

Uniform Distribution of Elongating RNA Polymerase II Complexes in Transcribed Gene Locus*

Received for publication, February 14, 2011, and in revised form, May 6, 2011. Published, JBC Papers in Press, May 23, 2011, DOI 10.1074/jbc.M111.230805

Kadri Peil, Signe Värvi, Marko Lööke, Kersti Kristjuhan, and Arnold Kristjuhan¹

From the Institute of Molecular and Cell Biology, University of Tartu, Riia 23, Tartu 51010, Estonia

The intensity of gene transcription is generally reflected by the level of RNA polymerase II (RNAPII) recruitment to the gene. However, genome-wide studies of polymerase occupancy indicate that RNAPII distribution varies among genes. In some loci more polymerases are found in the 5' region, whereas in other loci, in the 3' region of the gene. We studied the distribution of elongating RNAPII complexes at highly transcribed *GAL-VPS13* locus in *Saccharomyces cerevisiae* and found that in the cell population the amount of polymerases gradually decreased toward the 3' end of the gene. However, the conventional chromatin immunoprecipitation assay averages the signal from the cell population, and no data on single cell level can be gathered. To study the spacing of elongating polymerases on single chromosomes, we used a sequential chromatin immunoprecipitation assay for the detection of multiple RNAPII complexes on the same DNA fragment. Our results demonstrate uniform distribution of elongating polymerases throughout all regions of the *GAL-VPS13* gene.

Transcription of protein-coding genes by RNA polymerase II (RNAPII)² is highly regulated to respond properly to external and internal signals of cell growth and to the changes of cell environment. More than two thirds of constitutively expressed genes in budding yeast are transcribed at low levels. However, a set of genes, like those encoding histones, ribosomal proteins or glycolytic enzymes, are highly transcribed (1–3). In addition, transcription of many genes is rapidly up-regulated in response to various stress signals or changes in carbon sources in growth media, leading to the initiation of a new mRNA transcript in every 6–8 s (4). However, most of the constitutively expressed genes showed only a single transcript in the locus when the number of nascent mRNAs was analyzed in individual yeast cells (5).

Although the amount of RNAPII complexes on the gene locus is generally proportional to the rate of gene transcription, the average distribution of polymerases on actively transcribed loci is somewhat uneven. For example, in response to gene activation considerably higher amount of RNAPII is detected in the 5' than in the 3' region of the galactose-inducible *GAL-VPS13*

gene in budding yeast (6). Also, several studies of gene-specific and genome-wide distributions of RNAPII indicate that in some loci RNAPII is preferably located in the 5' region of the gene, whereas in other loci the elevated amounts of polymerases were detected in the 3' region or in both ends of genes (7–10).

The gradual loss of polymerase signal along the *GAL-VPS13* gene suggests that either the elongation rate of RNAPII accelerates toward the end of the gene, or alternatively, only a fraction of cells express full-length transcript, whereas others abort transcription at random positions in the gene. In the latter case, chromatin immunoprecipitation (ChIP) of RNAPII from the total lysate of cell population leads to deceptive enrichment of RNAPII in the 5' region because most of the cells have polymerases at the beginning of the gene, but only some of them have RNAPII also at the end of the gene.

To analyze the distribution of elongating RNA polymerases in different regions of a highly transcribed gene, we generated a model strain of *Saccharomyces cerevisiae* that simultaneously expresses RNAPII complexes with different epitope tags. We used sequential chromatin immunoprecipitation assay (re-ChIP) to analyze whether multiple RNAPII molecules co-occupy the same DNA fragment during transcription elongation and whether the spacing of polymerases differs in the 5' and 3' regions of actively transcribed *GAL-VPS13* locus. Our results indicate that elongating RNAPII complexes are uniformly distributed throughout the entire length of the gene.

EXPERIMENTAL PROCEDURES

Yeast Strains—All *S. cerevisiae* strains were congenic with strain W303 and are listed in Table 1. To generate *GAL-VPS13* genes with insertion of heterologous DNA sequence at different locations in the coding region, fragments of *Kluyveromyces fragilis VPS13* gene (nucleotides 8071–8170, 4165–4364, and 5113–5512) were amplified and ligated together to form a 700-bp fragment. This fragment was inserted into *GAL-VPS13* coding region at positions 151, 3221, 6071, or 9251 bp downstream from the start codon to create strains AKY271, AKY273, AKY275, and AKY276, respectively. In addition, all of these strains contain also triple E2 tags in C terminus of *RPB3* (tagged in the genuine *RPB3* locus) and two additional copies of *RPB3* with C-terminal triple E4 and myc tags (additional genes inserted into *HIS3* and *LEU2* loci, respectively).

ChIP and Re-ChIP Assays—Cells were grown overnight in YP medium containing 2% glucose or galactose as a carbon source before fixation for ChIP assay. For 30 min galactose induction cells were grown overnight in YP medium containing 2% raffinose, which was then replaced by 2% galactose for 30 min. Con-

* This work was supported by Wellcome Trust International Senior Research Fellowship Grant 081756 and European Molecular Biology Organization Installation Grant 1454.

⌘ Author's Choice—Final version full access.

¹ To whom correspondence should be addressed. Tel.: 372-737-5046; Fax: 372-742-0286; E-mail: arnoldk@ut.ee.

² The abbreviations used are: RNAPII, RNA polymerase II; qPCR, quantitative PCR; re-ChIP, sequential ChIP assays; WCE, whole cell extract.

Distribution of Elongating RNAPII

TABLE 1

Yeast strains used in this study

Strain	Genotype	Source
AKY271	W303 MATa GAL-VPS13 (<i>K. lactis</i> -vps13-insert at 0.1 kb)::TRP1 RPB3-3xE4tag::HIS3 RPB3-myc tag::LEU2 RPB3-3xE2tag::URA3	This study
AKY273	W303 MATa GAL-VPS13(<i>K. lactis</i> -vps13-insert at 3 kb)::TRP1 RPB3-3xE4tag::HIS3 RPB3-myc tag::LEU2 RPB3-3xE2tag::URA3	This study
AKY275	W303 MATa GAL-VPS13(<i>K. lactis</i> -vps13-insert at 6 kb)::TRP1 RPB3-3xE4tag::HIS3 RPB3-myc tag::LEU2 RPB3-3xE2tag::URA3	This study
AKY276	W303 MATa GAL-VPS13(<i>K. lactis</i> -vps13-insert at 9 kb)::TRP1 RPB3-3xE4tag::HIS3 RPB3-myc tag::LEU2 RPB3-3xE2tag::URA3	This study
AKY643	W303 MATa GAL-VPS13::TRP1 Rpb3-3xE2tag::natMX6	This study

ventional ChIP assays were performed as described (11). Whole cell extract from 1×10^7 cells was used for ChIP assays with antibodies directed against RNAPII (4H8; Upstate Biotechnology), anti-E4 tag (1E2; Quattromed), anti-myc tag (9E10; Abcam), and anti-E2-tag (5E11; Quattromed). Co-precipitated DNA was analyzed by real-time qPCR using the ABI Prism® 7900HT Fast Real-Time PCR System in standard conditions (40 cycles; 95 °C 15 s + 60 °C 1 min). Absolute qPCR SYBR Green reagents (ABgene and Fermentas) and 5×HOT FIRE-Pol® EvaGreen® quantitative PCR Mix (Solis BioDyne) were used. PCRs were done with primer pairs covering coding regions of *VPS13*, *FBA1*, and *GAL10* genes. The *VPS13* coding region primers amplified the sequences at 0.1, 2.6, 3.5, 5.5, 6.4, 8.5, and 9.3 kb downstream from the start codon. Three sets of primer pairs were used for detection of the *K. lactis* insertion sequence in *VPS13* ORF. The lengths of all qPCR products were about 150 bp. Sequences of all primers are available upon request.

Re-ChIP was adapted from Ref. 12. 5 ml of whole cell extract (WCE) was prepared from 1×10^9 cells (100-ml culture). Cells were fixed in 1% formaldehyde, resuspended in FA-lysis buffer (50 mM HEPES, pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, and protease inhibitors), and lysates were prepared as described (11). The first round of re-ChIP was performed with anti E2 tag antibody (5E11) from 1 ml of lysate, elution from Sepharose-protein A beads was done with 10 mM DTT. One quarter of the sample was taken out for qPCR analysis. The rest of the eluate was diluted 10-fold with FA-lysis buffer and divided into three samples for the second round of re-ChIP with anti-E4 (1E2), anti-myc, or anti-Rpb1 (4H8) antibodies. Co-precipitated DNA was analyzed by qPCR. Nontranscribed regions in the chromosome VIII and the telomeric region of chromosome VI right arm were used to set the level of background RNAPII signal in all immunoprecipitations. All ChIP experiments were repeated at least three times, and error bars represent S.D.

PCR Analysis of DNA Fragment Sizes in the Whole Cell Extract—Genomic DNA from the whole cell extract used in re-ChIP experiments was extracted with phenol:chloroform after proteinase K treatment, precipitated with ethanol, and dissolved in TE (10 mM Tris-HCl, 1 mM EDTA, pH 7.5). PCR with different *VPS13* primers was used to generate 700-, 1000-, 1500-, and 2000-bp fragments to analyze the length of DNA fragments in WCE. Unfragmented *S. cerevisiae* genomic DNA was used as a positive control.

RESULTS

The Average Amount of RNAPII in GAL-VPS13 Locus Decreases toward the End of the Gene—To study the distribution of RNAPII complexes on the coding region of a gene, we

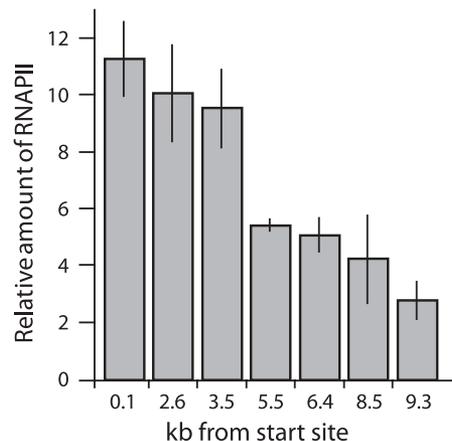


FIGURE 1. RNAPII level at the beginning of an inducible GAL-VPS13 gene is higher than at the end of the gene. Relative amount of RNAPII was analyzed with ChIP and qPCR with primers specific for different positions in *GAL-VPS13*. Samples for ChIP were taken 30 min after transcription induction with galactose and analyzed for RNAPII levels throughout the *VPS13* gene. Relative amount of RNAPII was determined by analyzing samples from cells grown in YP-galactose compared with cells grown in YP-raffinose. Error bars, S.D.

used *GAL-VPS13* as the model locus. This gene contains a galactose-inducible promoter in front of a 9.5-kb-long *VPS13* open reading frame, which allows us to induce or repress the gene with galactose or glucose, respectively (6). To analyze the distribution of elongating RNAPII complexes in more detail, we first generated a *S. cerevisiae* strain where the Rpb3 subunit of RNAPII was epitope-tagged with E2 tag. Using Rpb3 subunit of RNAPII excludes any possible bias of C-terminal domain-specific antibodies that are generally used for ChIP of RNAPII, but might have different affinity toward various phosphorylation forms of the C-terminal domain.

We induced transcription of *GAL-VPS13* in galactose-containing media and analyzed the distribution of RNAPII in the locus. The amount of RNAPII was dependent on the distance from the promoter, gradually diminishing toward the 3' end of the coding region and being approximately four times higher at the beginning of the gene (Fig. 1). As different pairs of PCR primers were used for detection of RNAPII in *GAL-VPS13* locus, it was possible that some of the effects were caused by uneven efficiency of qPCR primers. To analyze the presence of RNAPII at different regions of *GAL-VPS13* without primer bias, we made a panel of yeast strains containing insertions of a 700 bp DNA sequence (originated from *K. lactis* *VPS13* gene; see "Experimental Procedures") at different locations in the coding region of *GAL-VPS13* gene: at 0.1, 3, 6, and 9 kb from the promoter, respectively (Fig. 2A). This set of strains allowed us to monitor the amount of RNAPII on exactly the same DNA sequence at different distances from the promoter, thus excluding all differences that might be caused by the comparison of results obtained with different primers in qPCRs.

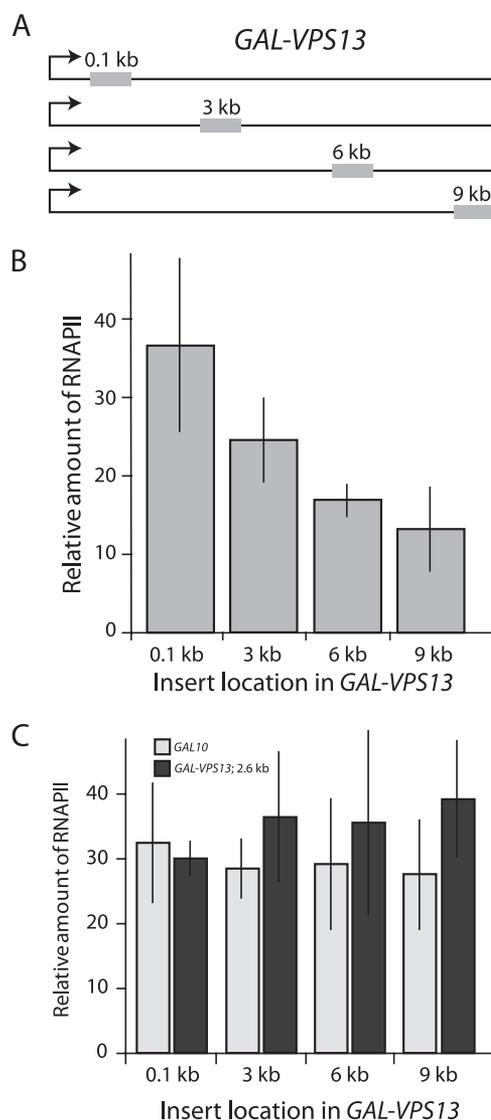


FIGURE 2. The relative amount of RNAPII on the insert DNA is dependent on the distance of the fragment from the promoter. *A*, scheme of yeast strains with *K. lactis* genomic DNA insertions at different locations inside *GAL-VPS13*. *B*, relative amount of RNAPII analyzed with primers specific for the insert DNA. Samples for ChIP were taken after overnight transcription induction with galactose and analyzed for RNAPII levels on the insert DNA in four different yeast strains. Relative amount of RNAPII was determined by analyzing samples from cells grown in YP-galactose compared with cells grown in YP-glucose. *C*, recruitment of RNAPII to *GAL10* and *VPS13* loci determined in yeast strains with insert DNA in *GAL-VPS13* gene. Relative amount of RNAPII was determined with primers specific for *GAL10* coding region and for 2.6-kb region of *VPS13* in the samples used in *B*. Error bars, S.D.

We determined the distribution of RNAPII at various locations throughout the *GAL-VPS13* locus after overnight induction of the gene. In all four strains the relative amount of RNAPII on the DNA insert was dependent on the distance of the fragment from the promoter, being approximately three times higher in the beginning of the gene compared with the end of the gene (Fig. 2*B*). To confirm identical induction of galactose-regulated genes in all four strains, we also compared the amount of RNAPII recruited to the coding region of *GAL10* gene and to 2.6 kb region of *GAL-VPS13* gene. Recruitment of RNAPII to these control regions was very similar in all strains (Fig. 2*C*). These results confirm that the differences of RNAPII

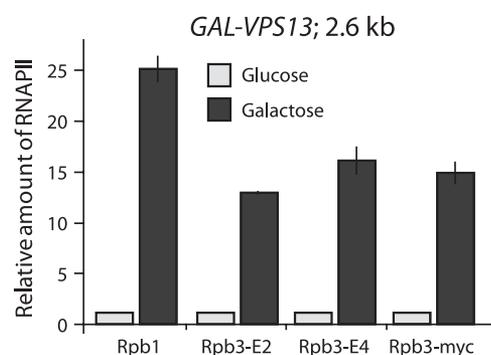


FIGURE 3. RNAPII complexes with different Rpb3 tags are equally recruited to the *GAL-VPS13* locus. Relative amount of RNAPII was analyzed with qPCR with primers specific for 2.6-kb region of *GAL-VPS13* in AKY271 strain. ChIP samples for RNAPII analysis were taken after overnight induction of *GAL-VPS13* with galactose. Relative amount of RNAPII was determined by analyzing samples from cells grown in YP-galactose compared with cells grown in YP-glucose. ChIP was performed with 4H8, 5E11, 1E2, or 9E10 antibodies, recognizing Rpb1, E2-tagged Rpb3, E4-tagged Rpb3, or myc-tagged Rpb3, respectively. Error bars, S.D.

levels at the beginning and at the end of the *GAL-VPS13* (Figs. 1*A* and 2*B*) reflected a real decrease of RNAPII density toward 3' region of the gene that were not caused by different experimental conditions or primers used in ChIP or qPCR assays.

Determination of RNAPII Distribution in the *GAL-VPS13* Locus—Because a conventional ChIP assay reflects the average density of RNAPIIs in population of cells, this method is not suitable to distinguish whether the higher signal of RNAPII in the beginning of the gene is detected due to higher density of polymerases on the 5' region of the gene in all cells, or alternatively, more polymerase is detected due to the initiation of *GAL-VPS13* transcription in most of the cells, whereas only in a fraction of them polymerases reach the end of the gene.

To be able to detect multiple transcribing RNAPII complexes in a single *GAL-VPS13* locus, we made a new strain, where the Rpb3 subunit of RNAPII was tagged with E2 epitope in its genomic locus, and in addition, two extra copies of *RPB3* genes, carrying either E4 or myc tags were inserted into *HIS3* and *LEU2* loci, respectively. Because the RNAPII complex contains a single subunit of Rpb3 protein, every RNAPII in this strain was tagged with E2, E4, or myc tag, but not with multiple tags simultaneously. Expression of all three Rpb3 proteins was confirmed by Western blotting (data not shown), and more importantly, all three Rpb3 versions were equally incorporated into functional RNAPII complexes because all of them were efficiently recruited to *GAL-VPS13* locus upon galactose induction (Fig. 3).

We used this strain for re-ChIP assay to detect multiple RNAPII complexes on the same DNA fragment and to determine the distribution of polymerases in *GAL-VPS13* locus. Two rounds of ChIPs were performed after the induction of *GAL-VPS13* transcription. For the first round we used an antibody recognizing the E2 tag to pull down RNAPII complexes carrying this epitope (*i.e.* one third of all polymerases in cell extract). If only a single RNAPII was bound to every DNA fragment, no co-precipitation of myc or E4 tags would be expected together with the E2 tag, and when this eluate is carried to the next round of immunoprecipitation with different antibody, no recovery of RNAPII-DNA complexes should be observed. However, if

Distribution of Elongating RNAPII

more than one RNAPII was present in the same protein-DNA complex, we would expect that polymerases carrying different Rpb3 tags were also co-precipitated. In this case, we should be able to recover DNA in the second round of immunoprecipitation with antibodies recognizing different Rpb3 tags. Importantly, successful recovery of DNA in the re-ChIP assay indicates that the whole RNAPII-DNA complex originated from the same cell because multiple polymerases were bound to the same DNA fragment. Therefore, the re-ChIP assay provides a unique opportunity to detect the presence of multiple transcribing polymerases on single chromatin fragment *in vivo*.

Equal amounts of the anti-E2 immunoprecipitation eluate were used as the starting material for the second round of ChIP, and the recovery of RNAPII with different antibodies was compared throughout the *GAL-VPS13* locus. In addition to anti-myc and anti-E4 antibodies, we also used the 4H8 antibody recognizing the Rpb1 subunit of RNAPII to monitor the general efficiency of RNAPII immunoprecipitation in the second round of re-ChIP. Because the 4H8 antibody recognizes all RNAPII complexes regardless of their different epitope tags on Rpb3 subunits, we defined the signal from this precipitation as the maximal possible recovery of RNAPII complexes in the second round of re-ChIP. As expected, precipitation of total RNAPII with 4H8 was always more efficient than with antibodies recognizing different tags in Rpb3. Re-precipitation of the eluate with anti-E4 or anti-myc antibodies recovered 30–40% of RNAPII complexes in the *GAL-VPS13* locus compared with precipitations with 4H8 antibody. Interestingly, we did not observe major differences in the efficiency of RNAPII re-ChIP in different regions of the *GAL-VPS13* gene. The recovery of differently tagged polymerases was rather uniform throughout the *GAL-VPS13* locus regardless of the distance from the promoter (Fig. 4A).

Taking into account that three differently tagged polymerases are present in our experimental system, it is unlikely to recover 100% of RNAPII-DNA complexes in the second round of re-ChIP with anti-myc or anti-E4 antibody. Two major factors influence the efficiency of re-ChIP in the second round: (i) the probability of finding two differently tagged RNAPII complexes on the same DNA fragment and (ii) the efficiency of cross-linking of RNAPII to DNA. If only one RNAPII is bound to every DNA fragment, the probability (p_1) of catching another epitope tag in the E2 precipitations equals 0. If two polymerase complexes are bound to the same DNA fragment, the probability (p_2) of catching another epitope tag is $0.33 \times X_{CL-eff}$ (one third of all possibilities multiplied by the RNAPII-DNA cross-linking efficiency coefficient X_{CL-eff}). If more polymerases are bound, the probability of catching the second epitope tag in the same DNA-RNAPII complex increases, and it can be calculated by the equation $p_{n+1} = p_n + (1 - p_n) \times 0.33 \times X_{CL-eff}$ where p_n is the probability of the presence of two differently tagged polymerases on DNA, if the total number of polymerases bound to this DNA is n . Accordingly, p_{n+1} is the probability of the presence of two differently tagged polymerases on DNA, if the total number of polymerases bound to this DNA is $n+1$.

We implemented this formula to estimate the number of transcribing polymerases in a single DNA-protein complex. Experimentally obtained values from re-ChIP assays (Fig. 4A)

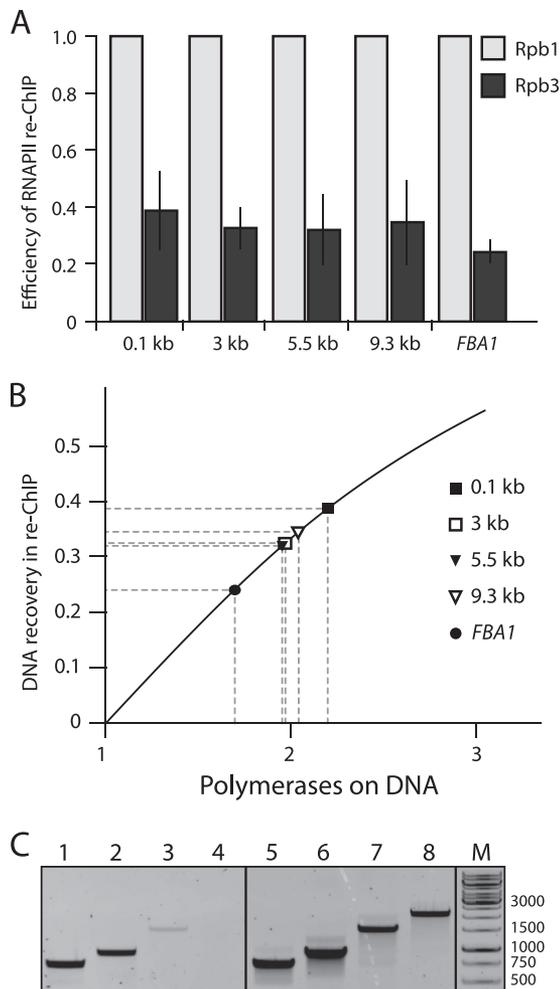


FIGURE 4. Analysis of the number of RNAPII complexes on a single DNA fragment. A, recovery of differently tagged RNAPII complexes in the second round of re-ChIP assay analyzed in *GAL-VPS13* and *FBA1* loci. The eluate from the first round immunoprecipitation of E2-tagged RNAPII complexes was re-precipitated with antibodies against E4 or myc tags to recover RNAPII-DNA complexes with multiple polymerases (columns *Rpb3*). Re-precipitation of total RNAPII complexes with 4H8 antibody (recognizing the Rpb1 subunit) in the second round of re-ChIP was set as 100% for all analyzed regions (columns *Rpb1*). The graph represents the average of three assays, error bars indicate S.D. B, estimation of the amount of RNAPII complexes on a single DNA fragment. The curve represents the theoretical probability of finding two differently tagged polymerases on the same DNA fragment as a function of the number of total polymerases bound to that DNA. The curve was calculated according to the equation $p_{n+1} = p_n + (1 - p_n) \times 0.33 \times X_{CL-eff}$ for conditions if $X_{CL-eff} = 1$ (see "Results" for details). Experimentally obtained values from A were fitted on top of the curve to estimate the number of polymerases in DNA-RNAPII complexes in *GAL-VPS13* and *FBA1* loci. C, analysis of the length of genomic DNA fragments in the WCE. DNA from the WCE used in re-ChIP experiments was extracted, amplified by PCR with different *VPS13* primers, and analyzed on agarose gel electrophoresis. Lanes 1–4, genomic DNA fragments from strain AKY273 were amplified by PCR producing 700-, 1000-, 1500-, and 2000-bp fragments. Lanes 5–8, unfragmented genomic DNA from *S. cerevisiae* was amplified with the same set of primers as on lanes 1–4. Lane M, DNA size markers.

were fitted to the probability curve of two-tagged multiple RNAPII complex detection (Fig. 4B). For simplicity, the curve was calculated with an assumption that the coefficient of cross-linking efficiency (X_{CL-eff}) equals 1, *i.e.* all RNAPII complexes were cross-linked to DNA during cell extract preparation. This assumption does likely overestimate the real cross-linking efficiency; however, it sets the border to define the minimal number of RNAPII complexes bound to the same DNA fragment. Our results show

that on average there are at least 2–2.5 polymerase complexes on the same DNA fragment in the *GAL-VPS13* locus (Fig. 4B). For comparison, we analyzed the density of RNAPIIs also on the *FBA1* gene, which has one of the highest transcription rates among constitutively expressed genes. The recovery of multiple RNAPII complexes in *FBA1* locus was less efficient, staying below 25% of recovery compared with 4H8 precipitations (Fig. 4A), corresponding to at least 1.8 polymerases on a single DNA fragment (Fig. 4B). We would like to emphasize that these calculations are based on the theoretical maximal efficiency of RNAPII cross-linking to the target DNA. Very likely, the real cross-linking efficiency was considerably lower, and therefore, we did not recover all multiple RNAPII complexes in the re-ChIP assays. This in turn means that those estimations should be considered as the minimal number of RNAPII complexes on a single DNA fragment.

Although the average size of DNA fragments in the WCE was about 500 bp as estimated by agarose gel-electrophoresis, we determined also the maximal length of DNA fragments in WCE by PCR. Genomic DNA was prepared from the WCE and was analyzed with a set of different *VPS13* primers to scan the maximal length of DNA. We estimated that the maximal length of DNA fragments was ~1000 bp, although the majority of them were shorter. Only a very weak signal was detected when DNA was amplified with primers generating 1500 bp of product, and no signal was detected for 2000 bp of product (Fig. 4C), ensuring that DNA fragments in WCE were short enough for reliable distinction of RNAPII-DNA complexes bound to different regions of the *GAL-VPS13* gene.

DISCUSSION

We used a re-ChIP assay to estimate the distribution of individual transcribing RNAPII complexes on a highly expressed gene in yeast. To distinguish between different RNAPII complexes, we made a yeast strain expressing three differently tagged forms of Rpb3 subunits of the RNAPII. Because only one Rpb3 protein is incorporated into 12-subunit RNAPII complex, every RNAPII in this strain was tagged with one of the three tags, but not with multiple tags simultaneously. We immunoprecipitated RNAPII from the cell extract with an antibody recognizing one of the epitope tags on Rpb3 and then re-precipitated the eluate with antibody against another tag to detect multiple RNAPII complexes in a single DNA-protein complex. As a prerequisite for successful re-ChIP assay, proteins must be bound to the common chromatin fragment, *i.e.* the whole protein-DNA complex originates from the same cell. Therefore re-ChIP provides an opportunity to detect the presence of multiple transcribing polymerases from individual cells.

When utilizing conventional ChIP assay to study the distribution of elongating RNAPII complexes in the highly transcribed *GAL-VPS13* locus, we noticed that upon induction the level of RNAPII recruitment to the 5' region of the gene was about three times higher than to the 3' end. Although this pattern of polymerase distribution can be expected at early stages of gene activation when transcription is being initiated simultaneously in the majority of cells, it remained the same for a long period and was clearly seen even in cells grown under inducing conditions overnight when the expression of galac-

tose-inducible genes had already been stabilized after initial burst of transcription (Fig. 2). This suggested that transcription of the gene was either prematurely terminated in a proportion of cells, or the velocity and spacing of elongating polymerases increased toward the end of the gene, or that both mechanisms contributed to the phenomenon.

Recent real-time measurements of mRNA synthesis and *in situ* hybridization assays have revealed remarkable variability of gene expression in individual yeast cells (5, 13). Although these methods are powerful tools to assess the absolute number of actively transcribing polymerases in the entire gene locus on single cell level, the distribution of individual RNAPII complexes across the locus remains obscure. We used an re-ChIP assay to determine the amount of transcribing polymerases in different regions of the *GAL-VPS13* gene and found that it remained fairly constant throughout the locus (Fig. 4). This suggests that deceptively uneven distribution of RNAPII in the *GAL-VPS13* locus in conventional ChIP assays (Figs. 1 and 2A) more likely reflects the heterogeneity of the cell population than the different spacing of polymerases in individual cells. This finding indicates that the results from conventional ChIP experiments should be interpreted with caution because they reflect the average signal from a population of cells. For example, the higher recovery of RNAPII in the 5' region of the gene has been generally interpreted as promoter-proximal pausing of polymerases. However, when analyzed in single cell level, the actual distribution of RNAPII complexes might be more uniform and therefore, many of the "paused" polymerases might actually represent prematurely terminating RNAPII complexes that never resume transcription.

We also applied re-ChIP assay to estimate the presence of multiple RNAPII complexes in some of the less intensively transcribed loci. In addition to *GAL-VPS13*, we also analyzed the coding regions of constitutively expressed *FBA1*, *TAN1*, and *ARN1* genes. From these, *FBA1* is a highly transcribed locus, whereas expression levels of *TAN1* and *ARN1* are relatively low (1–3). Recovery of multiple polymerases from the *FBA1* locus was slightly lower than from *GAL-VPS13* (Fig. 4), but we were unable to detect DNA from *ARN1* or *TAN1* loci in the re-ChIP assay (data not shown). This indicates that either the sensitivity of the re-ChIP assay is too low for detection of RNAPII complexes on lowly expressed genes, or that these loci are transcribed by single polymerases. In the latter case, the presence of multiple polymerases is not expected, and no DNA from these loci can be recovered in the re-ChIP assay. This possibility is also supported by several genome-wide studies showing that transcription of most of the genes is a rather rare event occurring only a few times during the cell cycle (1–3). Also, the analysis of gene transcription dynamics in living yeast cells has revealed that most constitutively expressed genes were transcribed by single polymerases (5).

If polymerases elongate at constant speed and abortion of transcription occurs randomly among elongating polymerases, the lower density of polymerases should be observed in the end of the gene, and larger polymerase-free regions should be found in the distal part of the transcription unit. However, our data show that the spacing between elongating polymerases remains stable in the entire transcribed locus (Fig. 4). This suggests that

either abortion of transcription occurs in a rather coordinated way affecting all polymerases on the gene, or the leading polymerase elongates slower than the following RNAPII complexes on the gene. In the latter case the faster polymerases can catch up the leading RNAPII and fill the caps left from spontaneously aborted polymerases. Both scenarios are likely to occur. For example, possible DNA damage in the gene locus might lead to the coordinated abortion of all elongating polymerases because the displacement and degradation of RNAPII complexes is required for efficient DNA repair in the coding regions of genes (14, 15). Also, the elongation speed of polymerases in the same locus might be different because the leading polymerase has to encounter generally more obstacles on its path than the following RNAPII complexes. For instance, opening the chromatin structure and temporal displacement of histones in the coding regions (6, 16–18) might slow down the leading polymerase and provide an opportunity for the following polymerases to catch up the leading RNAPII before the re-assembly of chromatin. Recent studies have provided evidence that pausing and backtracking of the elongating RNAPII occur very frequently *in vivo* (19, 20). Interestingly, in transcription factor IIS-deficient cells the global pattern of RNAPII pause sites correlate with nucleosome positions, indicating that nucleosomes represent the primary cause of transcription pausing *in vivo* (19), supporting the similar findings from RNAPII elongation studies on chromatin template *in vitro* (21–23).

Alternatively, the leading RNAPII might be more prone to spontaneous abortions than the following polymerases. It has been shown that many highly regulated genes are expressed as transcriptional bursts, where rapid initiations of multiple transcripts is followed by the relatively long periods of promoter inactivity (5, 24–26). In that case, the batch of elongating polymerases moves together through the entire gene locus. If the leading RNAPII aborts, the second polymerase in the row can take over the lead, and the spacing of the following polymerases stays constant. Although in this scenario the absolute number of transcribing polymerases decreases toward the end of the gene, the full-length transcript is made if the batch of RNAPII complexes is large enough and at least some polymerases can reach to the end of the gene. This might be the most efficient strategy for expression of inducible genes in response to external signals as multiple polymerases can more easily relieve the nucleosomal barrier on template and therefore can secure successful transcription of the gene (27, 28).

Taken together, our analysis of the distribution of RNAPII complexes on a highly expressed gene provides evidence that the density of elongating polymerases stays uniform throughout the transcribed locus. Our data also indicate that the results of assays averaging the signal from the cell population should be interpreted with care because the average results might mask the actual situation in individual cells.

Acknowledgments—We thank Drs. Tõnis Örd, Maida Remm, and Tanel Mahlakõiv for critical reading of the manuscript.

REFERENCES

1. Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S., and Young, R. A. (1998) *Cell* **95**, 717–728
2. Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M. (2008) *Science* **320**, 1344–1349
3. Pelechano, V., Chávez, S., and Pérez-Ortín, J. E. (2010) *PLoS One* **5**, e15442
4. Iyer, V., and Struhl, K. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5208–5212
5. Zenklusen, D., Larson, D. R., and Singer, R. H. (2008) *Nat. Struct. Mol. Biol.* **15**, 1263–1271
6. Kristjuhan, A., and Svejstrup, J. Q. (2004) *EMBO J.* **23**, 4243–4252
7. Glover-Cutter, K., Kim, S., Espinosa, J., and Bentley, D. L. (2008) *Nat. Struct. Mol. Biol.* **15**, 71–78
8. Kim, T. S., Liu, C. L., Yassour, M., Holik, J., Friedman, N., Buratowski, S., and Rando, O. J. (2010) *Genome Biol.* **11**, R75
9. Rodríguez-Gil, A., García-Martínez, J., Pelechano, V., Muñoz-Centeno Mde, L., Geli, V., Pérez-Ortín, J. E., and Chávez, S. (2010) *Nucleic Acids Res.* **38**, 4651–4664
10. Steinmetz, E. J., Warren, C. L., Kuehner, J. N., Panbehi, B., Ansari, A. Z., and Brow, D. A. (2006) *Mol. Cell* **24**, 735–746
11. Kristjuhan, A., Walker, J., Suka, N., Grunstein, M., Roberts, D., Cairns, B. R., and Svejstrup, J. Q. (2002) *Mol. Cell* **10**, 925–933
12. Métivier, R., Penot, G., Hübner, M. R., Reid, G., Brand, H., Kos, M., and Gannon, F. (2003) *Cell* **115**, 751–763
13. Larson, D. R., Zenklusen, D., Wu, B., Chao, J. A., and Singer, R. H. (2011) *Science* **332**, 475–478
14. Somesh, B. P., Reid, J., Liu, W. F., Søgaard, T. M., Erdjument-Bromage, H., Tempst, P., and Svejstrup, J. Q. (2005) *Cell* **121**, 913–923
15. Woudstra, E. C., Gilbert, C., Fellows, J., Jansen, L., Brouwer, J., Erdjument-Bromage, H., Tempst, P., and Svejstrup, J. Q. (2002) *Nature* **415**, 929–933
16. Lee, C. K., Shibata, Y., Rao, B., Strahl, B. D., and Lieb, J. D. (2004) *Nat. Genet.* **36**, 900–905
17. Schwabish, M. A., and Struhl, K. (2004) *Mol. Cell Biol.* **24**, 10111–10117
18. Värvi, C., Kristjuhan, K., and Kristjuhan, A. (2007) *Biochem. Biophys. Res. Commun.* **358**, 666–671
19. Churchman, L. S., and Weissman, J. S. (2011) *Nature* **469**, 368–373
20. Sigurdsson, S., Dirac-Svejstrup, A. B., and Svejstrup, J. Q. (2010) *Mol. Cell* **38**, 202–210
21. Bondarenko, V. A., Steele, L. M., Ujvári, A., Gaykalova, D. A., Kulaeva, O. I., Polikanov, Y. S., Luse, D. S., and Studitsky, V. M. (2006) *Mol. Cell* **24**, 469–479
22. Hodges, C., Bintu, L., Lubkowska, L., Kashlev, M., and Bustamante, C. (2009) *Science* **325**, 626–628
23. Kireeva, M. L., Hancock, B., Cremona, G. H., Walter, W., Studitsky, V. M., and Kashlev, M. (2005) *Mol. Cell* **18**, 97–108
24. Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y., and Tyagi, S. (2006) *PLoS Biol.* **4**, e309
25. Chubb, J. R., Trcek, T., Shenoy, S. M., and Singer, R. H. (2006) *Curr. Biol.* **16**, 1018–1025
26. Paré, A., Lemons, D., Kosman, D., Beaver, W., Freund, Y., and McGinnis, W. (2009) *Curr. Biol.* **19**, 2037–2042
27. Jin, J., Bai, L., Johnson, D. S., Fulbright, R. M., Kireeva, M. L., Kashlev, M., and Wang, M. D. (2010) *Nat. Struct. Mol. Biol.* **17**, 745–752
28. Kulaeva, O. I., Hsieh, F. K., and Studitsky, V. M. (2010) *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11325–11330