Dot1 or only one of the isoforms. Immunoblots using antibodies specific for mono-, di-, and trimethylated H3K79 showed no detectable difference in global methylation when one of the Dot1 isoforms was missing (Fig. 4B). In line with the fact that no effect on the global levels of H3K79 methylation was detected, either isoform alone was sufficient to complement the telomeric silencing defect in a \textit{dot1}Δ strain, when expressed from a single copy plasmid (Fig. 4C). Both isoforms were also as competent as wild-type Dot1 to disrupt telomeric silencing when overexpressed from the \textit{GAL1} promoter on a multi-copy plasmid (Fig. 4D). In addition, there was no difference in activation of the DNA damage response after UV irradiation (data not shown).

Genetic interactions of \textit{DOT1} indicate that Dot1 might have a function in cell wall biology \((9;10)\). To verify this notion the sensitivity to the antifungal agent Calcofluor White (CFW) was examined of yeast cells with or without Dot1. Altered sensitivity to CFW is an indication of a defect in chitin biosynthesis and/or cell wall biogenesis \((32;33)\). When yeast cells were grown on rich media containing galactose cells lacking Dot1 were more resistant to CFW (Fig. 5A), lending support to the idea that Dot1 is required for cell wall function. To examine whether the two Dot1 isoforms might have differential functions in cell wall biogenesis, we tested whether strains expressing either one of the isoforms showed altered sensitivity to CFW. Two strains were constructed in which the \textit{GAL1} promoter and a 3xHA tag were integrated either upstream of AUG1 of the endogenous \textit{DOT1} coding sequence \(P_{\text{GAL1}}-3\text{HA}-\text{DOT1}\) or immediately downstream of AUG17, thereby deleting the first seventeen amino acids \(P_{\text{GAL1}}-3\text{HA}-\text{DOT1}\Delta 1-17\). In both cases a single protein species was expressed (Fig. 5B). These promoter-integration strains were used instead of strains expressing Dot1 from a plasmid because CFW is unstable in synthetic media necessary to ensure propagation of the plasmids. The strain expressing only full length Dot1 from the \textit{GAL1} promoter was as sensitive to CFW as the wild-type strain. In contrast, the strain expressing only the short form of Dot1 was more resistant to CFW, just like a \textit{dot1}Δ strain.
Two Dot1 isoforms as a result of leaky scanning

(Fig. 5A). This result suggests that the N-terminus of Dot1 influences its cell wall-related function.

To further investigate the possible cell wall-related function of Dot1, the sensitivity of wild type or mutant cells to changes in osmolarity was tested by growing cells on media containing different concentrations of NaCl, sorbitol, glycerol or SDS. However, no differences were observed for any of the conditions (Fig. S1). To assess a possible function for Dot1 in cell wall biogenesis more directly, electron microscopy was used to visualize the yeast cell wall in wild type cells, dot1Δ cells, and cells expressing only full length Dot1 or only the short isoform from the GAL1 promoter. As a control, kre1Δ cells were used, which have been reported to lack the outer layer of the cell wall (34), as is also clear from Fig. 5C. Whereas the cell wall of wild type cells and dot1Δ cells has a finely delineated

![Figure 5. Dot1 is involved in cell wall function.](image)

(A) Sensitivity to Calcofluor White (CFW) of the following strains: wild type (UCC7164), dot1Δ (UCC7183), P_GAL1-3xHA-DOT1 (NKI1059) and P_GAL1-3xHA-DOT1Δ1-17 (NKI1060). Cells were plated in tenfold dilutions on rich media containing 2% galactose in the presence or absence of 20 μg/ml CFW. The image is representative of four independent experiments. (B) Immunoblot analysis of Dot1 expression in the strains described in panel A. A long and short exposure is shown to indicate the expression relative to endogenous Dot1. Pgk1 was used as a loading control. (C) The cell wall of the strains described in panel A was imaged using electron microscopy. A kre1Δ strain (YSC-1122) was used as a control.
outer layer, this layer was somewhat less structured and more rough in texture in cells overexpressing only full length Dot1, and especially in cells overexpressing only the short Dot1 isoform (Fig. 5C). This result supports the notion that Dot1 affects cell wall metabolism. However, this function of Dot1 is found only when Dot1 is overexpressed and is not differentially affected by the two isoforms.

DISCUSSION

Yeast cells express two isoforms of the H3K79 methyltransferase Dot1. Using mutational analysis we determined that alternative translation by leaky scanning results in the expression of a full length protein and a shorter Dot1 protein lacking the first sixteen amino acids. Starved cells lose expression of full length Dot1 and maintain the short Dot1 protein (Fig. 1D), suggesting that the leaky scanning of the DOT1 transcript is dependent on growth conditions.

The two isoforms showed indistinguishable localization, bulk H3K79 methylation, and silencing properties, indicating that the global activity of Dot1 is not affected by its N-terminus. However, yeast cells expressing either one of the two Dot1 isoforms showed a modest but reproducible difference in sensitivity to Calcofluor White (CFW). Since CFW has a high affinity for chitin in the cell wall of fungi, cells that are resistant to CFW are thought to have a defect in chitin biosynthesis and/or cell wall biogenesis (32;33). Interestingly, the Dot1 homolog in Drosophila, grappa, has also recently been implicated in stress resistance (35). Our results suggest that cells expressing the wild-type Dot1 products or only the full length protein are more sensitive to CFW than cells lacking Dot1 or expressing the short Dot1 protein only. Thus, the Dot1Δ1-17 protein does not complement the loss of Dot1, suggesting that the sensitivity to CFW is mediated by the N-terminal part that is missing in the Dot1 protein synthesized after leaky scanning. Inspection of the cell wall by electron microscopy revealed no difference between wild-type and dot1Δ cells, suggesting that the CFW resistance is not caused by large structural rearrangements of the cell wall. However, the outer layer of the cell wall had a slightly altered structure in cells overexpressing either isoform of Dot1 compared to wild type or dot1Δ cells. This effect was strongest in cells expressing only the short Dot1 isoform. These results suggest that loss of the Dot1 N-terminus causes subtle changes in cell wall metabolism probably related to chitin deposition, whereas overexpression of Dot1 causes more pronounced changes of the cell wall.

Two previous studies have reported genetic interactions between DOT1 and genes involved in cell wall assembly and chitin synthesis. Deletion of DOT1 is
Two Dot1 isoforms as a result of leaky scanning

Two Dot1 isoforms as a result of leaky scanning

synthetically lethal with deletion of either CHS5, encoding a protein that regulates targeting of the chitin synthase Chs3 to sites of polarized growth, or of GAS1, a gene required for cell wall assembly (9;10). CHS5 also shows genetic interactions with BRE1 and LGE1 (10), which encode subunits of the H2B ubiquitin ligase complex and are required for efficient H3K79 methylation (36;37). However, expression of GAS1 and CHS5 was unaffected in strains lacking Dot1 or expressing either one of the two isoforms in the absence or presence of 20 μg/ml CFW (data not shown). The same was true for four genes that are induced by high CFW concentrations (ARO9, YEA4, FUR4 and TUS1, [38]) (data not shown). In addition, in genome-wide expression profiling studies, loss of DOT1 had no or only a very minor effect on genome-wide transcription levels (39;40). We note that in contrast to the functions of Dot1 in telomeric silencing and the DNA damage response (20), our results suggest that the cell wall function of Dot1 is independent of global methylation levels (Fig. 4B). Therefore, mechanism by which Dot1 affects the yeast cell wall is still elusive.

The two isoforms of the Dot1 protein in yeast are the result of leaky scanning. Leaky scanning is a common phenomenon in mammals and plants (15) and a few examples have been described in yeast (29;41;42). However, it is unclear how frequent the leaky scanning mechanism is in yeast. In higher eukaryotes, there are strict requirements for the sequence context around the start codon to allow efficient translation initiation. The optimal sequence is defined by the Kozak rule: a purine at position -3 and/or a G at position +4 (15). In yeast, these residues seem to be less critical for the initiation of translation (24;25). However, a recent computational study demonstrated a strong bias in yeast for an A at position -3 and a C at position +5 (26). Our studies show that changes in position -3 and +5 had very little if any effect on leaky scanning of AUG1 of DOT1. Moreover, even replacement of the complete DOT1 5' UTR by an exogenous sequence did not significantly affect the rate of leaky scanning (Fig. 2F). Together these results strongly suggest that the 5' UTR of the DOT1 gene is not a factor in determining the efficiency of translation initiation.

Besides a suboptimal sequence context, the length of the mRNA leader sequence can also contribute to leaky scanning. Some mRNA transcripts have a 5' leader sequence that is too short for the ribosome to bind efficiently, leading to skipping of the first start codon (28;43). A leader shorter than twelve nucleotides has been shown to promote leaky scanning, and increasing the leader to more than twenty nucleotides restores proper initiation (44). Recently, the transcription start sites of most yeast genes have been mapped using RNA sequencing (45). This study showed that the mRNA transcript generated from the
The *DOT1* gene has a 5’ leader sequence of ~80 base pairs, which is longer than the median of all genes (45) and much longer than the minimal length (44). Thus, it is very unlikely that the first AUG is positioned too close to the 5’-end of the mRNA to be recognized by the ribosome. Translation can also be affected by the 3’ UTR of a transcript. In mammals and plants, binding of a microRNA to its recognition sequence in the 3’ UTR of a target mRNA can inhibit cap-dependent translation initiation by a mechanism that is not yet well understood (46). In addition, in specific tumor cells, the 3’ UTR of the Her-2 oncogene can override translational inhibition of the Her-2 mRNA, showing that such mechanisms are important in disease (47). In the case of the *DOT1* transcript, however, the 3’ UTR is not a key factor in translation initiation because introduction of a C-terminal tag and thereby replacing the endogenous *DOT1* 3’UTR with the 3’UTR of the *ADH1* gene did not affect the synthesis of the two Dot1 isoforms (Fig. 1G). This indicates that leaky scanning of the *DOT1* mRNA is affected by features in the open reading frame of the gene. Interestingly, several cases have been reported for a role of downstream coding sequences in translation initiation in other eukaryotes (26;48;49). However, mutagenesis of the sequence proximal to the initiator codon did not affect the rate of leaky scanning (Fig. 3), indicating that sequences more downstream are involved.

One possibility is that the structure of the Dot1 RNA molecule affects the efficiency of translation initiation. However, mutation of the sequences immediately downstream of the initiator codon that were predicted to disrupt a putative stem-loop did not affect the rate of leaky scanning (Fig. 3). Another possibility might be that the sequence between AUG1 and AUG17 contains several ‘uncommon’ codons for low-abundance tRNAs. If this is the case, production of the long isoform would be energetically more taxing for the cell and leaky scanning might therefore be favorable. However, inspection of the codons between AUG1 and AUG17 suggests that this is unlikely. Only three of the fifteen codons have the lowest usage frequency for the amino acid they encode. Furthermore, the average codon usage of the fifteen codons (0.024) is higher than the average codon usage for all yeast genes (0.016). Thus, the mechanism by which the open reading frame of the *DOT1* gene affects leaky scanning is still unclear.

Recent RNA sequencing of the yeast transcriptome predicts that transcripts of as many as 321 yeast genes (~6% of the yeast genome) contain small open reading frames (uORFs) upstream of the first in-frame start codon (45). Moreover, uORFs are present in almost half of all human and mouse transcripts (50). A number of uORFs have now been shown to control translation of the downstream gene (17;50) and a key question is how the start codons of the
Two Dot1 isoforms as a result of leaky scanning

uORFs of a gene are used or skipped to tune the efficiency of protein expression. The frequent occurrence in eukaryotic mRNAs of upstream AUGs or sub-optimal start codons indicates that alternative translation initiation plays an important role in regulation of protein expression and might be more widespread than previously thought (51). The discovery of alternative translation events like the one described here will be instrumental for understanding the mechanisms of translation start site selection.

MATERIALS AND METHODS

Yeast strains, plasmids and media

Yeast strains and plasmids used in this study are listed in Table 1. Yeast media were described previously (2). Silencing assays were performed using media containing 1 g/l 5-fluoroorotic acid (Toronto Research Chemicals). Calcofluor White (CFW) sensitivity of cells was tested using YEP media containing 20 µg/ml CFW (Sigma). Cycloheximide (Sigma) was added to the media at a concentration of 50 µg/ml to inhibit protein synthesis. NKI3031 was derived from a cross between BY4740 and BY4742 (52). NKI3049 and NKI3066 were derived from NKI3031 by integration of the GAL1 promoter in front of the DOT1 gene using plasmid pYM-N22 and PCR-mediated gene replacement (53); in NKI3049 the GAL1 promoter was integrated such that 62 bp of the endogenous DOT1 5' untranslated region was maintained, whereas in NKI3066 the endogenous DOT1 5' UTR was replaced by GAL1 and vector sequence present in pYM-N22. NKI1060 was constructed from UCC7164 using PCR-mediated gene replacement and plasmid pFA6a-KanMX-PGAL1-3HA (54). The single-copy pLEU2-DOT1 vector (pDOT1) and the multicopy GAL1-promoter pTRP1-DOT1 plasmid (pFvL18) were described previously (2). Mutations in the DOT1 gene in these plasmids were introduced by Quikchange (Stratagene) or a three-step PCR protocol and verified by sequencing. Details of the mutant DOT1 genes are shown in Fig. 2A. pFvL201 was generated by replacing the first methionine of DOT1 in pDOT1 by a single FLAG tag followed by a SGGSGG spacer. pFvL212 was derived from pRS314-DOT1 (single-copy pTRP1-DOT1) by integration of a GSGGSGGS spacer, single HA TAG and TAP tag behind the DOT1 gene; the TAP tag sequence was amplified from pBS1479 (55). pFvL215 was generated by deletion of amino acids 4-20 from pFvL212 by gap-repair. pFvL221 was generated from pFvL18 by integration of a GSGGSGG linker behind the DOT1 gene. pTW055 was derived from pFvL221 by insertion of the yeast-enhanced citrine tag from plasmid pKT211 (56) behind the linker using an AatII-AcII digest. Mutations in pTW055 were generated by integrating the Sall-EcoRI fragment from pGH03 and pGH04 in Sall-EcoRI digested pTW055 to generate pGH024 and pGH025, respectively.

UV irradiation

Exponentially growing cells (OD660 ~0.5) were arrested in G1/S with 5 µg/ml α-factor for 2 hrs, spun, resuspended in fresh media containing α-factor and plated on 14 cm Petri-dishes. The plates were irradiated with 100 J/m² UV light of 254 nm using a Stratalinker.
Trichloroacetic acid protein extracts were prepared 30 minutes after irradiation and compared to non-irradiated controls.

**Immunoblots**

Whole-cell extracts were obtained from approximately 5x10²⁷ cells by the classical glass beads breakage method using 200 µl of glass beads and SUMEB lysis buffer complemented with PMSF (1 mM), benzamidine (5 mM), pepstatin (1 µg/ml), leupeptin (1 µg/ml) and DTT (1 µM). Primary antibody incubations were performed overnight in Tris-buffered saline-Tween with 2% dry milk. The rabbit polyclonal antibodies against Dot1 and H3K79me1, -me2 and -me3 were described before. Commercially available antibodies that were used in this study are H3 (ab1791, Abcam), Pgk1 (A-6457, Invitrogen), and Rad53 (sc-6749, Santa Cruz).

**RNA structure prediction**

The Vienna RNAfold web server was used to predict secondary RNA structures and the effect of mutations in the DOT1 RNA sequence. Analysis of different sized fragments from the DOT1 region revealed a common predicted structure of the 5' part of the DOT1 mRNA.

**Confocal microscopy**

Log phase cells were fixed for 10 minutes in 4% formaldehyde in PBS, after which they were incubated with Hoechst 33342 (Invitrogen, 1 µg/ml) for 15 minutes to stain the nucleus. Cell wall staining was performed by incubation of the cells with ConA-Cy5 for 15 minutes. Cells were mounted in Vectashield mounting medium (Vector Laboratories). Samples were analyzed with a confocal laser scanning microscope equipped with a HCX PL APO lbd.bl 63x/NA 1.4 oil corrected objective lens (Leica). The acquisition software used was Leica LCS. Cells were imaged using a 405 nm, a 516 nm and a 633 nm laser to visualize Hoechst, Citrine and Cy5 respectively.

**Electron Microscopy**

Log phase cells were fixed in Karnovsky fixative containing 0.15% Ruthenium Red (BDH chemical) for 5 minutes at room temperature, for marking the outer membranes of the cells. After washing with 0.1 M cacodylate buffer pH 7.2 the cells were postfixed with 1% osmiumtetroxide, en-block stained with 0.5% uranylacetate (Ultrastain 1, Leica) followed by dehydration series and embedding in EMBed 812 (Aurion). The sections were examined using a CM10 microscope (FEI).
TABLE 1. Yeast strains and plasmids

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCC7164</td>
<td><strong>MATa ade2Δ::hisG his3Δ200 leu2Δ0 lys2Δ0 met15Δ0</strong> trp1Δ63 ura3Δ0 ADE2-TEL-VR URA3-TEL-VII</td>
<td>(20)</td>
</tr>
<tr>
<td>UCC7183</td>
<td><strong>MATa ade2Δ::hisG his3Δ200 leu2Δ0 lys2Δ0 met15Δ0</strong> trp1Δ63 ura3Δ0 ADE2-TEL-VR URA3-TEL-VII dot1Δ::KanMX</td>
<td>(20)</td>
</tr>
<tr>
<td>NKI1059</td>
<td><strong>MATa lys2Δ0 trp1Δ63 his3Δ200 ade2Δ::hisG ura3Δ0 leu2Δ0 met15Δ0 ADE2-TEL-VR URA3-TEL-VII dot1Δ::KanMX-PGAL1-3HA-DOT1</strong></td>
<td>(20)</td>
</tr>
<tr>
<td>NKI1060</td>
<td><strong>MATa lys2Δ0 trp1Δ63 his3Δ200 ade2Δ::hisG ura3Δ0 leu2Δ0 met15Δ0 ADE2-TEL-VR URA3-TEL-VII dot1Δ::KanMX-PGAL1-3HA-dot1(Δ1-17)</strong></td>
<td>This study</td>
</tr>
<tr>
<td>NKI3031</td>
<td><strong>MATa leu2Δ0 lys2Δ0 ura3Δ0</strong></td>
<td>This study</td>
</tr>
<tr>
<td>NKI3049</td>
<td><strong>MATa leu2Δ0 lys2Δ0 ura3Δ0 dot1Δ::KanMX-PGAL1-DOT1</strong></td>
<td>This study</td>
</tr>
<tr>
<td>NKI3066</td>
<td><strong>MATa leu2Δ0 lys2Δ0 ura3Δ0 dot1Δ::KanMX-PGAL1-DOT1</strong></td>
<td>This study</td>
</tr>
<tr>
<td>YSC-1122</td>
<td><strong>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 kre1Δ::KanMX</strong></td>
<td>Open Biosystems</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRS315</td>
<td><strong>LEU2 CEN</strong></td>
<td>(2)</td>
</tr>
<tr>
<td>pDOT1</td>
<td><strong>LEU2 CEN DOT1</strong></td>
<td>(2)</td>
</tr>
<tr>
<td>pGH01</td>
<td><strong>LEU2 CEN dot1-M1T</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pGH02</td>
<td><strong>LEU2 CEN dot1-M17L</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pFF018</td>
<td><strong>LEU2 CEN dot1-17M</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pFF025</td>
<td><strong>LEU2 CEN dot1-initiation mutant 1</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pFF026</td>
<td><strong>LEU2 CEN dot1-initiation mutant 2</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pFF027</td>
<td><strong>LEU2 CEN dot1-initiation mutant 3</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pGH18</td>
<td><strong>LEU2 CEN dot1- frameshift mutant 1</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pGH21</td>
<td><strong>LEU2 CEN dot1- frameshift mutant 2</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pFF028</td>
<td><strong>LEU2 CEN DOT1 (A-6C, A04C, G4U, G5U, C6A)</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pFF029</td>
<td><strong>LEU2 CEN DOT1 (C60G)</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pFF030</td>
<td><strong>LEU2 CEN DOT1 (C6U)</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pFFL212</td>
<td><strong>TRP1 CEN DOT1-spacer-HA-TAP</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pFFL215</td>
<td><strong>TRP1 CEN dot1(Δ4-20)-spacer-HA-TAP</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pTCG</td>
<td><strong>TRP1 2μ PGAL1</strong></td>
<td>(2)</td>
</tr>
<tr>
<td>pFFL18</td>
<td><strong>TRP1 2μ PGAL1-DOT1 (G8A)</strong></td>
<td>(2)</td>
</tr>
<tr>
<td>pGH03</td>
<td><strong>TRP1 2μ PGAL1-dot1-M1T</strong></td>
<td>This study</td>
</tr>
<tr>
<td>pGH04</td>
<td><strong>TRP1 2μ PGAL1-dot1-M17L</strong></td>
<td>This study</td>
</tr>
</tbody>
</table>
pGH06  TRP1 2μ PGAL1-DOT1 (G8A, U22A, C23G, A24U)  This study
pFvL221  TRP1 2μ PGAL1-DOT1-linker\(^6\)  This study
pTW055  TRP1 2μ PGAL1-DOT1-linker-YECitrine\(^6\)  This study
pGH024  TRP1 2μ PGAL1-dot1-M1T-linker-YECitrine\(^6\)  This study
pGH025  TRP1 2μ PGAL1-dot1-M17L-linker-YECitrine\(^6\)  This study

1 The DOT1 gene is flanked by its endogenous 5' context in NKI3049; in NKI3066 the 5' sequence has been replaced by vector sequence (pYM-N22) and 5' sequence of the GAL1 gene.
2 The following point mutations were introduced in the three initiation mutants (numbers are relative to A of AUG1): G-3 to A-3 in pFF025; G+5 to C+5 in pFF026 and the double mutation G-3 to A-3 and G+5 to C+5 in pFF027.
3 The following base insertions (bold-faced and underlined) were made in the frameshift mutants:
FSM1 (pGH1B): ATG GCC GGT CAA GAA AGT ATA TCA AAA TAA • (TAA = termination codon)
FSM2 (pGH21): ATG GCC GGT CAA GAA AGT A\(\text{CT}^\text{A}\) ATC AAA TAA • (TAA = termination codon)
4 pFvL201 contains a flexible spacer between the FLAG tag and the Dot1 protein, consisting of the amino acids SGGSGG.
5 pFvL212 and pFvL215 contain a flexible spacer between the Dot1 protein and the HA tag, consisting of the amino acids GSGGSGGS.
6 pFvL221, pTW055, pGH024 and pGH025 contain a linker sequence at the very C-terminus of the Dot1 protein, consisting of the amino acids GSGGSGGS.

ACKNOWLEDGEMENTS

We thank M. Fornerod, R. Agami, and members of the van Leeuwen lab for discussions and/or critical reading of the manuscript, T. van Welsem and N. Ong for technical assistance and help with plasmid constructions.
This work was supported by the EU 6th framework program NOE ‘The Epigenome’ [LSHG-CT-2004-503433]; the Leukemia and Lymphoma Society (Special Fellow 3409-04 to F.v.L.); and a VIDI grant from the Netherlands Organisation for Scientific Research to F.v.L.
REFERENCES


Two Dot1 isoforms as a result of leaky scanning


SUPPLEMENTARY INFORMATION

Figure S1. Cells lacking Dot1 or expressing only one Dot1 isoform are not sensitive to changes in osmolarity.

Yeast cells were spotted in tenfold dilutions on YEP media containing 2% galactose as the carbon source. The plates contained the following additives: 0.3M NaCl, 1M NaCl, 1M sorbitol, 1.5M sorbitol, 2.5M glycerol, 0.001% SDS or 0.01% SDS. The following strains were used: wild type (UCC7164), dot1Δ (UCC7183), P_{GAL1}-3xHA-DOT1 (NKI1059) and P_{GAL1}-3xHA-DOT1Δ1-17 (NKI1060). The known osmosensitive strain hog1Δ (YSC-2724, Open Biosystems) was used as a positive control.