Tellier *et al. BMC Res Notes (2019) 12:554*
https://doi.org/10.1186/s13104-019-4582-8

BMC Research Notes

DATA NOTE

Open Access

Effect of CFIm68 knockdown on RNA polymerase II transcription

Michael Tellier*, Jessica G. Hardy, Chris J. Norbury and Shona Murphy

Abstract

Objectives: Transcription of eukaryotic protein-coding genes by RNA polymerase II (pol II) is highly regulated at initiation, elongation and termination. Transcription is also coordinated with co-transcriptional processing of the emerging pre-mRNA by capping, splicing, and cleavage and polyadenylation. Polyadenylation (poly(A)) site recognition, which defines the end of the mRNA, relies on the cleavage and polyadenylation (CPA) complex. It was previously observed that knocking-down proteins of the CPA complex affects not only recognition of the poly(A) site but also results in increased pausing of pol II at the beginning of genes. This finding suggests that the CPA complex plays a role in regulating pol II turnover after transcription initiation.

Data description: To explore this possibility, we knocked-down a subunit of the cleavage factor I (CFIm), CFIm68, which is part of the CPA complex and involved in alternative polyadenylation, and performed pol II ChIP-seq in absence or presence of a transcription elongation inhibitor. In addition, we performed pol II ChIP-qPCR on a subset of protein coding genes after knocking down CFIm68.

Keywords: RNA polymerase II, CFIm68, Transcription, Pol II pausing, Cleavage and polyadenylation factors

Objective

Transcription of a eukaryotic protein-coding gene by pol II requires several steps, including transcription initiation, elongation, and termination. During transcription, co-transcriptional processes such as mRNA capping, splicing, and cleavage and polyadenylation also occur and are required for the production of a mature mRNA. The end of a protein-coding gene is defined by one or more poly(A) sites and recognition of a poly(A) site is essential for the cleavage and polyadenylation of the mRNA [1, 2]. Approximately 85 proteins make up the cleavage and polyadenylation (CPA) complex and are distributed between four multi-subunits complexes that regulate poly(A) site recognition, pre-mRNA cleavage, and polyadenylation [2, 3]. The four complexes are cleavage and polyadenylation specificity factor (CPSF), cleavage stimulation factor (CstF), and cleavage factors I (CFIm) and II (CFIIm) [3]. CFIm is composed of two CFIm25 subunits, which binds the pre-mRNA, and two larger subunits, CFIm59 and CFIm68 [4, 5]. CFIm binds the pre-mRNA 40–50 nt upstream of the poly(A) site but its role in pre-mRNA cleavage remains unclear [6]. However, previous studies have shown a shift towards proximal poly(A) site usage following depletion of CFIm25 or CFIm68 [7–9], suggesting a role of CFIm in promoting distal poly(A) site recognition and longer mRNA 3′UTRs [10].

Some proteins of the CPA complex, including CstF64, CPSF73, and the CPA-associated termination factor Xrn2 have been shown to regulate pol II activity at the beginning and end of the transcription cycle [11, 12]. To determine whether depletion of CFIm also affects pol II pausing and transcription, we used a CRISPR/Cas9 approach to reduce the expression of two subunits of CFIm, CFIm25 and CFIm68 [8, 13], and performed pol II ChIP-seq in the CFIm68KD cell line in absence or presence of an inhibitor of cyclin-dependent kinase (CDK)9, whose activity regulates pol II pause release and entry into productive elongation [14].

*Correspondence: michael.tellier@path.ox.ac.uk
Sir William Dunn School of Pathology, University of Oxford, South Park Roads, Oxford OX1 3RE, UK

Data description

Cell culture

HEK293 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Sigma) supplemented with 10% fetal bovine serum (FBS, Gibco) and 100 units/ml penicillin + 100 μg/ml streptomycin (Gibco). The CFIm68KD cell lines, and its respective control HEK293 Flp, were previously described [8]. The cell lines were treated prior to ChIP-seq with DMSO or 100 μM 5,6-dichlorobenzimidazone-1-β-D-ribofuranoside (DRB, Sigma) for 30 min (Table 1).

ChIP-qPCR and ChIP-seq

ChIP was performed as previously described [15]. Briefly, cells were crosslinked at room temperature with 1% formaldehyde and quenched with 125 mM glycine for 5 min. Nuclear extracts were sonicated twice on a Bioruptor (Diagenode) for 15 min at high amplitude, 30 s ON/30 s OFF. 80 μg of chromatin was incubated overnight at 4 °C with 2 μg of an antibody against IgG (sc-2027, Santa Cruz) or against pol II (sc-899X, Santa Cruz). After recovery of immune complexes with BSA-saturated protein G Dynabeads and extensive washes, crosslinks were reversed by incubation at 65 °C for 5 h. After ethanol precipitation and proteinase K treatment, DNA was purified using a MinElute PCR Purification Kit (Qiagen). A single replicate of ChIP samples were sequenced on an Illumina HiSeq 4000 with 75 bp paired-end reads (Wellcome Trust Centre for Human Genetics, University of Oxford). For ChIP-qPCR, the list of primers can be found in Additional file 1. Pol II ChIP-qPCR were done in biological triplicates and can be found in Additional file 2. Statistical test: unpaired *t* test, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Bioinformatics analysis

Adapters were trimmed with Cutadapt v. 1.9.1 [16] with the following constant parameters: --minimum-length 10

Table 1 Overview of data files

Label	Name of data file/data set	File types (file extension)	Data repository and identifier (DOI or accession number)
Data file 1 [20, 21]	293 Flp-In Input	Fastq.gz (raw files), bigwig (processed files)	ENA accession number (fastq.gz): https://identifiers.org/ena.embl:SRX2915405 GEO accession number (bigwig): https://identifiers.org/geo:GSM2666454
Data file 2 [20, 21]	293 Flp-In Pol II	Fastq.gz (raw files), bigwig (processed files)	ENA accession number (fastq.gz): https://identifiers.org/ena.embl:SRX2915406 GEO accession number (bigwig): https://identifiers.org/geo:GSM2666455
Data file 3 [20, 21]	293 Flp-In DRB Input	Fastq.gz (raw files), bigwig (processed files)	ENA accession number (fastq.gz): https://identifiers.org/ena.embl:SRX6095228 GEO accession number (bigwig): https://identifiers.org/geo:GSM3898200
Data file 4 [20, 21]	293 Flp-In DRB Pol II	Fastq.gz (raw files), bigwig (processed files)	ENA accession number (fastq.gz): https://identifiers.org/ena.embl:SRX6095229 GEO accession number (bigwig): https://identifiers.org/geo:GSM3898201
Data file 5 [20, 21]	68KD Input	Fastq.gz (raw files), bigwig (processed files)	ENA accession number (fastq.gz): https://identifiers.org/ena.embl:SRX2915407 GEO accession number (bigwig): https://identifiers.org/geo:GSM2666456
Data file 6 [20, 21]	68KD Pol II	Fastq.gz (raw files), bigwig (processed files)	ENA accession number (fastq.gz): https://identifiers.org/ena.embl:SRX2915408 GEO accession number (bigwig): https://identifiers.org/geo:GSM2666457
Data file 7 [20, 21]	CFIm68KD DRB Input	Fastq.gz (raw files), bigwig (processed files)	ENA accession number (fastq.gz): https://identifiers.org/ena.embl:SRX6095230 GEO accession number (bigwig): https://identifiers.org/geo:GSM3898202
Data file 8 [20, 21]	CFIm68KD DRB Pol II	Fastq.gz (raw files), bigwig (processed files)	ENA accession number (fastq.gz): https://identifiers.org/ena.embl:SRX6095231 GEO accession number (bigwig): https://identifiers.org/geo:GSM3898203
Additional file 1 [22]	List of primers	.docx	https://doi.org/10.6084/m9.figshare.9159869
Additional file 2 [23]	Pol II ChIP-qPCR results	.pdf	https://doi.org/10.6084/m9.figshare.9159878

–q 15, 10 --max-n 1. Obtained sequences were mapped to the human hg19 reference sequence with Bowtie2 v. 2.2.5 [17]. Unmapped reads were removed with SAMtools v. 1.3.1 [18]. Mapped reads were then de-duplicated using Picard to remove PCR duplicates. Bam files were sorted and indexed with SAMtools. Bigwig files were created with a FPKM (Fragments per kilobase per million mapped reads) normalization by employing deepTools2 v. 2.2.4 [19] bamCoverage tool with the following parameters: -bs 10 -normalizeToRPKM -e –p max. Metaprofiles were created with deepTools2 computeMatrix tool.

Limitations

The knockdown of CFIm68 was not complete and may therefore may not be sufficient to completely abrogate the role of CFIm68 in pol II pausing and transcription regulation. The ChIP-seq were performed only once and in only one cell line; HEK293. We also performed pol II ChIP-qPCR on a limited number of protein-coding genes.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s13104-019-4582-8.

Additional file 1. List of primers, List of primers used for the ChIP-qPCR presented in additional file.

Additional file 2. Pol II ChIP-qPCR results, results of the pol II ChIP-qPCR on three genes in the HEK293 Flp-In and CFIm68KD cell lines.

Abbreviations

Pol II: RNA polymerase II; DRB: 5,6-dichlorobenzimidazone-1-β-D-ribofuranoside; ChIP: chromatin immunoprecipitation; DMEM: Dulbecco's Modified Eagle's Medium; FBS: fetal bovine serum; FPKM: fragments per kilobase per million mapped reads; 3′UTR: 3′ untranslated region; CDK9: cyclin-dependent kinase 9; CPA: cleavage and polyadenylation complex; CFIm: cleavage factor I.

Acknowledgements

We thank the Oxford Genomics Centre at the Wellcome Centre for Human Genetics (funded by Wellcome Trust Grant Reference 203141/Z/16/Z) for the generation of sequencing data.

Authors' contributions

MT, JGH, CJN and SM designed different aspects of the research. JGH produced the CFIm68KD cell line, MT and JGH carried out the ChIP-seq, MT performed the bioinformatics analysis and the ChIP-qPCR. MT, CJN and SM drafted the manuscript. All authors have read and approved the final manuscript.

Funding

This work was supported by a Wellcome Trust Senior Investigator Grant WT106134AIA to SM and a Cancer Research UK (CR-UK) Grant Number C38302/A13012, through an Oxford Cancer Research Centre Prize DPhil Studentship to JGH. The funders had no role in the design of the study and collection, analysis and interpretation of data and in writing the manuscript.

Availability of data materials

The data described in this Data note can be freely and openly accessed on the GEO website under the Accession Number: GSE99955 [20] and in ENA under the Accession Number PRJNA390279 [21]. Please see Table 1 and reference list for details and links to the data.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 20 June 2019 Accepted: 21 August 2019
Published online: 02 September 2019

References

1. Derti A, Garrett-Engele P, Macisaac KD, Stevens RC, Sriram S, Chen R, et al. A quantitative atlas of polyadenylation in five mammals. Genome Res. 2012;22(6):1173–83.
2. Shi Y, Manley JL. The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. Genes Dev. 2015;29(9):889–97.
3. Takagaki Y, Ryner LC, Manley JL. Four factors are required for 3′-end cleavage of pre-mRNAs. Genes Dev. 1989;3(11):1711–24.
4. Kim S, Yamamoto J, Chen Y, Aida M, Wada T, Handa H, et al. Evidence that cleavage factor Im is a heterotetrameric protein complex controlling alternative polyadenylation. Genes Cells. 2010;15(9):1003–13.
5. Yang Q, Gilmartin GM, Doublie S. Structural basis of UGUA recognition by the Nudix protein CFI(m)25 and implications for a regulatory role in mRNA 3′ processing. Proc Natl Acad Sci USA. 2010;107(22):10062–7.
6. Ruegsegger U, Beyer K, Keller W. Purification and characterization of human cleavage factor Im involved in the 3′ end processing of messenger RNA precursors. J Biol Chem. 1996;271(11):6107–13.
7. Kubo T, Wada T, Yamaguchi Y, Shimizu A, Handa H. Knock-down of 25 kDa subunit of cleavage factor Im in Hela cells alters alternative polyadenylation within 3′-UTRs. Nucleic Acids Res. 2006;34(21):6264–71.
8. Hardy JG, Tellier M, Murphy S, Norbury CJ. The RS domain of human CFIm68 plays a key role in selection between alternative sites of pre-mRNA cleavage and polyadenylation. 2017. https://doi.org/10.1101/177980.
9. Zhu Y, Wang X, Forouzmand E, Jeong J, Qiao F, Sowd GA, et al. Molecular mechanisms for CFIm-mediated regulation of mRNA alternative polyadenylation. Mol Cell. 2018;69(1):62–74.
10. Hardy JG, Norbury CJ. Cleavage factor Im (CFIm) as a regulator of alternative polyadenylation. Biochem Soc Trans. 2016;44(4):1051–7.
11. Brannan K, Kim H, Erickson B, Glover-Cutter K, Kim S, Fong N, et al. mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. Mol Cell. 2012;46(3):311–24.
12. Nojima T, Gomes T, Grosso ARF, Kimura H, Dye MJ, Dhir S, et al. Mammalian NET-Seq reveals genome-wide nascent transcription coupled to RNA processing. Cell. 2015;161(3):526–40.
13. Tellier M, Hardy JG, Norbury CJ, Murphy S. Effect of CFIm25 knockout on RNA polymerase II transcription. BMC Res Notes. 2018;11(1):894.
14. Laitem C, Zaborowska J, Isa NF, Kufs J, Dienstbier M, Murphy S. CDK9 inhibitors define elongation checkpoints at both ends of RNA polymerase II-transcribed genes. Nat Struct Mol Biol. 2015;22(5):396–403.
15. Laitem C, Zaborowska J, Tellier M, Yamaguchi Y, Cao Q, Egloff S, et al. CTCF regulates NELF, DSIF and P-TEFb recruitment during transcription. Transcription. 2015;6(5):79–90.
16. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011;17(1):3.
17. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–9.
18. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078–9.

19. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016;44(W1):W160–5.
20. Tellier M, Hardy JG, Norbury CJ, Murphy S. Effect of CFIm68 knockdown on RNA polymerase II transcription. GEO GSE99955. 2019. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE99955.
21. Tellier M, Hardy JG, Norbury CJ, Murphy S. Effect of CFIm68 knockdown on RNA polymerase II transcription. ENA PRJNA390279. 2019. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA390279.
22. Tellier M, Hardy JG, Norbury CJ, Murphy S. Effect of CFIm68 knockdown on RNA polymerase II transcription. Figshare. 2019. https://doi.org/10.6084/m9.figshare.9159869.
23. Tellier M, Hardy JG, Norbury CJ, Murphy S. Effect of CFIm68 knockdown on RNA polymerase II transcription. Figshare. 2019. https://doi.org/10.6084/m9.figshare.9159878.

Publisher's Note