bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Programmed Delayed Splicing: A Mechanism for Timed Inflammatory Gene Expression

Devdoot S. Majumdar*[1], Luke Frankiw[1], Christian H. Burns[1], Yvette Garcia-Flores[1], David Baltimore[1]

[1]Division of Biology, California Institute of Technology.

*Correspondence to baltimo@caltech.edu

SUMMARY:

Inflammation involves timed gene expression, suggesting that the fine-tuned onset, amplitude, and termination of expression of hundreds of genes is of critical importance to organismal homeostasis. Recent study of post-transcriptional regulation of inflammatory gene expression led to the suggestion of a regulatory role for pre-mRNA splicing. Here, using a hybrid capture approach to purify incompletely spliced, chromatin-associated pre-mRNAs, we use deep sequencing to study pre-mRNA splicing of the NF-kB transcriptome. By freezing transcription and examining subsequent splicing of complete transcripts, we find many introns splice tens to hundreds of times slower than average. In many cases, this is attributable to poor splice donor sequences that are evolutionarily conserved. When these introns were altered by ~2 base pairs to yield stronger splice donors, gene expression levels increased markedly for several genes in the context of a reporter system. We propose that such splice sites represent a regulatory mechanism that determines the timing of production of the mRNAs from certain inflammatory genes and may also limit mRNA expression from these genes. Further work will be needed to understand the roles of this regulation in the inflammatory response. The suggestion of extensive temporal regulation of pre-mRNA splicing as a regulatory process in inflammation raises the question of where else in biology there may be timed processes with a similar underlying cause.

KEYWORDS: INFLAMMATION, INNATE IMMUNITY, MACROPHAGE, GENE EXPRESSION, SPLICING

INTRODUCTION

Gene expression in response to an inflammatory stimulus begins rapidly and is tightly controlled by conventional means (transcription and protein turnover (Chen and Chen, 2013; Gautier et al., 2012; Smale et al., 2014)) and by an expanding list of modalities that have gained in appreciation as being general regulatory strategies (RNA stabilization, RNA deadenylation, ribosomal

bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

regulation, microRNA regulation, as examples (Hao and Baltimore, 2009; Leppek et al., 2013; O'Connell et al., 2012; Wan et al., 2007)). We and others have recently investigated the role of RNA splicing kinetics – independently of alternative splicing – in gene expression (Hao and Baltimore, 2013; Pandya-Jones et al., 2013; Rabani et al., 2011; Rabani et al., 2014). In macrophages, an inflammatory stimulus leads to upregulation of expression of pre-mRNAs from hundreds of genes, providing an experimentally favorable system to investigate whether differential kinetics of pre-mRNA splicing may control the timing of gene expression following an inducing stimulus.

Pre-mRNA conversion to mRNA has been implicated in regulation of gene expression in diverse systems. As part of the cellular response to various environmental stressors, mRNAs for ribosomal proteins were shown to be downregulated due to decreased splicing efficiency in yeast (Bergkessel et al., 2011). Global changes in efficiency of pre-mRNA splicing have been shown to be a developmental prerequisite for *Drosophila* early embryonic development (Guilgur et al., 2014). The developing vertebrate embryo obeys a ‘segmentation clock’ determining body segment length whose very timing relies on delays attributable to control of the splicing rate of the *Hes7* transcriptional repressor (Takashima et al., 2011).

In certain well-studied cases, as with the cytokine TNFα, regulatory mechanisms modulating RNA levels exert significant physiological effects (Eissa et al., 1996; Hargreaves et al., 2009; Kontoyiannis et al., 1999; Mahtani et al., 2001; Mino et al., 2015; Rao et al., 2005; Ruggiero et al., 2009; Stoecklin et al., 2003). The insight that TNFα contains AU-rich elements in its 3’ untranslated region that act as mRNA degradation signals (Han et al., 1990), and subsequent observations that a mouse in which these AU-rich elements were removed results shows a robust autoimmune phenotype (Kontoyiannis et al.), was an early indication of the importance of precisely tuned mRNA levels in the regulation of inflammation to avoid autoimmunity. Given the role of pre-mRNA splicing in biogenesis of mature mRNA, we and others (Bhatt et al., 2012; Cho et al., 2014; Davis-Turak et al., 2015; Grabherr et al., 2011; Hao and Baltimore, 2013) suggested that regulation of splicing kinetics may influence the gene expression kinetics that define the inflammatory cascade.

To examine the timing of intron removal from 230 different transcripts induced by TNFα in macrophages, we have developed a method for highly enriching transcript populations for mRNAs of interest, which is followed by deep sequencing of the largely pre-mRNA populations we purify. The induction of transcripts and removal of introns can be quantified with precision, with lifetimes of introns determined by blocking transcription early after induction. Among genes whose mRNAs appear more slowly after induction, we identify ones containing introns with poor binding sites for

bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

splicing factor U1, usually finding one per transcript. We call these “bottleneck” introns. Among the most rapidly induced genes we find no such introns. We show that in these pre-mRNAs, the sequence of the U1 binding site is critical for the speed of intron removal by building mini-genes with “repaired” introns and showing that these splice at the canonical rate (identified as about 20 seconds after polymerase has passed that point). We propose that bottleneck introns are important for determining either the rate of degradation of pre-mRNAs or the rate of appearance of mature mRNA or both.

RESULTS

It has been established that pre-mRNA is highly enriched in the chromatin-associated, polyadenylated RNA fraction (Tilgner et al.). To examine splicing events in pre-mRNAs, we performed time-course experiments with TNF-stimulated bone marrow-derived macrophages (BMMs), isolating RNA after biochemical separation of chromatin-associated material (Bhatt et al.). Using a hybrid-capture approach, we targeted sequencing toward transcripts of 230 genes previously identified as TNF-induced inflammatory mRNAs (Ramirez-Carrozzi et al., 2009). Purification of cDNA corresponding to these inflammatory transcripts involved reverse transcription of chromatin-associated RNA using oligo(dT), and capture of desired cDNAs using biotinylated probes complementary to the last exon of each gene of interest. Oligo(dT) priming and the choice of the last exon as a capture target enabled us to sequence complete transcripts from the standpoint of the splicing machinery, because all introns will have been transcribed in such transcripts. The hybrid capture strategy, based on a published approach (Engreitz et al., 2013), involved: (1) microarray printing of 12,000 150-bp ssDNAs designed from tiled fragments of the last exon of each gene of interest; (2) conversion to a pool of biotinylated ssRNA probes using PCR followed by in vitro transcription with biotinylated ribonucloetides; (3) hybridization of ssRNA pools to cDNA from each biological experiment, and (4) and streptavidin-coated bead-mediated capture of the transcripts of interest (Fig. S1). This approach resulted in a ~30-fold enrichment of genes of interest, with 70% of the sequenced reads corresponding to the genes of interest; RNA submitted only to poly(A) selection contained only 2% of such reads (Fig. S2A). In this way we could analyze 1024 introns from genes induced by an inflammatory stimulus (Fig S2B).

The chromatin-associated, selected transcripts from many time points following induction were sequenced and are represented as a read density histogram, wherein thousands of sequencing reads are histogrammed along the length of the gene, illustrating the relative abundance of sequence

bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

elements such as introns and exons within each gene's total population of reads. Such a representation of one transcript—that of *NFKBIA*, the gene encoding IκB-α— is shown as a time course evolving over 1 hour following TNF induction in Fig. 1A; the read density histograms are shown on a $\log_{10}$ scale, accentuating differences along the gene. The data from single time points is normalized to the highest frequency bin from that time point to allow comparison between time points and to emphasize the kinetics of mRNA biogenesis apparent from the dynamics of appearance/disappearance of sequence elements of IκB-α. New transcription is seen at 6 minutes for this RNA and most others—this is particularly obvious if a linear unnormalized scale of read density histograms is used (Fig. S3). The corresponding log-scaled histograms permit visualization of the intronic signal as a function of time after induction. Furthermore, from Fig. 1A it is evident that at all time points following induction, the 5'-proximal introns have been totally removed from the sequenced transcripts, indicating that the selection against partial transcripts is quite complete. Whereas excision of the first intron is always observed, the middle three introns are seen at intermediate states of excision in all time points such that intron definition for these introns is readily observable from read density histograms. We attribute this first exon excision largely to cotranscriptional splicing, consistent with other genome-wide splicing studies (Bentley, 2014). Strikingly, the final intron deviates significantly in its kinetic trajectory, as its read density does not obey a similar relative reduction. This might be due to a lag in terminal intron splicing (Carrillo Oesterreich et al., 2010) or a feature of splicing that accompanies transcript release from chromatin.

To better quantify the observed dynamics, we adapted the Coefficient of Splicing (CoSI) (Figure 1B), which quantifies the extent of splicing as a ratio of spliced to total (spliced and unspliced) junction reads such that CoSI values of ~1 and ~0 imply near-complete splicing and virtually unspliced states, respectively (Tilgner et al., 2012b). Though we observed a decrease in read density as a function of distance from the 3' of the gene (Fig. S3), presumably as a consequence of premature termination of the reverse transcriptase during copying of the pre-mRNA, the use of CoSI allows for an intron-specific splicing score regardless of read densities at neighboring introns. Using the CoSI metric, a time course plotting of the extent of splicing showed a dip in CoSI at ~6 minutes (Fig. 1C) corresponding to the aforementioned accumulation of new, unspliced transcripts. The splicing dynamics of each NFKBIA intron can be inferred from the CoSI dynamics, and the notable difference in splicing between the 5' and 3' introns is demonstrated by the amplitude of the 'dip' at 6 minutes and the time required for each intron to return to Co-SI of 1.

A surprising heterogeneity in CoSI was observed among all inflammatory introns (Fig 2A), implying diversity in their propensity to be spliced (Fig. 2). When considering all 1,024 introns in the

bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

chromatin-associated TNF time course we find most introns very near to CoSI ~1 relatively soon after induction, indicating that most introns do not remain unspliced long after induction begins at 6 minutes (Fig. 2A). Unexpectedly, although the median CoSI value remains high, we identified considerable heterogeneity among introns, some at and remaining near CoSI values <0.5, indicating relatively poor splicing, found very late into the time course.

As an example of a poorly spliced intron, chemokine CXCL10 intron 2 (Fig. 2A) is notable as it remains poorly spliced despite clear excision of neighboring introns, remaining quite unspliced even ~30 minutes after induction. It is possible that this intron undergoes splicing after its nascent chromatin-associated state, as is likely the case for the 3'-terminal intron of NFKBIA. It is also possible that this intron targets CXCL10 transcripts for degradation and the relatively fixed nature of intron 2's splicing status throughout the time course is a function of a constant rate of degradation. Introns with low CoSI at late time points post-TNF induction were considered putative 'bottleneck introns'– borrowing from the language that accompanied the discovery and characterization of slowly splicing U12-type introns (Patel et al., 2002). These introns were so slow to splice that they may intrinsically delay gene expression. Notably, the distribution of CoSI values of the entire dataset (Figure 2A and 2C) was very broad, and though most introns spliced immediately, many introns showed evidence of splicing bottlenecks, noticeable by their significant deviation from mean CoSI. At 10 and 60 minutes post-induction, 14% and 11% of introns, respectively, had CoSI values below one standard deviation of the mean (0.86±0.25 and 0.91±0.19, respectively). Shown as examples are CD40, DAXX, and IRF7 (Fig. 2B), genes whose immunological and inflammatory importance is well-established in studies with knockout mice (Honda et al.; Lei et al.; Michaelson et al.).

To quantify splicing kinetics, we used Actinomycin D (actD) to freeze transcription and followed the loss of intron and accumulation of splice junctions. In these experiments, splicing was analyzed at many time points immediately following actD treatment on the same 230 transcripts of interest and selected using hybrid capture from the total pool of cellular RNA rather than chromatin-associated RNA. Fitting the accumulation of spliced transcripts (as measured by CoSI) with an exponential distribution, we were able to extrapolate intron excision half-lives (Fig. 3). Because total cellular RNA was used, observed rates were independent of chromatin localization. We find intron half-lives that range from 20-40s (56% of introns splicing in this timing window) to several minutes, reflecting the considerable heterogeneity that is observed from CoSI differences.

A delay in splicing at certain sites could simply confer a delay in gene expression (by ~5 minutes in the slow case shown in Figure 4, IκBε), or, as is seen in yeast studies, it could result in both gene expression delay and gene expression diminution due to degradation of slowly splicing

bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

pre-mRNA (Koodathingal et al., 2010). Prior studies placed IκBε in a delayed splicing category (Hao and Baltimore), suggesting a pronounced splicing delay relative to rapidly induced genes. To understand a potential mechanistic basis for these differences in splicing time, each intron within our dataset was assessed computationally for the concurrence of its 5' splice donor sequence to a consensus sequence (Pessa et al., 2006). The 5' splice donor is a highly conserved sequence that directly base pairs with splicing factor U1 (Freund et al., 2005); deviation from consensus sequence confers a significantly reduced ability to engage the splicing machinery.

A maximum entropy model was used to calculate an intron quality score measuring extent of deviation from consensus splice sequence (Fig. S4). Among the inflammatory transcripts studied, many examples of introns with poor 5' donor scores were identified such as IRF7 and IL12b, where lower scores indicate significant deviation from consensus. We suggest that having non-consensus splice sites may be a regulatory mechanism affecting gene expression. We considered that splicing might show profound differences in the previously defined categories of induction (immediate/early/intermediate/late) characteristic of the inflammatory gene expression kinetics (Bhatt et al., 2012). We found that the 'immediate' genes showed consistently fast splicing (highest CoSI values) at all of their introns but the most 3" but that the other three groups shared similar CoSI distributions (Fig. S6). Using the bioinformatics "intron quality score" we also found that the introns of the later gene classes have significantly higher scoring 5' splice donor sequences (Fig. S4). Therefore, from experimental measurement of splicing and sequence-based prediction, genes expressed immediately following inflammatory stimulus are spliced faster, whereas all other inflammatory genes have a complex and heterogeneous distribution of splicing efficiency that does not stratify cleanly into the later kinetic categories (early/intermediate/late). Slowly splicing introns are found throughout these later kinetic categories in similar abundance, perhaps playing very gene-specific roles in diverse kinetic categories.

To test whether delays in splicing result in changes to gene expression, we identified a set of introns with the following criteria: (1) introns that splice poorly as defined by RNA-Seq, (2) introns that contain a low-scoring (non-consensus) 5' splice donor, and (3) introns whose weak 5' splice donor is evolutionarily conserved across many mammalian species. These introns were tested in the context of a splicing reporter expressed on a bidirectional promoter (Mukherji et al.). For each intron of interest, the reporter construct consists of a single transcript containing: (1) the 5' neighboring exon from the gene of interest, the intron of interest, and the 3' neighboring exon; (2) a 2A 'self-cleaving' peptide; and (3) the GFP gene. In the opposite orientation but from the same promoter, a blue fluorescent protein (BFP) mRNA is made in equal amounts to the intron-GFP construct ,GFP

bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

fluorescence of cells transfected with this reporter is a readout of splicing efficiency of the intron-GFP construct, when normalized to BFP fluorescence levels. Transfected into HEK293T cells and expressed for 24 hours, this bidirectional reporter enabled us to understand, at steady state, whether gene expression is affected by slow splicing introns. Measured using flow cytometry, the slope of the line corresponding to BFP:GFP ratios provides a relative metric of splicing efficiency, where slopes ~.9 and ~0.1 imply efficient and inefficient splicing, respectively.

To test the effect of a poor splice donor, we 'rescued' some poorly splicing introns in the context of the reporter by mutation to consensus splice donor 'GTAAG.' For instance, for IL12 intron 3, the splice donor sequence of "GTAAT" that is conserved among many mammalian species, was altered to "GTAAG" (Fig. S5). Expression levels of the reporter construct with the wild-type IL12 intron were found to be about half (57%) of the levels of the same construct with a single base pair alteration to make the stronger splice donor. *IRF7* intron 5 was tested against a 'fixed' intron as well as a wild-type intron from an actin gene, both resulting in two-fold improvements of gene expression. In one case, expression of *TFEC* (transcription factor EC) was not altered by splice site repair, suggesting that other mechanisms beyond 5'-splice site deficiencies may be involved in mediating slow splicing. Generally, when the BFP:GFP slopes of wild-type and mutated introns were compared by taking a ratio of their slopes, a change of ~2 nucleotides dramatically altered the slope of the line (Fig. 5a).

DISCUSSION

In this study we sought to understand splicing kinetics of the large number of genes that comprise the inflammatory response and to assess whether splicing itself might play a regulatory role in inflammation. We developed a targeted sequencing strategy, purifying transcripts containing each gene's terminal exon. This approach allowed us to sequence the 1,024 introns within inflammatory genes and permitted direct assessment of the structures of nearly-completed transcripts. We found considerable heterogeneity in splicing efficiency among these introns. In studying evolutionarily conserved weak 5' splice donors, we have isolated one cause of slow appearance of mRNA following a pulse of stimulus; many other slowly spliced introns without such sequences were also identified in this study and suggest other regulatory mechanisms may be responsible.

Crucially, the hybrid capture approach averts a common ambiguity in analyzing splicing kinetics of not being able to differentiate completed pre-mRNA from nascent transcripts during an induction pulse – this often leads to an overestimation of the unspliced status of early introns and

bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

complicates quantification of splicing kinetics. To the contrary, we rarely found genes containing unspliced first introns. This was true of chromatin-associated RNA and of whole cell RNA after actD stalling. These effects are consistent with the emerging model of co-transcriptional pre-mRNA splicing, where the splicing machinery has been suggested to lag 3-5kb behind the polymerase. Indeed, several recent global studies of RNA splicing bolster the claim that much pre-mRNA is spliced co-transcriptionally: 74% (yeast) or 75-84% (human) of introns are found to be at least 50% spliced by the time of transcription termination in several other studies (Ameur et al., 2011; Brugiolo et al., 2013; Carrillo Oesterreich et al., 2010; Girard et al., 2012; Khodor et al., 2011; Tilgner et al., 2012a). Surprisingly, this ~80% figure remains constant whether total RNA or chromatin-associated RNA is measured, implying that our choice to analyze chromatin-associated RNA does not significantly overrepresent splicing intermediates.

We found that most introns are spliced very efficiently, appearing and disappearing as a rapid dip of CoSI immediately following induction, returning to a CoSI of ~0.95 within minutes after induction. Notably, the distribution of CoSI values of the entire dataset (Figure 2) was very broad. Though most introns spliced immediately, there were several 'bottleneck introns'. In order to determine more specifically the rates of slowly splicing introns, studies employing actD to stall transcription and examine intron splicing half-lives corroborated the idea that there is tremendous intron-to-intron heterogeneity. Whereas most introns spliced within 20-40s, some were delayed significantly (upwards of 5 minutes). Of note, however, is that our 20-40s rate of splicing is somewhat at odds with other figures in the literature of 8-10 minutes for intron excision after a washout of the drug D-ribofuranosylbenzimidazole. There is some debate as to the perturbative role of actD in splicing, with one report observing that splicing intermediates in the context of the MS2 reporter system are prematurely liberated from chromatin upon actD treatment. Even in this case, the rapid actD-based rates, are likely underestimating even faster kinetics if one considers that the co-transcriptional splicing machinery targets chromatin mRNA faster than released mRNA (Martin et al.). However, even in the absence of actD, stimulation revealed that most of IKBα's introns are spliced in less than two minutes when one takes into account the 'dip' in CoSI due to induction and the time to reach steady CoSI levels (Figure 2). While the terminal intron of IKBα appears to have a longer half-life, this unique feature of terminal introns is consistent with prior studies (Carrillo Oesterreich et al., 2010).

In testing gene expression differences in bottleneck introns among introns with poor splice sites that are also evolutionarily conserved, we found that steady state levels of reporter proteins were upregulated when the 5' splice donor sequence was mutated to the consensus sequence

bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

'GTAAG' in all cases but one. Attenuated U1 binding provides a mechanistic insight for bottleneck introns that were chosen for their weak 5' splice donors. This implies that at the level of splicing, either due to delays in expression or perhaps degradation due to delayed expression, significant differences in gene expression arise from small differences in nucleotide sequence. We also find many slowly spliced introns not explained by weak 5' splice donors. In some cases, we find multiple bottlenecks introns per gene, as in the case of IRF7, where one bottleneck (intron 5) was attributable to an evolutionarily conserved weak 5' splice donor while another (intron 7) was observed experimentally but of unknown cause. This may be due to any of a number of potential mechanisms that may also serve in tuning the speed of splicing: cis-regulatory protein recruitment, 3' splice acceptor sequence or other sequence elements, or alterations of RNA polymerase speed or chromatin marks or three-dimensional gene structure.

Central to our inquiry is the enigmatic nature of these bottlenecks remaining in physiologically critical genes, often evolutionarily conserved, and yet intrinsically mediating an inefficiency in gene expression. We posit that the gene expression changes that are shown in bone marrow-derived macrophages offer a regulatory strategy to slow up and maybe restrict expression of genes in a manner dependent on the composition of mRNA processing factors in the cell ('the splicing landscape'), the cell type, or the stimulus type in question. Recent studies have demonstrated global changes in intron retention preferences in B cell lymphomagenesis and granulocyte differentiation (Koh et al., 2015). In a similar manner, we suggest that selection of splicing and kinetics of splicing might allow a previously unappreciated level of specificity to gene expression decisions in cells presented with an inflammatory stimulus. Juneau (Juneau et al., 2006), Brinster (Brinster et al., 1988), Shabalina (Shabalina et al., 2010), Paretneu (Parenteau et al., 2008), Furger (Furger et al., 2002), Kornbliht (Kornblihtt et al., 2004), Kroun {Damgaard, 2008 #241},

Induction with TNF is a particularly favorable situation because many of the genes we examined were up-regulated in their transcription within 4-6 minutes of adding inducer (Supplemental Fig. 2B), allowing examination of large numbers of pre-mRNA transcripts. This, in concert with the hybrid capture approach that provides a large number of junctional sequencing reads, has permitted unique insight into the kinetics of splicing of mature transcripts and revealed surprising heterogeneity. We suggest that this methodology and analysis could have wider applicability for other gene induction situations.

EXPERIMENTAL PROCEDURES:

bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Cells. C56BL6/J mice were sacrificed via CO_2 euthanasia and sterilized with 70% ethanol. Femur and tibia bones harvested and stripped of muscle tissue. Bone marrow cells were resuspended in 20mL of fresh DMEM. 2.5e6 bone-marrow cells plated in a 15-cm dish in 20mL of BMDM Media (DMEM, 20% FBS, 30% L929 condition media, and 1% Pen/Strep) and grown at 5% CO_2 and 37°C. BMDM media completely replaced on day 3 as well as a supplemental addition of 5mL L929 condition media on day 5.

RNA. Total RNA was prepared with Trizol (ThermoFisher) as per manufacturer instructions. Isolation of cytoplasmic, nucleoplasmic and chromatin-associated RNA fractions isolated used a modification of a previously published method (Bhatt et al., 2012). Full methods available in supplement.

cDNA Pulldown. cDNA is added biotinylated RNA probes, generated by Ampliscribe T7-Flash Biotin Kit (Epicentre), and incubated at 74°C for 4.5min to denature followed by addition of 1 volume of 2X hybridization (HYB) buffer (1M LiCl, 40mM Tris-HCl (pH 7.5), 20mM EDTA (pH 8.0), 4M Urea, 0.5% Triton X-100, 1% SDS, 0.2% Na-deoxycholate). Reaction incubated at 70°C for 30min. 0.3mg BioMag streptavidin beads (Bang Laboratories Inc.), washed 3 times in 1X HYB buffer, added and reaction incubated at 70°C and 1100rpm for 20min to capture cDNA-probe complex. Beads pelleted on magnet, followed by 2 washes of 150µL with preheated 1X HYB at 70°C, 1 wash of 150µL with wash #4 (160mM LiCl, 20mM Tris-HCl (pH 7.5), 10mM EDTA (pH 8.0), 2M Urea, 0.25% Triton X-100, 0.5% SDS, 0.1% Na-deoxycholate), and 1 final wash with wash #5 (40mM LiCl, 20mM Tris-HCl (pH 7.5), 10mM EDTA (pH 8.0), 2M Urea, 0.25% Triton X-100, 0.5% SDS, 0.1% Na-deoxycholate). Beads resuspended in 35µL of base elution buffer (125mM NaOH, 10mM EDTA (pH 8.0), 10mM Tris-HCl (pH 7.5) and incubated at 74°C and 1100rpm for 5min. Beads pelleted and 30µL cDNA containing supernatant removed to a new tube. Solution neutralized with 6.25µL neutralization buffer (800mM HCl, 160mM Tris-HCl (pH 7.5), 20mM EDTA (pH 8.0)). Immediately after neutralization, cDNA purified by 1.0X Sera-Mag treatment as described previously above and eluted in 45µL and stored at -80°C.

Sequence data was acquired using a HiSeq 2500 instrument (Illumina) at 50bp single-end resolution. We will deposit data in NCBI GEO. All code will be made freely available on GitHub.

ACKNOWLEDGEMENTS:

The authors would like to thank Alex Shishkin and Mitchell Guttman (Dept. of Biology, Caltech) for assistance with hybrid capture strategy design; and Ann-Jay Tong Stephen Smale, Doug Black and Amy-Pandya Jones (Dept. of Biology, University of California, Los Angeles) for insights and advice; and Sergei Manakov, Evelyn Stuwe, Dubravka Pezic, Igor Antoshechkin, Sagar Damle, and Alok Joglekar (Dept. of Biology, California Institute of Technology) for experimental and computational assistance. This work was funded from a grant from NIH and from an endowment provided by the Raymond and Beverly Sackler Foundation.

AUTHOR CONTRIBUTIONS:

DSM/LF/CB/DB designed the experiments and wrote the manuscript. DSM/LF/CB/YG designed and conducted experiments, and DSM/LF performed bioinformatics analysis on the data.

bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

REFERENCES:

Ameur, A., Zaghlool, A., Halvardson, J., Wetterbom, A., Gyllensten, U., Cavelier, L., and Feuk, L. (2011). Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. Nat Struct Mol Biol *18*, 1435-1440.
Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. Nature Reviews Genetics *15*, 163-175.
Bergkessel, M., Whitworth, G.B., and Guthrie, C. (2011). Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast. RNA *17*, 1461-1478.
Bhatt, D.M., Pandya-Jones, A., Tong, A.-J., Barozzi, I., Lissner, M.M., Natoli, G., Black, D.L., and Smale, S.T. (2012). Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. Cell *150*, 279-290.
Brinster, R.L., Allen, J.M., Behringer, R.R., Gelinas, R.E., and Palmiter, R.D. (1988). Introns increase transcriptional efficiency in transgenic mice. Proc Natl Acad Sci U S A *85*, 836-840.
Brugiolo, M., Herzel, L., and Neugebauer, K.M. (2013). Counting on co-transcriptional splicing. F1000Prime Rep *5*, 9.
Carrillo Oesterreich, F., Preibisch, S., and Neugebauer, K.M. (2010). Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. Mol Cell *40*, 571-581.
Chen, J., and Chen, Z.J. (2013). Regulation of NF-kappaB by ubiquitination. Curr Opin Immunol *25*, 4-12.
Cho, V., Mei, Y., Sanny, A., Chan, S., Enders, A., Bertram, E.M., Tan, A., Goodnow, C.C., and Andrews, T.D. (2014). The RNA-binding protein hnRNPLL induces a T cell alternative splicing program delineated by differential intron retention in polyadenylated RNA. Genome Biol *15*, R26.
Davis-Turak, J.C., Allison, K., Shokhirev, M.N., Ponomarenko, P., Tsimring, L.S., Glass, C.K., Johnson, T.L., and Hoffmann, A. (2015). Considering the kinetics of mRNA synthesis in the analysis of the genome and epigenome reveals determinants of co-transcriptional splicing. Nucleic acids research *43*, 699-707.
Eissa, N.T., Strauss, A.J., Haggerty, C.M., Choo, E.K., Chu, S.C., and Moss, J. (1996). Alternative splicing of human inducible nitric-oxide synthase mRNA tissue-specific regulation and induction by cytokines. Journal of biological chemistry *271*, 27184-27187.
Engreitz, J.M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., Kadri, S., Xing, J., Goren, A., and Lander, E.S. (2013). The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. Science *341*, 1237973.
Freund, M., Hicks, M.J., Konermann, C., Otte, M., Hertel, K.J., and Schaal, H. (2005). Extended base pair complementarity between U1 snRNA and the 5' splice site does not inhibit splicing in higher eukaryotes, but rather increases 5' splice site recognition. Nucleic Acids Res *33*, 5112-5119.
Furger, A., O'Sullivan, J.M., Binnie, A., Lee, B.A., and Proudfoot, N.J. (2002). Promoter proximal splice sites enhance transcription. Genes Dev *16*, 2792-2799.
Gautier, E.L., Shay, T., Miller, J., Greter, M., Jakubzick, C., Ivanov, S., Helft, J., Chow, A., Elpek, K.G., Gordonov, S.*, et al.* (2012). Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. Nat Immunol *13*, 1118-1128.
Girard, C., Will, C.L., Peng, J., Makarov, E.M., Kastner, B., Lemm, I., Urlaub, H., Hartmuth, K., and Luhrmann, R. (2012). Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. Nat Commun *3*, 994.
Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., and Zeng, Q. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature biotechnology *29*, 644-652.
Guilgur, L.G., Prudencio, P., Sobral, D., Liszekova, D., Rosa, A., and Martinho, R.G. (2014). Requirement for highly efficient pre-mRNA splicing during Drosophila early embryonic development. Elife *3*, e02181.
Han, J., Brown, T., and Beutler, B. (1990). Endotoxin-responsive sequences control cachectin/tumor necrosis factor biosynthesis at the translational level. The Journal of experimental medicine *171*, 465-475.
Hao, S., and Baltimore, D. (2009). The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules. Nat Immunol *10*, 281-288.
Hao, S., and Baltimore, D. (2013). RNA splicing regulates the temporal order of TNF-induced gene expression. Proc Natl Acad Sci U S A *110*, 11934-11939.

bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Hargreaves, D.C., Horng, T., and Medzhitov, R. (2009). Control of inducible gene expression by signal-dependent transcriptional elongation. Cell *138*, 129-145.
Honda, K., Yanai, H., Negishi, H., Asagiri, M., Sato, M., Mizutani, T., Shimada, N., Ohba, Y., Takaoka, A., Yoshida, N.*, et al.* (2005). IRF-7 is the master regulator of type-I interferon-dependent immune responses. Nature *434*, 772-777.
Juneau, K., Miranda, M., Hillenmeyer, M.E., Nislow, C., and Davis, R.W. (2006). Introns regulate RNA and protein abundance in yeast. Genetics *174*, 511-518.
Khodor, Y.L., Rodriguez, J., Abruzzi, K.C., Tang, C.H., Marr, M.T., 2nd, and Rosbash, M. (2011). Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in Drosophila. Genes Dev *25*, 2502-2512.
Koh, C.M., Bezzi, M., Low, D.H., Ang, W.X., Teo, S.X., Gay, F.P., Al-Haddawi, M., Tan, S.Y., Osato, M., and Sabò, A. (2015). MYC regulates the core pre-mRNA splicing machinery as an essential step in lymphomagenesis. Nature.
Kontoyiannis, D., Pasparakis, M., Pizarro, T.T., Cominelli, F., and Kollias, G. (1999). Impaired on/off regulation of TNF biosynthesis in mice lacking TNF AU-rich elements: implications for joint and gut-associated immunopathologies. Immunity *10*, 387-398.
Koodathingal, P., Novak, T., Piccirilli, J.A., and Staley, J.P. (2010). The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing. Molecular cell *39*, 385-395.
Kornblihtt, A.R., de la Mata, M., Fededa, J.P., Munoz, M.J., and Nogues, G. (2004). Multiple links between transcription and splicing. RNA *10*, 1489-1498.
Lei, X.F., Ohkawara, Y., Stampfli, M.R., Mastruzzo, C., Marr, R.A., Snider, D., Xing, Z., and Jordana, M. (1998). Disruption of antigen-induced inflammatory responses in CD40 ligand knockout mice. J Clin Invest *101*, 1342-1353.
Leppek, K., Schott, J., Reitter, S., Poetz, F., Hammond, M.C., and Stoecklin, G. (2013). Roquin promotes constitutive mRNA decay via a conserved class of stem-loop recognition motifs. Cell *153*, 869-881.
Mahtani, K.R., Brook, M., Dean, J.L., Sully, G., Saklatvala, J., and Clark, A.R. (2001). Mitogen-activated protein kinase p38 controls the expression and posttranslational modification of tristetraprolin, a regulator of tumor necrosis factor alpha mRNA stability. Molecular and Cellular Biology *21*, 6461-6469.
Martin, R.M., Rino, J., Carvalho, C., Kirchhausen, T., and Carmo-Fonseca, M. (2013). Live-cell visualization of pre-mRNA splicing with single-molecule sensitivity. Cell Rep *4*, 1144-1155.
Michaelson, J.S., Bader, D., Kuo, F., Kozak, C., and Leder, P. (1999). Loss of Daxx, a promiscuously interacting protein, results in extensive apoptosis in early mouse development. Genes Dev *13*, 1918-1923.
Mino, T., Murakawa, Y., Fukao, A., Vandenbon, A., Wessels, H.-H., Ori, D., Uehata, T., Tartey, S., Akira, S., and Suzuki, Y. (2015). Regnase-1 and Roquin Regulate a Common Element in Inflammatory mRNAs by Spatiotemporally Distinct Mechanisms. Cell *161*, 1058-1073.
Mukherji, S., Ebert, M.S., Zheng, G.X., Tsang, J.S., Sharp, P.A., and van Oudenaarden, A. (2011). MicroRNAs can generate thresholds in target gene expression. Nat Genet *43*, 854-859.
O'Connell, R.M., Rao, D.S., and Baltimore, D. (2012). microRNA regulation of inflammatory responses. Annu Rev Immunol *30*, 295-312.
Pandya-Jones, A., Bhatt, D.M., Lin, C.-H., Tong, A.-J., Smale, S.T., and Black, D.L. (2013). Splicing kinetics and transcript release from the chromatin compartment limit the rate of Lipid A-induced gene expression. RNA *19*, 811-827.
Parenteau, J., Durand, M., Veronneau, S., Lacombe, A.A., Morin, G., Guerin, V., Cecez, B., Gervais-Bird, J., Koh, C.S., Brunelle, D.*, et al.* (2008). Deletion of many yeast introns reveals a minority of genes that require splicing for function. Mol Biol Cell *19*, 1932-1941.
Patel, A.A., McCarthy, M., and Steitz, J.A. (2002). The splicing of U12-type introns can be a rate-limiting step in gene expression. EMBO J *21*, 3804-3815.
Pessa, H.K., Ruokolainen, A., and Frilander, M.J. (2006). The abundance of the spliceosomal snRNPs is not limiting the splicing of U12-type introns. RNA *12*, 1883-1892.
Rabani, M., Levin, J.Z., Fan, L., Adiconis, X., Raychowdhury, R., Garber, M., Gnirke, A., Nusbaum, C., Hacohen, N., Friedman, N.*, et al.* (2011). Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. Nat Biotechnol *29*, 436-442.
Rabani, M., Raychowdhury, R., Jovanovic, M., Rooney, M., Stumpo, D.J., Pauli, A., Hacohen, N., Schier, A.F., Blackshear, P.J., Friedman, N.*, et al.* (2014). High-resolution sequencing and modeling identifies distinct dynamic RNA regulatory strategies. Cell *159*, 1698-1710.

bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Ramirez-Carrozzi, V.R., Braas, D., Bhatt, D.M., Cheng, C.S., Hong, C., Doty, K.R., Black, J.C., Hoffmann, A., Carey, M., and Smale, S.T. (2009). A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. Cell *138*, 114-128.

Rao, N., Nguyen, S., Ngo, K., and Fung-Leung, W.-P. (2005). A novel splice variant of interleukin-1 receptor (IL-1R)-associated kinase 1 plays a negative regulatory role in Toll/IL-1R-induced inflammatory signaling. Molecular and cellular biology *25*, 6521-6532.

Ruggiero, T., Trabucchi, M., De Santa, F., Zupo, S., Harfe, B.D., McManus, M.T., Rosenfeld, M.G., Briata, P., and Gherzi, R. (2009). LPS induces KH-type splicing regulatory protein-dependent processing of microRNA-155 precursors in macrophages. The FASEB Journal *23*, 2898-2908.

Shabalina, S.A., Ogurtsov, A.Y., Spiridonov, A.N., Novichkov, P.S., Spiridonov, N.A., and Koonin, E.V. (2010). Distinct patterns of expression and evolution of intronless and intron-containing mammalian genes. Mol Biol Evol *27*, 1745-1749.

Smale, S.T., Tarakhovsky, A., and Natoli, G. (2014). Chromatin contributions to the regulation of innate immunity. Annu Rev Immunol *32*, 489-511.

Stoecklin, G., Lu, M., Rattenbacher, B., and Moroni, C. (2003). A constitutive decay element promotes tumor necrosis factor alpha mRNA degradation via an AU-rich element-independent pathway. Molecular and cellular biology *23*, 3506-3515.

Takashima, Y., Ohtsuka, T., Gonzalez, A., Miyachi, H., and Kageyama, R. (2011). Intronic delay is essential for oscillatory expression in the segmentation clock. Proc Natl Acad Sci U S A *108*, 3300-3305.

Tilgner, H., Knowles, D.G., Johnson, R., Davis, C.A., Chakrabortty, S., Djebali, S., Curado, J., Snyder, M., Gingeras, T.R., and Guigo, R. (2012a). Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. Genome Res *22*, 1616-1625.

Tilgner, H., Knowles, D.G., Johnson, R., Davis, C.A., Chakrabortty, S., Djebali, S., Curado, J., Snyder, M., Gingeras, T.R., and Guigó, R. (2012b). Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. Genome research *22*, 1616-1625.

Wan, F., Anderson, D.E., Barnitz, R.A., Snow, A., Bidere, N., Zheng, L., Hegde, V., Lam, L.T., Staudt, L.M., Levens, D.*, et al.* (2007). Ribosomal protein S3: a KH domain subunit in NF-kappaB complexes that mediates selective gene regulation. Cell *131*, 927-939.

bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

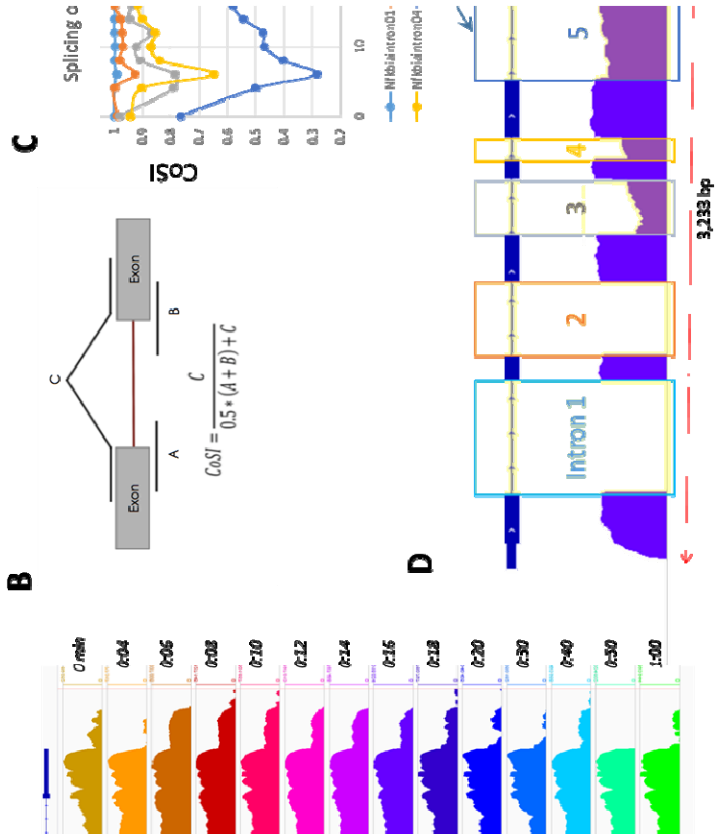
Fig. 1. Sequencing of complete, chromatin-associated pre-mRNA during inflammatory stimulus reveals differential splicing dynamics among introns of IKBα. (A) Histogram of reads corresponding to the TNF-induced expression and splicing of IKBα pre-mRNA of BMDMs. RNA-Seq was performed on chromatin associated RNA, enriched for NFkB genes as a function of a TNF stimulation timecourse, time shown in minutes after stimulation. Reads are histogrammed in $\log_{10}$ scale and normalized to each time point's maximum value. (B) The Coefficient of Splicing (CoSI) metric quantifies extent of splicing as a function of time, expressed as a ratio of reads from each splice junction to total junctional reads. Dynamics of IKBα splicing as a function of each intron's CoSI is shown (C), where 1=spliced and 0=unspliced, with corresponding introns highlighted in sample timepoint.

Fig. 2. Heterogeneity of splicing at each intron reveals splicing 'bottlenecks.' The Co-SI of each intron per timepoint is shown as a function of the entire inflammatory mRNA dataset as box-whisker plot (A). Each point represents an intron of one of 230 genes, revealing high rates of splicing (median Co-SI indicated by bar near 1.0 for each timepoint) for most genes with significant outliers. As an example, CXCL10 intron 2 (red arrowhead) is represented by the datapoint with arrowhead, and a histogram of reads is shown to demonstrate relative unspliced nature of this intron, which is not involved in alternative splicing. (B) Several similar introns that are relatively unspliced are found throughout the inflammatory transcriptome; shown are bottleneck introns within CD40, DAXX, and IRF7 as examples in the context of their neighboring introns.

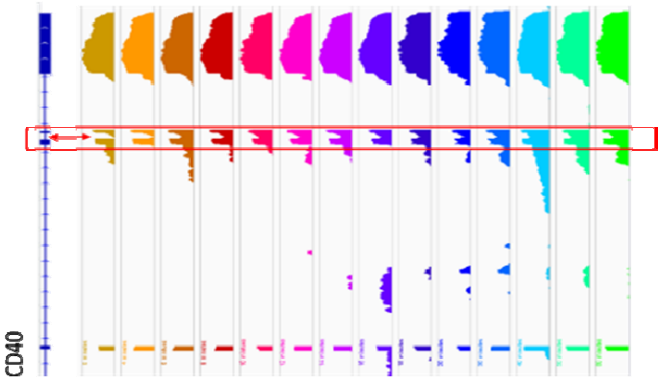
Fig. 3. Splicing kinetics of inflammatory introns are heterogeneous, ranging from seconds to minutes. CoSI of introns representing various splicing rates are measured and fit to half-lives. Cells were treated with Actinomycin D-treated, from which hybrid capture of genes of interest and sequencing was performed on total (unfractionated) RNA. Shown are four representative samples of splicing kinetics.

Fig. 4. Bottleneck introns can be repaired, and account for significant alterations to gene expression. (A) Intron-GFP splicing reporters for each wild-type intron (red) and modified intron (green) are shown as BFP:GFP ratio. (B) Ratio of WT:Fixed slopes is shown; whereas Tfec expression is not altered by improved 5' sequence, Malt1 intron sequence is significantly impaired owing to its 5' donor sequence, exhibiting a roughly 5-fold impairment in gene expression due to the 5' splice donor.

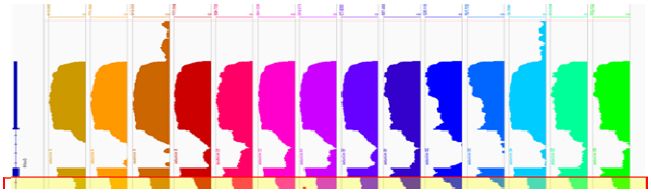
bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



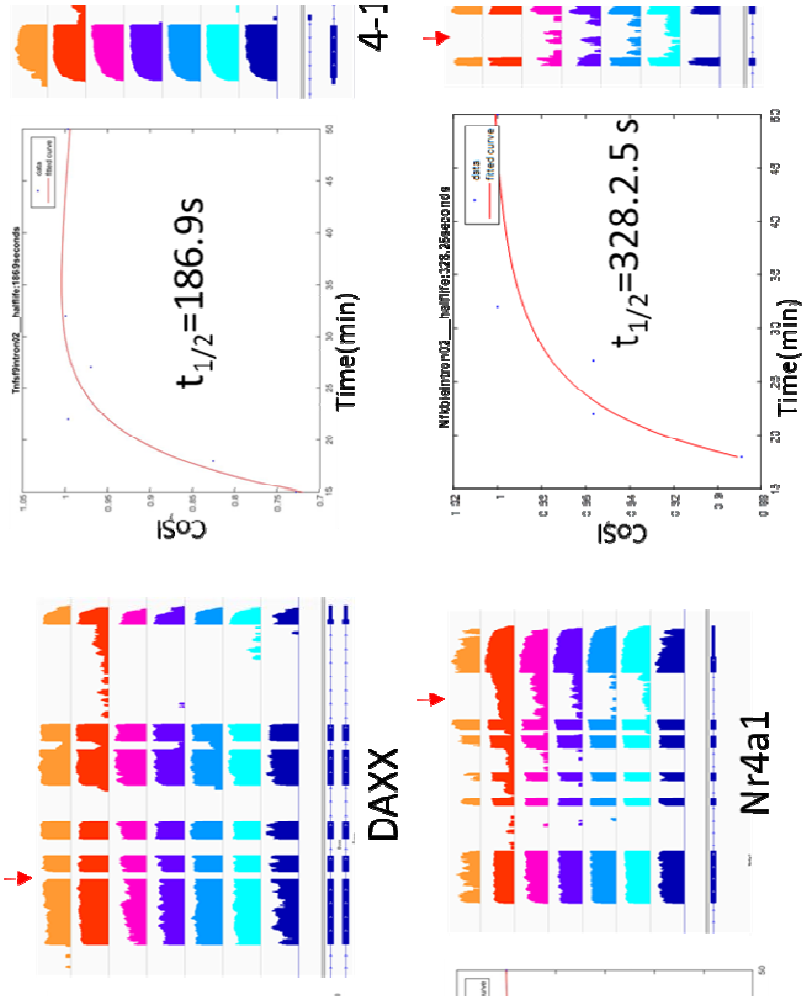
bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



B



bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/443796; this version posted October 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.