A genetic screen for terminator function in yeast identifies a role for a new functional domain in termination factor Nab3

Travis J. Loya, Emory University
Thomas W. O'Rourke, Emory University
Daniel Reines, Emory University

Journal Title: Nucleic Acids Research
Volume: Volume 40, Number 15
Publisher: Oxford University Press (OUP): Policy C - Option B | 2012-05-07, Pages 7476-7491
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1093/nar/gks377
Permanent URL: http://pid.emory.edu/ark:/25593/f75p6

Final published version: http://nar.oxfordjournals.org/content/40/15/7476

Copyright information:
© Loya et al. 2012 Published by Oxford University Press
This is an Open Access work distributed under the terms of the Creative Commons Attribution-Noncommercial 3.0 Unported License (http://creativecommons.org/licenses/by-nc/3.0/).

Accessed March 8, 2022 1:08 AM EST
A genetic screen for terminator function in yeast identifies a role for a new functional domain in termination factor Nab3

Travis J. Loya, Thomas W. O'Rourke and Daniel Reines*

Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322, USA

Received March 22, 2012; Revised April 6, 2012; Accepted April 12, 2012

ABSTRACT

The yeast IMD2 gene encodes an enzyme involved in GTP synthesis. Its expression is controlled by guanine nucleotides through a set of alternate start sites and an intervening transcriptional terminator. In the off state, transcription results in a short non-coding RNA that starts upstream of the gene. Transcription terminates via the Nrd1–Nab3–Sen1 complex and is degraded by the nuclear exosome. Using a sensitive terminator read-through assay, we identified trans-acting Terminator Override (TOV) genes that operate this terminator. Four genes were identified: the RNA polymerase II phosphatase SSU72, the RNA polymerase II binding protein PCF11, the TRAMP subunit TRF4 and the hnRNP-like, NAB3. The TOV phenotype can be explained by the loss of function of these gene products as described in models in which termination and RNA degradation are coupled to the phosphorylation state of RNA polymerase II’s repeat domain. The most interesting mutations were those found in NAB3, which led to the finding that the removal of merely three carboxy-terminal amino acids compromised Nab3’s function. This region of previously unknown function is distant from the protein’s well-known RNA binding and Nrd1 binding domains. Structural homology modeling suggests this Nab3 ‘tail’ forms an α-helical multimerization domain that helps assemble it onto an RNA substrate.

INTRODUCTION

An interesting role of transcription termination in controlling nucleotide-regulated genes has recently been brought to light in Saccharomyces cerevisiae (1,2). IMD2 encodes IMP dehydrogenase, an important enzyme in the de novo guanine nucleotide biosynthetic pathway. Its transcription is regulated by cellular guanine nucleotides through alternative start sites and an intergenic terminator (IT) (Figure 1A). When GTP is abundant, transcripts initiate at an upstream G-residue. Due to the strong terminator, this initiation choice results in a short, non-coding transcript that is degraded. In this manner, IMD2 attains the ‘off’ state when IMP dehydrogenase activity is not needed (high GTP levels). Termination employs the RNA binding heterodimer Nrd1–Nab3 and the putative helicase Sen1, all encoded by essential genes (3–5). This system operates to terminate a number of other small RNAs, including snRNAs, snoRNAs and cryptic unstable transcripts (CUTs) (6). The CUT upstream of IMD2 is degraded by the nuclear exosome (3,7). When IMD2 expression is needed at times of GTP scarcity, RNA polymerase II (pol II) initiates at an adenine downstream of the terminator due to the unavailability of GTP needed for initiation at the upstream site (Figure 1A) (4).

NAB3 encodes a protein with an RNA recognition motif (RRM), which is important for cell growth and allows the protein to interact with the sequence UCUU in an RNA substrate (8–11). This central core of Nab3 has been expressed in Escherichia coli and dimerizes with Nrd1. The resulting heterodimer binds RNA with high affinity (12). The protein is peculiar in that it has an amino-terminal region very rich in aspartate (D) and glutamate (E) and a carboxy-terminal region rich in proline (P) and glutamine (Q), with the respective regions showing long runs of D, E, P and Q homopolymers. The function of these repeat sequence domains is not known. The amino-terminal D/E-rich region is dispensable for cell viability, whereas the carboxy terminal ≈30% of the protein,
including the P/Q region, is not (12). The Nab3–Nrd1 complex is thought to recruit the Sen1 helicase to appropriate elongation complexes to provoke termination (13–15). Nrd1 possesses a pol II carboxy-terminal domain (CTD) interaction region that preferentially binds to the pol II CTD phosphorylated on ser5 of its repeating heptapeptide unit (16). This interaction, along with Nab3’s recognition of its binding site in RNA, is thought to contribute to the specificity of the system in enabling the termination complex to recognize only some promoter proximal elongation complexes, particularly those that generate short RNAs (12,14–20). In addition, Nrd1 possesses an RRM and recognizes RNA containing the GUA[A/G] sequence (8,21,22). Association of the two RRM-containing proteins Nab3 and Nrd1 yields cooperative RNA binding (12). A number of proteome-wide yeast protein–protein interaction maps have documented the physical interactions between Nrd1, Nab3 and Sen1 (23,24). The phosphorylation pattern of the repeat domain of the large subunit of pol II appears to be a marker for promoter proximal Nab3–Nrd1-dependent termination of short RNAs and may explain some of the specificity of termination for short RNAs (16,25).

In addition to participating in this alternative form of transcription termination, the Nrd1–Nab3–Sen1 complex also recruits the nuclear exosome’s ribonuclease activity to degrade or process many of the resulting small RNAs (26).

The coupling of termination and RNA hydrolysis employs the TRAMP complex, which is composed of a copy of the non-canonical polyA polymerase Trf4 or Trf5, a copy of the zinc knuckle protein Air1 or Air2 and the Mtr4 RNA helicase (19,20,27,28). The TRAMP complex interacts with Nrd1 as well as the nuclear exosome and stimulates the latter’s ribonuclease function (15,19,28,29). Through this coupling role, TRF4 is involved in the metabolism of CUTs and the processing of small non-coding RNAs (30).

Here we use a genetic screen in *Saccharomyces cerevisiae* to identify TOV (‘terminator override’) genes. Based upon an analysis of RNA metabolism, two classes of mutants could be identified. One group appeared to be primarily RNA degradation deficient and two independent isolates were shown to contain mutations in the TRAMP component, *TRF4*. The other class displayed terminator read-through defects. Two of the mutant strains had alterations in *SSU72*. One had a mutation in *PCF11*. Another two mutants in this class fell within the *NAB3* gene. The *tov* changes represent new mutant alleles for each of these termination-related genes. The *nab3* alleles were particularly interesting since they revealed a new portion of this hnRNP-like protein in a short region of its carboxy-terminal amino acids. Structural homology suggests that this region is functionally akin to an α-helix in hnRNP C, which serves to multimerize that RNA binding protein on substrate RNAs. These

Figure 1. (A) Model for the regulation of IMD2 by GTP. The choice of start sites for *IMD2* transcription varies as a function of GTP levels (4). Under GTP-replete conditions, initiation starts upstream (thin bent arrow) and terminates at the IT. The resulting non-coding RNA (short wavy line) is degraded by the exosome and *IMD2* mRNA is not produced. When GTP is scarce, nucleotide-starved pol II bypasses the upstream start in favor of an adenine-initiated downstream start (bold bent arrow). The IT is bypassed, full-length mRNA is made (long wavy line) and expression of *IMD2* results. (B) *IMD2*’s terminator quenches GFP reporter expression in yeast. Yeast strains harboring a plasmid that contains (DY1514; center patch), or lacks (DY1513; left patch) the *IMD2* intergenic terminator between the *GAL1* promoter and GFP reading frame, were patched onto galactose-containing medium and photographed under white light (top) or blue light to monitor GFP fluorescence (bottom). A strain with a plasmid containing a previously described point mutation in the *IMD2* terminator (36) was tested in the patch on the right.
findings show the effectiveness of sorting live fluorescent yeast as a genetic screen and of identifying mutations by whole genome sequencing. They also reveal a new degree of complexity in the Nab3 protein.

MATERIALS AND METHODS

Cell growth and media

Yeast strains used in this study are presented in Table 1. Cells were grown in rich media (YPD) or standard selective ‘drop-out’ media (SC-ura, SC-leu or SC-ura-leu) at 30°C unless otherwise indicated. For galactose induction, cells were grown in SC-ura in which glucose was replaced by raffinose as a carbon source and 2% galactose was added as indicated or cells were grown continuously in galactose. Plasmids were introduced into yeast strains using the lithium acetate procedure (31).

Plasmid/strain construction and genomic integration

Oligonucleotides used in this study are listed in Table 2. The plasmid pREF-GFP was assembled by a three-way ligation of a PCR product encoding the yeast IMD2 IT (HindIII-PstI cut product made with primers ‘oPur5P-F170’ and ‘oR90-PstI’) into the plasmid pYES2 (Invitrogen; cut with HindIII and XhoI along with a PCR product encoding the GFP open reading frame (PstI-XhoI cut PCR product made with primers ‘oGFP-stop’ and ‘oGFP-atg’). The kanamycin resistance gene from pFA6a-KanMX4 (33) was excised by digestion with EcoRI and BamHI, filled in with DNA polymerase and inserted into the blunted BglII site of the ampicillin resistance gene in pREF-GFP to yield pREF-GFP-Kan.

The plasmid pRS315-nab3Q784X was made by generating a PCR product from the genome of the mutant strain 1A1 using the oligonucleotides ‘Nab3Prom-UP-BamHI’ and ‘Nab3cds-Dwn-XhoI’. The product was cut with XhoI and BamHI and inserted into similarly cut pRS315. The nab3-11 allele was amplified by PCR using the same primers and genomic DNA from strain ACY1223. It was digested with XhoI and BamHI and inserted into pRS315, to yield plasmid pnab3-11. A PCR product of wild-type NAB3 was generated from BY4742 genomic DNA using the same primers and inserted into pRS316 cut with BamHI and XhoI to yield pRS316-NAB3.

The Nab3 tail deletion series was made by subcloning a 1.2 kbp HindIII-XhoI fragment encoding the 3′-end of the coding sequence and untranslated region into similarly cut pBluescript II KS(−) (Stratagene, Inc.). This plasmid was cut with BspMI and XhoI into which was inserted a series of oligonucleotide duplexes with compatible overhangs and which encoded portions of the Nab3 tail (‘Δ1’: Nab3Δ1 tail Fwd’ and Nab3Δ1 tail REV; ‘Δ2’: Nab3Δ2 tail Fwd’ and Nab3Δ2 tail REV; ‘Δ3’: Nab3-15aa-tail Fw’ and Nab3-15aa-tail Re’ and full length (‘FL’): Nab3-18aa-tail Fw’ and Nab3-18aa-tail Re’). These constructs generated a series of intermediate plasmids from which fragments were removed to build the deletion derivatives of Nab3. Restriction fragments encoding these rebuilt 3′-ends were individually removed from their respective Bluescript vectors by digestion with HindIII and XhoI and inserted into similarly cut pRS315-NAB3 from which the natural sequence was removed. This resulted in the final plasmids pRS315-nab3CAΔ1, pRS315-nab3CAΔ2, pRS315-nab3CA3 and pRS315-nab3FL.

pTRF4 was obtained from Dr A. Corbett (Emory U.) and contained the wild-type copy of TRF4 inserted into pRS315. The plasmid ptrf4-13A was assembled from a XhoI and SacI digested PCR product made from genomic DNA of strain 2A1 (tov2) using oligonucleotides ‘AC4037-SacI-Trf4-Fwd’ and ‘AC4040-XhoI-Trf4-Rev’ and inserted into similarly cut pRS315. Similarly, ptrf4S430stop was made by amplifying the TRF4 allele in genomic DNA from 2C1 (tov6) and inserting it into pRS315 as was done for ptrf4-13A. pRS315-SSU72 and pRS315-ssu72-2D were made by inserting a PCR product made from wild-type or 2D1 genomic DNA, respectively (using primers SSU72-Bam-For and SSU72-Xho-Rev), into BamHI and XhoI cut pRS315. pRS315-PCF11 was made by amplifying from genomic DNA using the oligonucleotide ‘Pcf11BamHIFor’ and ‘Pcf11SalIIRev’, cutting the product with BamHI and SalI and inserting it into pRS315.

Strains DY1513 and DY1514 were made by transforming BY4742 (Open Biosystems) with the plasmids pGAL-GFP-kan and pREF-GFP-kan, respectively. Isolates that were selected by FACS were named 1A1, 2A1, 3A1, 4A1, 2B1, 2C1 and 2D1 and assigned tov numbers one through seven, respectively. These strains (and their parent strain, DY1514) were cured of the reporter plasmid by growth on 5-flouro-orotic acid (FOA) and named 1A1F, 2A1F, 3A1F, 4A1F, 2B1F, 2C1F, 2D1F and DY1514F, respectively. Strain 2A1 was transformed with pTRF4 or ptrf4S430stop to generate DY3029 and DY3092, respectively. Strain 2B1 was transformed with pREF-GFP-kan and either pRS315-PCF11 (DY3628) or pRS315 (DY3627).

To make a family of strains lacking chromosomal NAB3, a diploid heterozygous for a kanMX disruption of the gene (Open Biosystems YSC1021-669423) was transformed with pRS315-nab3Q784X to yield DY2222. This strain was sporulated and a resulting haploid (DY2233) lacking chromosomal NAB3 and containing the plasmid was isolated. This strain was transformed with pREF-GFP-kan to yield DY2240. DY2233 was transformed with URA3-marked pnab3-11 and screened for the loss of LEU2-marked pRS315-nab3Q784X to generate DY3029. DY3029 was independently transformed with pRS315-nab3CAΔ1, pRS315-nab3CAΔ2, pRS315-nab3CA3 or pRS315-nab3FL and each was grown on FOA to select for loss of pnab3-11. The resulting strains were transformed with pREF-GFP-kan to generate DY3060, DY3061, DY3038 and DY3036, respectively.

In some strains, a genomic copy of the GFP reporter transcription unit was integrated into a gene-poor region of chromosome XII. The oligonucleotides ‘5′GalGFP(C)’ and ‘3′GalGFP(C)’ were used to amplify the integration cassette from plasmids pREF-GFP-kan or pGal-GFP-kan. The former plasmid contains IMD2’s IT sequence between the GAL1 promoter and the GFP reading
Table 1. Yeast strains used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 [pREF-GFP kan (URA3)] tov1</td>
</tr>
<tr>
<td>A1F</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov1</td>
</tr>
<tr>
<td>2A1</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 [pREF-GFP kan (URA3)] tov2</td>
</tr>
<tr>
<td>2A1F</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov2</td>
</tr>
<tr>
<td>2B1</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 [pREF-GFP kan (URA3)] tov5</td>
</tr>
<tr>
<td>2B1F</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov5</td>
</tr>
<tr>
<td>3C1</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 [pREF-GFP kan (URA3)] tov6</td>
</tr>
<tr>
<td>3C1F</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov6</td>
</tr>
<tr>
<td>3A1</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 [pREF-GFP kan (URA3)] tov7</td>
</tr>
<tr>
<td>3A1F</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov7</td>
</tr>
<tr>
<td>4A1</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 [pREF-GFP kan (URA3)] tov4</td>
</tr>
<tr>
<td>4A1F</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov4</td>
</tr>
<tr>
<td>ACY1223</td>
<td>MATa ade2 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 nab3-11</td>
</tr>
<tr>
<td>DY1513</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 [pRS315 nab3 (LEU2)]</td>
</tr>
<tr>
<td>DY1514</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 [pRS315 nab3 (LEU2)]</td>
</tr>
<tr>
<td>DY1514F</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0</td>
</tr>
<tr>
<td>DY2044</td>
<td>MATa lys2-A14 ade5-1 his7-2 leu2-3,112 trp1::KANMX trp1-289 ura3-52 [pREF-GFP kan (URA3)]</td>
</tr>
<tr>
<td>DY2222</td>
<td>MATa his3A1 his3A1 leu2A0/leu2A0 lys2A0/LYS2 met15A0/MET15 NAB3/nab3::kanMX ura3A0/ura3A0 [pREF-GFP kan (URA3)]</td>
</tr>
<tr>
<td>DY2233</td>
<td>MATa his3A1 leu2A0 (lys2, met15 status unknown) nab3::kanMX ura3A0</td>
</tr>
<tr>
<td>DY2240</td>
<td>MATa his3A1 leu2A0 (lys2, met15 status unknown) nab3::kanMX ura3A0</td>
</tr>
<tr>
<td>DY3009</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 chr. XII::GAL-GFP-kan</td>
</tr>
<tr>
<td>DY3001</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 chr. XII::GAL-GFP-kan</td>
</tr>
<tr>
<td>DY3003</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov1 chr. XII::GAL-GFP-kan</td>
</tr>
<tr>
<td>DY30059</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov2 chr. XII::GAL-GFP-kan</td>
</tr>
<tr>
<td>DY30079</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov3 chr. XII::GAL-GFP-kan</td>
</tr>
<tr>
<td>DY3009</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov4 chr. XII::GAL-GFP-kan</td>
</tr>
<tr>
<td>DY30109</td>
<td>MATa ade2 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 nab3-11 [pNAB3]</td>
</tr>
<tr>
<td>DY3011</td>
<td>MATa ade2 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 nab3-11 [pRS315]</td>
</tr>
<tr>
<td>DY3012</td>
<td>MATa ade2 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 nab3-11 [pRS315-nab3-784X]</td>
</tr>
<tr>
<td>DY3013</td>
<td>MATa ade2 can1-100 his3-11,15 leu2-3,112 trp1-1 ura3-1 nab3-11 [pRS315-nab3-784X]</td>
</tr>
<tr>
<td>DY3018</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov1 [pREF-GFP kan (URA3)] [pRS315 (LEU2)]</td>
</tr>
<tr>
<td>DY3019</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov1 [pREF-GFP kan (URA3)]</td>
</tr>
<tr>
<td>DY30229</td>
<td>MATa his3A1 leu2A0 (lys2, met15 status unknown) nab3::kanMX ura3A0</td>
</tr>
<tr>
<td>DY30329</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov2 [pREF-GFP kan (URA3)] [pTRF4 (LEU2)]</td>
</tr>
<tr>
<td>DY3036</td>
<td>MATa his3A1 leu2A0 (lys2, met15 status unknown) nab3::kanMX ura3A0</td>
</tr>
<tr>
<td>DY3038</td>
<td>MATa his3A1 leu2A0 (lys2 and met15 status unknown) nab3::kanMX ura3A0</td>
</tr>
<tr>
<td>DY3060</td>
<td>MATa his3A1 leu2A0 (lys2, met15 status unknown) nab3::kanMX</td>
</tr>
<tr>
<td>DY3061</td>
<td>MATa his3A1 leu2A0 (lys2, met15 status unknown) nab3::kanMX</td>
</tr>
<tr>
<td>DY3078</td>
<td>MATa his3A1 ade5-1 his7-2 leu2-3,112 trp1::KANMX TRF4 trp1-289 ura3-52</td>
</tr>
<tr>
<td>DY3082</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 [pREF-GFP kan (PFR4 (LEU2))</td>
</tr>
<tr>
<td>DY3083</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 [pREF-GFP kan (URA3)] [pTRF4-13A (LEU2)]</td>
</tr>
<tr>
<td>DY3084</td>
<td>MATa his3A1 ade5-1 his7-2 leu2-3,112 trp1::KANMX TRF4 trp1-289 ura3-52</td>
</tr>
<tr>
<td>DY3087</td>
<td>MATa his3A1 ade5-1 his7-2 leu2-3,112 trp1::KANMX TRF4 trp1-289 ura3-52</td>
</tr>
<tr>
<td>DY3092</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov2 [pREF-GFP kan (URA3)] [pTRF4-5430stop (LEU2)]</td>
</tr>
<tr>
<td>DY3114</td>
<td>MATa ura3-52 trp1::A63 leu2A1 [pREF-GFP kan (URA3)]</td>
</tr>
<tr>
<td>DY3115</td>
<td>MATa ura3-52 trp1::A63 leu2A1 [pREF-GFP kan (URA3)]</td>
</tr>
<tr>
<td>DY3625</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 [pREF-GFP kan (URA3)]</td>
</tr>
<tr>
<td>DY3626</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 [pREF-GFP kan (URA3)] tov7 [pRS315-SSU72 (LEU2)]</td>
</tr>
<tr>
<td>DY3627</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov5 [pREF-GFP kan (URA3)] [pRS315-SSU72-2D (LEU2)]</td>
</tr>
<tr>
<td>DY3628</td>
<td>MATa his3A1 leu2A0 lys2A0 ura3A0 tov5 [pREF-GFP kan (URA3)] [pRS315-PFC11 (LEU2)] [pRS315-PFC11 (LEU2)]</td>
</tr>
<tr>
<td>H291</td>
<td>MATa trp1-32 leu2A0 [pDEF-HAIR-SSU72-U3A]</td>
</tr>
<tr>
<td>H291</td>
<td>MATa trp1-32 leu2A0 [pDEF-HAIR-SSU72-U3A]</td>
</tr>
<tr>
<td>H291</td>
<td>MATa trp1-32 leu2A0 [pDEF-HAIR-SSU72-U3A]</td>
</tr>
<tr>
<td>H291</td>
<td>MATa trp1-32 leu2A0 [pDEF-HAIR-SSU72-U3A]</td>
</tr>
<tr>
<td>YH99,</td>
<td>MATa his3A1 ade5-1 his7-2 leu2-3,112 trp1::KANMX TRF4 trp1-289 ura3-52</td>
</tr>
<tr>
<td>YH99,</td>
<td>MATa his3A1 ade5-1 his7-2 leu2-3,112 trp1::KANMX TRF4 trp1-289 ura3-52</td>
</tr>
<tr>
<td>YH99,</td>
<td>MATa his3A1 ade5-1 his7-2 leu2-3,112 trp1::KANMX TRF4 trp1-289 ura3-52</td>
</tr>
<tr>
<td>YH99,</td>
<td>MATa his3A1 ade5-1 his7-2 leu2-3,112 trp1::KANMX TRF4 trp1-289 ura3-52</td>
</tr>
</tbody>
</table>

A. Corbett-Emory U.

This study

Open Biosystems
frame, the latter does not. The PCR product from pREF-GFP-kan was transformed into DY1514F, 1A1F, 2A1F, 3A1F and 4A1F using high-efficiency lithium acetate transformation and G418 resistant isolates were identified by illumination with a Dark Reader lamp and replica plating to a density of 1.33 OD/ml in 1.5 ml and were sorted at 488 nm excitation wavelength.

**Fluorescence-activated cell sorting and analysis**

Parental strain DY1514 was grown overnight at 30°C in SC-ura (raffinose) with 2% galactose. Cells were resuspended to a density of 1.33 OD/ml in 1.5 ml and were sorted at 488 nm using a FACSaria II (Becton–Dickinson) and FACSDiva software. Selected cells were sorted at 488 nm using a FACSAria II (Becton–Dickinson) and FACSDiva software. Selected cells were selected and named H98 or H291, respectively. DY3114 and DY3115 were made by transforming pRS315-SSU72 or pRS315-ssu72-2D, respectively. DY3625 and DY3626 were made following transformation of 2D1 with pRS315-SSU72 or pRS315-ssu72-2D, respectively. DY3114 and DY3115 were made by transforming pREF-GFP-kan into H98 or H291, respectively.

**Northern blots**

Cultures were grown to an optical density (600 nm) of 0.4–0.6, washed in water and frozen at –80°C. Where indicated, 0.5 mM guanine for 30 min or 15 μg/ml mycophenolic acid for 2 hr was added to the media. RNA was extracted and separated on 1% agarose gels for transfer to Zeta- Probe GT Genomic filters as described (34–36). Probes were generated by PCR from plasmid DNA (for GFP) or yeast genomic DNA using the primer pairs ‘SNR33up/SNR33down’, ‘oIMD2P-F500B’/‘oIMD2P-R00’ and ‘oGFP-atg’/‘oGFP-stop’ for SNR33, IMD2 and GFP, respectively. Probes were labeled using random primers (Invitrogen), Klenow DNA polymerase and α-32P-dATP, (Perkin Elmer). Filters were hybridized at 42°C overnight, washed and exposed to HyBlot CL film (Denville Scientific Inc.) with an intensifying screen.

**Whole genome sequencing**

Genomic DNA was purified and subjected to high-throughput sequencing by HiSeq2000 (Illumina Inc.). In brief, the purified DNA was sheared to an average fragment size of 300 bp using acoustic focusing. The SPRI works platform was used to create libraries, which was submitted to Illumina. Whole genome shotgun Illumina HiSeq sequenced was used to generate draft genome assemblies.
were sequenced on an Illumina HiSeq2000 with a read length of 50 bases plus 7 bases for the multiplex tags. The resulting sequence was extracted and de-multiplexed using Illumina’s SCS2.8 software. The data were then analyzed using the CLC Genomics Workbench package. Imported reads were trimmed to remove low-quality sequence as well as any reads of <36 bases in length. Reads were mapped to the published yeast genome obtained from the Saccharomyces Genome Database, yielding an average coverage of 95× across the genome. Variations (SNPs and InDels) were called using a minimum Q score of 20 and a coverage minimum of 25×. This yielded over 500 SNPs and 350 InDels per strain. We focused on SNPs in protein-encoding sequences to look for differences between strains. Using data from other similar strains as well as gene annotation, variations that were common or unlikely to be related to the phenotype were removed from consideration. All variations in the reduced subset were manually checked to ensure that the variation call was accurate.

Western blotting
Yeast cells were collected, washed with water and lysed with SDS loading buffer before being resolved on a 5% polyacrylamide gel. Proteins were transferred to nitrocellulose electrophoretically and probed with monoclonal antibodies against Pgk1 (2C12; Invitrogen, Inc.) or Nab3 (2F12-2H2; Dr M. Swanson, U. Fla.). Signal was detected using horse radish peroxidase-conjugated anti-mouse IgG (Amersham-General Electric, Co.), chemiluminescent substrate (Pierce, Inc.) and XOMAT film (Kodak).

RESULTS
A genetic screen for terminator function
We previously characterized IMD2’s IT, an ≈100bp element that blocks elongation and provokes exosome-mediated degradation of the resulting short RNAs (3,37). When placed downstream of the GALI promoter, the IT is a potent terminator. cis-Acting mutations were identified in the IT that inactivated its activity allowing full length RNA synthesis (4,37). We reasoned that including the terminator between an inducible promoter and the GFP reading frame could form the basis of a sensitive genetic screen for trans-acting mutations in genes that encode proteins needed for IT function. Loss of terminator activity would be observed as transcriptional read-through yielding fluorescent cells that could be captured by fluorescence-activated cell sorting (FACS) and clonally isolated.

As a proof of principle, we introduced into yeast a plasmid with the IT inserted between the GALI promoter and GFP encoding sequences and plated them on galactose-containing medium. The presence of the IT effectively prevented what was otherwise strong GFP fluorescence (Figure 1, center versus left patches). When we included the version of the IT bearing a point mutation known to give transcriptional read-through (4,37), fluorescence was restored (right patch). Thus, the IT functions as a potent switch for interrupting GFP expression in live yeast cells.

Cells transformed with the GAL-IT-GFP reporter (Figure 1, center patch) were induced with galactose and subjected to FACS sorting. We expected that cells with spontaneous, trans-acting mutations that reduce IT function, would yield mutants that score positive for fluorescence. From $10^7$ starting cells, we collected ≈0.1% of those with the strongest fluorescent signal. Figure 2A shows the profile of the sorted (red) and post-sorted (blue) populations. The latter group of cells was plated onto galactose-containing media and seven of the strongest fluorescent colonies were pursued (tov1–tov7). The fluorescence intensity of a population grown from each founder colony was quantified by flow cytometry (Figure 2B). For most isolates, a bimodal pattern was obtained in which a peak at ≈$10^2$ fluorescent units was seen along with a strong fluorescent signal at ≈$10^3$–$10^4$ units. The former represents the background of endogenous fluorescence from live yeast cells without a GFP-containing plasmid (data not shown). This peak is also the signal from cells of the starting strain that contains the terminator-repressed GFP plasmid, demonstrating the efficiency of the IT (Figure 2B, ‘WT’, blue-green line and 2C, ‘WT’ red line). The tov mutants clearly gained GFP fluorescence relative to the wild-type population. A small peak of cells with a fluorescence value of ≈$10^2$ in the tov mutant populations are non-expressers. This subset results from reporter plasmid loss that has been described previously (38,39). Even when this group is included in the calculation, mean fluorescence from each of the tov populations was substantially increased over the parent strain, with the maxima of positive peaks showing values of $>10^3$ units for tov2 and tov6 and $>10^4$ fluorescent units for the remaining strains (Figure 2B). These latter values closely approached the level of expression of a wild-type cell with a reporter plasmid that has no terminator placed between the GALI promoter and GFP (data not shown). To ensure that the genetic changes observed were not linked to the reporter plasmid, strains were cured of their URA3-marked plasmids by growth on FOA and retransformed with fresh plasmid. All tov strains fluoresced upon retransformation (data not shown).

To obtain a more accurate measure of mean fluorescence intensity differences, we integrated the GAL-IT-GFP reporter into a relatively gene-free region of chromosome XII for four of the tov strains, tov1, tov2, tov7 and tov4. This obviated the need for an extrachromosomal reporter plasmid and eliminated the problem of plasmid loss. As a result, bimodality disappeared and single peaks were obtained from populations that also displayed lower fluorescent intensities due to the strictly single copy nature of the integrated transcription units (Figure 2C). These four mutant strains displayed mean intensities 6 (tov2)–37 (tov1)-fold higher than the parental wild-type strain. All four tov strains retained full expression of wild GFP in the absence of the IT, ruling out an indirect cause for fluorescent changes, such as a defective galactose response (data not shown).

None of the isolates displayed a temperature-sensitive growth phenotype at 37°C (data not shown). When the
tov1, tov2, tov3 and tov4 mutant haploid strains were mated to an isogenic wild-type strain, GFP fluorescence was lost in the presence of galactose, demonstrating that the mutations were recessive (data not shown).

**Termination and/or RNA degradation is defective in tov mutants**

To test if pol II in these GFP-positive mutants bypassed the IT, we analyzed RNA derived from the reporter plasmid by northern blotting with a GFP probe. In wild-type cells, placement of the IT between the GAL1 promoter and GFP effectively turned off mRNA production (Figure 3, left two panels GAL-GFP versus GAL-IT-GFP), whereas in all the tov isolates, expression was restored following galactose induction, albeit to varying levels. These results are consistent with the fluorescence data suggesting that the mutants suffered from terminator override at the level of transcription.

![Figure 2](image)
For most of the other proteins that couple exosomal degradation to termination. IMD2 through the endogenous tions (Figure 4A), consistent with the idea that pol II reads growth (‘no addition’) and repressed (‘+ guanine’) condi-

tions of endogenous IMD2 length detect the short upstream CUT (Figure 1A), as well as full

short upstream transcript is made but normally degraded). with 0.5 mM guanine (the repressed state in which the

wild-type was observed during standard vegetative

growth. For tov6 cells, IMD2’s short upstream RNA was strongly

stabilized, similar to what was seen previously for an

RRP6-deletion strain (3). This result suggested that tov2 and tov6 harbored mutations in exosomal subunits or proteins that couple exosomal degradation to termination. For most of the other tov strains, a readily detectable increase in the level of full length IMD2 mRNA over wild-type was observed during standard vegetative growth (‘no addition’) and repressed (‘+guanine’) conditions (Figure 4A), consistent with the idea that pol II reads the through the endogenous IMD2 terminator in these mutants.

To test if the mutations resulted in changes in expression of endogenous IMD2, northern blots were probed to detect the short upstream CUT (Figure 1A), as well as full length IMD2 mRNA, as described previously (3). RNA was prepared from untreated cells and from cells grown with 0.5 mM guanine (the repressed state in which the short upstream transcript is made but normally degraded). Cells treated with 15 μg/ml mycophenolic acid, a drug that inhibits IMP dehydrogenase and reduces GTP levels (the induced state), were also analyzed (Figure 4A). For tov2 and tov6 cells, IMD2’s short upstream RNA was strongly stabilized, similar to what was seen previously for an RRP6-deletion strain (3). This result suggested that tov2 and tov6 harbored mutations in exosomal subunits or proteins that couple exosomal degradation to termination. For most of the other tov strains, a readily detectable increase in the level of full length IMD2 mRNA over wild-type was observed during standard vegetative growth (‘no addition’) and repressed (‘+guanine’) conditions (Figure 4A), consistent with the idea that pol II reads through the endogenous IMD2 terminator in these mutants.

To test the generality of these observations, we stripped and re-probed the northern filter to examine a small non-coding RNA, snR33, whose biogenesis is known to involve the Nrd1-Nab3-SEN1 machinery (14,16,17). The resulting terminated RNA is a precursor that is trimmed resulti

ng terminated RNA is a precursor that is trimmed

further downstream than

truncated after its first 264 amino acids. This truncated

protein would lack its nucleotide binding domain and

needed for polyadenylation activity (41,42). The

strain had a C–G substitution in

that changed

TRF4 strain had a C–G substitution in

domain called the ‘central’ region that is conserved

between canonical polyA polymerases (41). Both

truncated protein would lack its nucleotide binding domain and one of the three aspartic acids of the catalytic triad

needed for polyadenylation activity (41,42). The

TRF4 strain had a C–G substitution in

that changed

TRF4 strain had a C–G substitution in

protein function to both strains (orange and red lines, Figure 5A). Sequencing revealed that, relative to its parent strain or the Saccharomyces Genome Database reference sequence, tov2 and tov6 have alterations in the coding sequence of TRF4, a subunit of the TRAMP complex. Tov2 had an additional adenine in a tract of 12 A-residues in TRF4, which produces a frame shift resulting in a Trf4 protein truncated after its first 264 amino acids. This truncated protein would lack its nucleotide binding domain and one of the three aspartic acids of the catalytic triad needed for polyadenylation activity (41,42). The

tov2 strain had a C–G substitution in TRF4 that changed

SerR380 to a stop codon. This truncates the protein

further downstream than tov2’s frame shift in the domain called the ‘central’ region that is conserved between canonical polyA polymerases (41). Both changes were confirmed after amplifying the region and sequencing the products by conventional Sanger methods.

If these mutations were causal for the terminator override phenotype, then a copy of wild-type TRF4 on a plasmid should extinguish the fluorescence of the tov2 and tov6 strains containing the GAL-IT-GFP reporter plasmid. Indeed, the wild-type TRF4 restored terminator function to both strains (orange and red lines, Figure 5A). The tov4 alleles were Amplified from tov2 and tov6, inserted into plasmids and tested for their abilities to cross-complement the tov6 and tov2 strains, respectively. Neither mutant allele could rescue the GFP-expressing phenotype of the other (blue and green lines, Figure 5A), yielding cells with strong fluorescence levels similar to that of cells with empty plasmids (data not shown). These results provide evidence that the tov2 and tov6

Figure 3. Northern blot of GFP mRNA. Cells were grown in the presence or absence of galactose as indicated. Total RNA was isolated from the indicated primary tov strains, the isogenic wild-type strain containing the GAL-IT-GFP plasmid (DY1514) or isogenic wild-type strain with a GAL-GFP reporter plasmid lacking the IT (DY1513). The agarose gel was blotted to a membrane and the filter was probed with a radio-labeled PCR product containing the GFP reading frame. The 12 lanes on the left were from a single gel and filter. The six lanes on the right were from a different single gel and filter (the origin of the higher mobility, cross-hybridizing band seen in all lanes is unknown).
mutations are two different, non-complementing, loss-of-function, mutations in TRF4. If tov2 and tov6’s phenotypes were due to their respective mutations in TRF4, then a complete deletion of TRF4 should yield GFP-positive cells by flow cytometry after transformation with the GAL-IT-GFP reporter plasmid. This was indeed the case (orange versus blue, Figure 5B). Wild-type TRF4 on a plasmid restored substantial IT function (red). This provided independent evidence that the nonsense mutations in TRF4 found in tov2 and tov6 were causally involved in their terminator read-through phenotype.

TOV genes and pol II’s CTD: tov3 and tov7 are allelic to SSU72 and tov5 is allelic to PCF11

Whole genome sequencing also revealed that the tov3 and tov7 strains both harbored the same nucleotide substitution in SSU72, an RNA polymerase II phosphatase previously implicated in termination. This was indeed the case (orange versus blue, Figure 5B). Wild-type TRF4 on a plasmid restored substantial IT function (red). This provided independent evidence that the nonsense mutations in TRF4 found in tov2 and tov6 were causally involved in their terminator read-through phenotype.

TOV genes and pol II’s CTD: tov3 and tov7 are allelic to SSU72 and tov5 is allelic to PCF11

Whole genome sequencing also revealed that the tov3 and tov7 strains both harbored the same nucleotide substitution in SSU72, an RNA polymerase II phosphatase previously implicated in termination. The A-T change was confirmed by Sanger sequencing of a PCR product from tov3 and tov7 genomic DNA and was unique to these two strains. The resulting change converts leucine84 to a phenylalanine. To prove that the ssu72 mutation was causal, we introduced plasmids containing either the wild-type or tov3 (= tov7) allele of SSU72 into the tov7 strain with the GAL-IT-GFP reporter plasmid. Strong termination was restored to the tov strain by the wild-type, but not the mutant, copy of SSU72 (Figure 6A, red versus blue). Similar results were observed for the tov3 strain containing an integrated copy of the GAL-IT-GFP reporter (data not shown).

We also tested if an independently generated, conditional allele of SSU72 (degron strain H291; courtesy of M. Hampsey) (Table 1), showed read-through of the IMD2 IT. Following a shift to the non-permissive temperature, cells containing the GAL-IT-GFP reporter were GFP positive relative to control cells (6B). Together the experiments of Figure 6 show that Ssu72 functions at IMD2 and is the causal basis for terminator override in tov3 and tov7.

Similarly, genome sequencing revealed that the tov5 strain contained a point mutation in PCF11 gene that was not found in any other tov strains. PCF11 is an essential gene that encodes a termination-related protein known to associate with the pol II large subunit’s CTD. Conventional Sanger sequencing confirmed this change in tov5, which changes tyrosine94 to serine. A wild-type copy of PCF11 on a plasmid was effective at reversing GFP reporter expression when transformed into a tov5 strain (Figure 6C). Both PCF11 and SSU72 strongly implicate the phosphorylation state of the CTD in IMD2’s regulation via the IT.

tov1 and tov4 are allelic to NAB3 and reveal an important function for the protein’s extreme carboxy-terminal residues

Early in the process of characterizing the tov genes, we explored the possibility that mutations fell in genes previously known to control termination and processing of
small RNAs such as CUTs and snoRNAs. We selected SEN1, RRP6, NAB3 and NRD1 as candidate open reading frames for direct conventional sequencing using individually amplified regions from the genomic DNA of some of the tov strains. All strains were wild-type for these except tov1 and tov4 that showed novel TAA nonsense mutations in NAB3. Both fell within a stretch of CAA codons, encoding a tract of 16 glutamine residues near the carboxy terminus. This resulted in the loss of 19 and 27 amino acids from the end of carboxy-terminal end of Nab3 (Figure 7). Thus, it appeared that there was a short region of Nab3’s carboxyl-terminal tail that was functionally important for termination. Interestingly, 15 of the 16 carboxy-terminal residues are completely conserved among six *Saccharomyces* species (asterisks), even though the adjacent sequence, including the glutamine repeat tract, is somewhat variable (Figure 7).

To confirm that the terminator override phenotype was due to the tov1 mutation resulting in the loss of 19 carboxy-terminal amino acids (henceforth referred to as nab3-Q784X), we transformed a diploid heterozygous for a deletion of NAB3, with a plasmid encoding the nab3-Q784X allele expressed from its own promoter. The diploid was sporulated and a haploid was selected that bore the deletion of chromosomal NAB3 and contained the plasmid expressing nab3-Q784X. When this strain was transformed with the GAL-IT-GFP reporter plasmid to test for terminator function, GFP was expressed at levels comparable to the original mutant strain, providing strong evidence that this genetic change alone was sufficient to yield the terminator override phenotype (Figure 8A, left sector). In addition, when wild-type NAB3 was introduced on a plasmid, terminator function (loss of GFP) was restored to both the tov1 (Figure 8A, right sector) and tov4 (data not shown) strains. Genetic complementation was also observed when tov1 cells were mated to a tov2 strain in which NAB3 was wild-type (data not shown).

To define how much of this carboxy-terminal Nab3 ‘tail’ is required for termination, we constructed a nested set of plasmids that deleted the last 3, 6, 9, 12 or 15 amino acids near the carboxy terminus of Nab3, encoding a tract of 16 glutamine residues near the carboxy terminus. This resulted in the loss of 19 and 27 amino acids from the end of carboxy-terminal end of Nab3 (Figure 7). Thus, it appeared that there was a short region of Nab3’s carboxyl-terminal tail that was functionally important for termination. Interestingly, 15 of the 16 carboxy-terminal residues are completely conserved among six *Saccharomyces* species (asterisks), even though the adjacent sequence, including the glutamine repeat tract, is somewhat variable (Figure 7).

To confirm that the terminator override phenotype was due to the tov1 mutation resulting in the loss of 19 carboxy-terminal amino acids (henceforth referred to as nab3-Q784X), we transformed a diploid heterozygous for a deletion of NAB3, with a plasmid encoding the nab3-Q784X allele expressed from its own promoter. The diploid was sporulated and a haploid was selected that bore the deletion of chromosomal NAB3 and contained the plasmid expressing nab3-Q784X. When this strain was transformed with the GAL-IT-GFP reporter plasmid to test for terminator function, GFP was expressed at levels comparable to the original mutant strain, providing strong evidence that this genetic change alone was sufficient to yield the terminator override phenotype (Figure 8A, left sector). In addition, when wild-type NAB3 was introduced on a plasmid, terminator function (loss of GFP) was restored to both the tov1 (Figure 8A, right sector) and tov4 (data not shown) strains. Genetic complementation was also observed when tov1 cells were mated to a tov2 strain in which NAB3 was wild-type (data not shown).

To define how much of this carboxy-terminal Nab3 ‘tail’ is required for termination, we constructed a nested set of plasmids that deleted the last 3, 6, 9, 12 or 15 amino acids near the carboxy terminus of Nab3, encoding a tract of 16 glutamine residues near the carboxy terminus. This resulted in the loss of 19 and 27 amino acids from the end of carboxy-terminal end of Nab3 (Figure 7). Thus, it appeared that there was a short region of Nab3’s carboxyl-terminal tail that was functionally important for termination. Interestingly, 15 of the 16 carboxy-terminal residues are completely conserved among six *Saccharomyces* species (asterisks), even though the adjacent sequence, including the glutamine repeat tract, is somewhat variable (Figure 7).

To confirm that the terminator override phenotype was due to the tov1 mutation resulting in the loss of 19 carboxy-terminal amino acids (henceforth referred to as nab3-Q784X), we transformed a diploid heterozygous for a deletion of NAB3, with a plasmid encoding the nab3-Q784X allele expressed from its own promoter. The diploid was sporulated and a haploid was selected that bore the deletion of chromosomal NAB3 and contained the plasmid expressing nab3-Q784X. When this strain was transformed with the GAL-IT-GFP reporter plasmid to test for terminator function, GFP was expressed at levels comparable to the original mutant strain, providing strong evidence that this genetic change alone was sufficient to yield the terminator override phenotype (Figure 8A, left sector). In addition, when wild-type NAB3 was introduced on a plasmid, terminator function (loss of GFP) was restored to both the tov1 (Figure 8A, right sector) and tov4 (data not shown) strains. Genetic complementation was also observed when tov1 cells were mated to a tov2 strain in which NAB3 was wild-type (data not shown).

To define how much of this carboxy-terminal Nab3 ‘tail’ is required for termination, we constructed a nested set of plasmids that deleted the last 3, 6, 9, 12 or 15 amino acids near the carboxy terminus of Nab3, encoding a tract of 16 glutamine residues near the carboxy terminus. This resulted in the loss of 19 and 27 amino acids from the end of carboxy-terminal end of Nab3 (Figure 7). Thus, it appeared that there was a short region of Nab3’s carboxyl-terminal tail that was functionally important for termination. Interestingly, 15 of the 16 carboxy-terminal residues are completely conserved among six *Saccharomyces* species (asterisks), even though the adjacent sequence, including the glutamine repeat tract, is somewhat variable (Figure 7).

To confirm that the terminator override phenotype was due to the tov1 mutation resulting in the loss of 19 carboxy-terminal amino acids (henceforth referred to as nab3-Q784X), we transformed a diploid heterozygous for a deletion of NAB3, with a plasmid encoding the nab3-Q784X allele expressed from its own promoter. The diploid was sporulated and a haploid was selected that bore the deletion of chromosomal NAB3 and contained the plasmid expressing nab3-Q784X. When this strain was transformed with the GAL-IT-GFP reporter plasmid to test for terminator function, GFP was expressed at levels comparable to the original mutant strain, providing strong evidence that this genetic change alone was sufficient to yield the terminator override phenotype (Figure 8A, left sector). In addition, when wild-type NAB3 was introduced on a plasmid, terminator function (loss of GFP) was restored to both the tov1 (Figure 8A, right sector) and tov4 (data not shown) strains. Genetic complementation was also observed when tov1 cells were mated to a tov2 strain in which NAB3 was wild-type (data not shown).
residues of the protein. The tov1 strain was transformed with each of these and tested for GFP expression. None could rescue termination function; all constructs yielded strongly fluorescent cells (data not shown). This suggests that the loss of as few as three of Nab3’s carboxy-terminal amino acids sufficed to generate the tov phenotype. This was independently confirmed for the construct lacking the carboxy-terminal three amino acids when it was expressed in a strain deleted for the chromosomal copy of NAB3 (Figure 8B, blue versus red). To refine further the minimum portion of the carboxy-terminal tail that supports efficient termination, we generated Nab3

Figure 6. Terminator override in ssu72 (tov7) and pcf11 (tov5) mutants. (A) A tov7 strain transformed with the GAL-IT-GFP reporter plasmid and either a plasmid with a wild-type copy of SSU72 (DY3625) or its L84F allele (‘ssu72-2D’; DY3626), were induced with galactose and subjected to flow cytometry. (B) Yeast strains containing an intact (DY3114) or degron derivative (DY3115) of SSU72 were transformed with the GAL-IT-GFP reporter plasmid, induced with galactose and analyzed by flow cytometry. Each curve is the result of averaging five runs of 10 000 cells each. (C) A tov5 strain was transformed with the GAL-IT-GFP reporter plasmid and either a plasmid with a wild-type copy of PCF11 (DY3628) or the empty pRS315 plasmid (DY3627) was induced with galactose and analyzed by flow cytometry.
Figure 7. Map of Nab3 with expansion of the carboxy-terminal sequence and its alignments. A schematic map (not to scale) of the domains of *S. cerevisiae* Nab3 is shown at the top with the approximate positions of the D/E- and P/Q-rich regions. The RRM piece analyzed by X-ray crystallography (10) is delineated as are the positions of the missense changes in the nab3-11 mutant. A CLUSTALW alignment of the indicated regions of Nab3 from *S. cerevisiae* (boxed), *S. bayanus*, *S. paradoxus*, *S. kluveri*, *S. kudriavzevii* and *S. mikatae*, is shown in the centre (43,44). Shading designates identities (yellow), strong similarities (pink) and weak similarities (green). The positions of nonsense mutations (stop signs) for the tov1 and tov4 isolates are shown above the alignment. The structural homology alignment generated by the PHYRE algorithm is shown for the known structure of a portion of hnRNP C (template) and the Nab3 tail (query).

Figure 8. Testing of wild-type and mutant NAB3 function. (A) tov1 strains containing the GAL-IT-GFP reporter plasmid and either an empty control vector control (DY3018; top sector) or a plasmid with the NAB3 gene (DY3019; right sector), were grown on galactose along with a strain (DY2240; left sector) deleted for chromosomal NAB3 and containing a plasmid expressing the tov1 mutation of NAB3 (nab3-Q784X). The plate was photographed under blue light to detect GFP fluorescence. (B) A yeast strain lacking its chromosomal copy of NAB3 and containing the GAL-IT-GFP reporter plasmid was transformed with a second plasmid containing the wild-type (‘NAB3FL’; DY3036), the 3D (‘nab3C31’; DY3060) or the 1D (‘nab3C31’; DY3061) derivatives of NAB3, as indicated. Cells were grown in the presence of galactose and analyzed by flow cytometry. (C) A strain with the nab3-11 allele was transformed with a plasmid containing either the wild-type NAB3 (DY3019), the tov1 allele nab3-Q784X (DY3012), the nab3-11 allele (DY3011) or an empty vector (DY3013). Cultures were serially diluted 10-fold and the dilutions were spotted on selective medium for growth at 22°C or 30°C, as indicated. (D) Western blot for Nab3 in wild-type (DY1514F) and the indicated mutant strains (1A1F = ‘nab3-Q784X’, 4A1F = ‘nab3-Q776X’, DY30229 = ‘nab3-11’). Pgk1 served as a loading control. The nab3-11 strain was grown at the permissive temperature of 22°C.
expressing constructs lacking the last one or two amino acids (Figure 8B, orange and green). These versions of Nab3 were partially defective for termination as reflected by their fluorescence profile. This result indicates that the last amino acid in the Nab3 polypeptide is functionally important and the final three amino acids are critical for Nab3 to function properly in termination.

Previous studies of Nab3’s function exploited the temperature-sensitive allele nab3-11, which has two amino acid substitutions (F371L and P374T) in the protein’s RNA binding domain (Figure 7, top) (4,8,19). To test if the nab3-Q784X mutation could be distinguished from this previously characterized mutation, we transformed a nab3-11 strain with the plasmid expressing nab3-Q784X. Plasmids containing wild-type NAB3 or nab3-Q784X rescued the temperature sensitivity seen in the parent nab3-11 strain (Figure 8C). This result demonstrates that the nab3-Q784X protein is thermostable and has a functional deficit distinct from that of the well-studied nab3-11 mutant that has an altered RRM.

To learn if the truncated proteins encoded by the tov1 and tov4 alleles of NAB3 were stable proteins (nab3-Q784X and nab3-Q776X, respectively), we carried out western blotting using a previously described monoclonal antibody against Nab3 (11). Both truncated Nab3 derivatives (tov1 and tov4) were at least as abundant as wild-type Nab3 (Figure 8D). The mutant proteins also migrated faster than wild-type Nab3 consistent with their slightly shorter lengths.

**DISCUSSION**

**IMD2** expression is regulated by intracellular guanine nucleotides through an IT that can be considered a eukaryotic attenuator (1–4). In a set of proof-of-principle experiments, we have exploited a genetic screen to identify alleles of genes encoding trans-acting factors that operate this terminator. Prior studies by us and others have exploited GFP fluorescence and copper resistance to identify cis-acting sequences involved in termination, as well as start site choice at IMD2 (4,37). Here, we observed that the fluorescence sorting approach was sensitive enough to identify spontaneous mutations in at least four trans-acting genes known to participate in small RNA biogenesis and metabolism, TRF4, SSU72, PCF11, and NAB3. This set of base and amino acid changes are summarized in Table 3. It is doubtful that this screen has been saturated and it is likely that mutagenesis of our starting strain could provide additional genes and alleles that can add to our understanding of the function of this, or even other, terminators. In addition, high-throughput sequencing of genomes from a number of mutants was effective at identifying causal mutations. Whole genome sequencing reveals numerous strain to strain differences, hence identification of genes expected to affect the process under study is straightforward, as observed here through complementation with plasmid-borne wild-type candidates. Finding genes whose role may be unforeseen would be more difficult, requiring, for example, complementation with a genomic library. The four tov genes identified here represent the following components of the transcription termination machinery: (i) Trf4, part of the TRAMP complex, (ii) Ssu72, an enzyme that modulates RNA polymerase II’s CTD’s phosphorylation state, (iii) Pcf11, a protein that binds to phospho-CTD and (iv) Nab3, an RNA binding, hnRNP-like protein. Further study of the individual resulting mutant proteins should be of mechanistic value in understanding the distinct regulated termination event at IMD2.

The TRAMP complex links the termination and ribonuclease machineries by directing terminated RNAs for degradation via the unconventional polyadenylation activity of the Trf4 polyA polymerase (29). It facilitates termination and exosomal hydrolysis of transcripts through a physical association with both the Nrd1–Nab3 and exosomal protein complexes (30). Deletion of TRF4 stabilizes a number of CUTs (7,29,30). Our results confirm that the loss of TRF4 takes the IMD2 CUT out of the degradation pathway (Figure 4). In addition, termination read-through was seen in the trf4 mutants examined here, giving the impression that TRF4 influences both the efficiency of CUT degradation and the tightness of termination. The appearance of this read-through transcript could be due to (i) the sparing of a constitutively produced, extended (non-terminated) RNA that would otherwise be degraded if TRAMP was intact and/or (ii) loss of termination efficiency because TRAMP modifies the termination reaction, perhaps through its interaction with Nab3-Nrd1. A less likely possibility is that the trf4 mutation shifts initiation from the upstream to the downstream start site as an alternative means of bypassing the IT.

The Ssu72 phosphatase and the phosphorylation state of the CTD of the large subunit of pol II, are implicated in transcription termination at a number of genes encoding small non-coding RNAs and mRNAs (13,16,17,25,45–49). While a number of SSU72 mutations have been identified by genetic screens, none have yielded the change we found here (L84F). The atomic level structure of Drosophila Ssu72 shows that the cognate leucine (leu82) lines the groove that accommodates the pol II CTD substrate (49). When site-directed mutagenesis was used to probe the importance of the groove residues by changing them to alanine, the substitution of leu82 did not impact phosphatase activity toward a CTD peptide. We suggest that a change to the larger phenylalanine residue obtained in the screen reported here, would influence peptide

<table>
<thead>
<tr>
<th>tov isolate</th>
<th>Candidate gene</th>
<th>Base change</th>
<th>Amino acid change</th>
</tr>
</thead>
<tbody>
<tr>
<td>tov1</td>
<td>NAB3</td>
<td>C2350T</td>
<td>Q784X</td>
</tr>
<tr>
<td>tov2</td>
<td>TRF4</td>
<td>A795ins</td>
<td>K264fs</td>
</tr>
<tr>
<td>tov3</td>
<td>SSU72</td>
<td>A252T</td>
<td>L84F</td>
</tr>
<tr>
<td>tov4</td>
<td>NAB3</td>
<td>C2326T</td>
<td>Q776X</td>
</tr>
<tr>
<td>tov5</td>
<td>PCF11</td>
<td>A281C</td>
<td>Y94S</td>
</tr>
<tr>
<td>tov6</td>
<td>TRF4</td>
<td>C1289G</td>
<td>S430X</td>
</tr>
<tr>
<td>tov7</td>
<td>SSU72</td>
<td>A252T</td>
<td>L84F</td>
</tr>
</tbody>
</table>

Table 3. tov Gene identification
binding and phosphatase activity. In any case, this report provides biological evidence in favor of the importance of this substrate binding site residue. Our finding contrasts with an earlier study using a copper resistance assay in which a mutation in SSU72 known to affect termination did not strongly perturb the IMD2 CUT terminator’s function (4). A higher sensitivity of the fluorescent assay presented here could be part of the basis for this difference. One hypothesis for the L84F mutation’s impact upon termination is that the accumulation of ser5-CTD phosphorylation when phosphatase activity is impaired, relaxes the coupling of Nab3–Nrd1’s association with the IMD2 CUT in preparation for termination, resulting in higher read-through of the terminator.

The structure of Pcf11 and the details of its binding to pol II’s CTD have also been studied at the atomic level (50,51). The CTD-interaction domain of Pcf11 contains an eight helix hydrophobic core. The tov5 mutation of PCF11 is a new allele that changes a tyrosine in a conserved position important for this helical bundle’s packing, to a serine. The protein’s folding is likely to be changed; whether this results in a reduction of Pcf11 levels and/or interferes with CTD binding, remains to be determined. Pcf11 is known to play a role in transcription termination for mRNA and non-coding RNAs (17,52–55). Recent work has shown that it associates with IMD2 and the A66D allele (Pcf11-9) disrupts termination at the IMD2 CUT (48). The two phospho-CTD binding proteins, Nrd1 and Pcf11 influence termination in a complex manner in response to a gene-specific CTD modification code (48). Further work is needed to learn if Nrd1 and Pcf11 cooperate at this terminator.

The most interesting of the tov mutations was that found in Nab3, a protein whose study has been hampered by the inability to produce a full length recombinant protein (12). NAB3 was previously demonstrated to be involved in termination at IMD2’s IT using the well-studied nab3-11 temperature-sensitive allele (4). However, data presented herein draw attention to a distinct and very small part of Nab3 that has not previously been recognized as important, but which is required for termination at IMD2. The critical RRM of Nab3 and its Nrd1 interaction region are located approximately centrally in the 802 amino acid protein (Figure 7) (8). They lie in the most evolutionarily conserved part of the protein in which lethal mutations have been described (8,9). Biochemical and structural analyses have shown that the RNA binding domain forms a canonical ββββββββ RRM sufficient for high-affinity binding to RNA when paired with a corresponding RRM from Nrd1 (9,10,12,18). Difficulty in expressing full length recombinant Nab3, probably due to its extensive polyQ and polyD/E tracts, has meant there has been little, if any, biochemical analysis of either portion of the protein flanking the RRM, much less the complete protein (12). Hence, the function of the bulk of the Nab3 protein remains mysterious. The spontaneous tov1 and tov4 strains is not required for cell viability, it is important for the termination reaction and optimal cell proliferation (data not shown). Precisely what the role of this tail region is remains to be learned. It is unlikely to bind RNA (9). A clue to its function is offered by the result of computational analyses using the PHYRE and HHpred algorithms, which assess structural relatedness (56,57). Both analyses predict an α-helical structure for the Nab3 tail and the best fit was to a similar region conserved in the hnRNP C family of proteins (Figure 7 and data not shown). Evidence from NMR and biophysical studies indicates that this element of the human hnRNP C protein forms a tetrameric α-helical bundle (58). This suggests that the CTD-interaction domain of Nab3 is a protein–protein interaction domain, perhaps one that enables Nab3 to multimerize with itself much the way hnRNP C can (59,60). It is intriguing to consider this possibility given that substrate RNAs often contain multiple tandem Nrd1 and Nab3 recognition sites (18). The organization of such sites is variable between different RNA substrates and the sites often appear redundant by mutational analysis (18). Perhaps the degree of multimerization of the Nab3–Nrd1 complex on RNA accounts for some of the variability between transcripts in their dependence upon Nrd1 and Nab3 for termination or processing. The concept of polymerization of Nrd1–Nab3 dimers on RNA was presented previously based upon (i) the multiplicity of RNA binding sites per transcript and (ii) the repeats in the CTD that may bind several copies of Nrd1 in the vicinity of the nascent transcript (12). Self association of the Nab3 tail could offer an additional source of scaffolding through which an array of Nrd1–Nab3 dimers assemble onto RNA, potentially forming a higher order packaging of this yeast ribonucleoprotein.

Regulation of IMD2 expression has proven useful as a model to study the Nrd1–Nab3–Sen1 termination system, CTD phosphorylation and the degradation of CUTs. Additional work on the mutations reported here should shed light on this complex mechanism.

ACKNOWLEDGEMENTS

We thank R. Karaffa and S. Durham of the Emory FACS Core Facility for cytometry services, the Emory GRA Genome Center for sequencing services,
REFERENCES


