Figures and figure supplements

m^6A modification of U6 snRNA modulates usage of two major classes of pre-mRNA 5’ splice site

Matthew T Parker et al.
Figure 1. Loss of FIO1 causes early flowering and reduced splicing of pre-mRNA encoding the floral repressor MAF2. (A) Schematic showing the format of the two-step mutant screen. (B) Scatter plot showing allele fractions of ethyl methanesulfonate (EMS)-induced G to A transitions in pooled phenotypically normal plants (blue) and early flowering sisters (orange). Dark orange diamonds show SNPs predicted to have a significant impact on the functional expression of a protein-coding gene (nonsense mutations and splice site mutations). The green line shows the test statistic for the G-test.
between the allele fractions found in early flowering and normal plants, which has been smoothed using a tri-cube kernel with a window size of two megabases (Mb). The 2.6 Mb mapping interval containing FIO1 is highlighted in grey. (C) Gene track showing the three fio1 mutant alleles used in this study. fio1-1 is an EMS mutant with a G→A transition at the −1 position of the 3’ splice site (3’SS) in intron 2 of FIO1 (Kim et al., 2008). This causes activation of a cryptic 3’SS 15 nt downstream, and the loss of 5 aa of sequence from the FIO1 open reading frame (shown in orange). fio1-3 is a T-DNA insertion mutant (SALK_084201) in the first exon of the FIO1, disrupting the gene (region downstream of insertion shown in light blue). fio1-4 is an EMS mutant with a G→A transition at the +1 position of the 5’ splice site (5’SS) in intron 2 of FIO1. This causes activation of a cryptic 5’SS 69 nt upstream, and the loss of 23 aa of sequence from the FIO1 open reading frame (shown in orange). (D) Changes in splicing efficiency determined by RT-quantitative PCR (qPCR) visualised as a regression scatterplot showing the change in spliced to retained ratio of MAF2 intron 3 in fio1-1 and fio1-3 at a range of temperatures. Shaded regions show bootstrapped 95% CI for regression lines. (E) Boxplot showing the change in flowering time (days to flowering) observed in the fio1-1, fio1-3, and maf2 mutants at a range of temperatures. (F) Photographs showing the early flowering phenotypes of fio1-1, fio1-3, and fio1-4 mutants.
Figure 1—figure supplement 1. (A) RT-PCR products separated by agarose gel electrophoresis used to identify enhanced MAF2 intron 3 retention in the EMS-129 line, later renamed to fio1-4. (B) Photograph of the F2 segregating population used to map fio1-4.
Figure 2. FIO1-dependent m^6^A modification of poly(A)+mRNA is rare or absent, but FIO1 is required for U6 snRNA methylation. (A) Liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis. Boxplots showing that the ratio of m^6^A to A in poly(A)+mRNA is only modestly reduced in fio1-1 mutants compared to Col-0. In contrast, m^6^A levels are significantly reduced in the fip37-4 mutant. (B) The intersection of m^6^A modification sites detected by nanopore direct RNA sequencing of poly(A)+RNA purified from Col-0, fip37-4, and fio1-1, visualised using an upset plot. All sites shown have significant differences in modification level in a three-way comparison between fip37-4, fio1-1, and Col-0. Bars show size of intersections between sites which are significant in each two-way comparison. Total intersection sizes displayed in black above each bar. A comparison to previously identified m^6^A sites using the orthogonal technique, miCLIP is included: orange and blue bar fractions show the number of sites within each set intersection that have or do not have an miCLIP peak (Parker et al., 2020) within 5 nt, respectively. Percentage of intersections with miCLIP support is displayed in orange above each bar. A small number of sites are significant in the three-way comparison, but in neither two-way comparison (far left bar). (C–D) Detection of RNAs immunoprecipitated from Col-0 and fio1 mutant alleles with anti-m^6^A antibodies using RT-qPCR analysis. The data are presented as strip-plots with mean and 95% CIs, showing the enrichment of U6 and U2 snRNAs over input using (C) Synaptic systems #202 003 anti-m^6^A antibody and RNA purified from Col-0 and fio1-1, and (D) Millipore ABE572 anti-m^6^A antibody and RNA purified from Col-0, fio1-1, fio1-3, and fio1-4. Y axes show −ΔCt (m^6^A-IP — input) corrected for input dilution factor. Strip-plots show mean values for three or four independent RT-qPCR amplifications on each biological replicate immunopurification experiment.
Figure 2—figure supplement 1. Analysis of RNA modification sites detected by nanopore direct sequencing of RNA purified from *fip-37-4* and *fio1-1*. 
(A) Boxplot showing effect sizes for modification sites which have significant changes in modification rate in both *fip37-4* and *fio1-1* (intersection shown in Figure 3—figure supplement 1A). 
(B) Sequence logo identified at FIP37-dependent modification sites detected by Yanocomp. 
(C) Bar plot showing the mean density of FIP37-dependent modification sites in different genic features annotated in Araport11. Error bars are bootstrapped 95% CIs for means. 
(D) Sequence logo identified at modification sites which are significant in *fio1-1* vs Col-0 comparison, but not in *fip37-4* vs Col-0, detected by Yanocomp, which are also supported by orthogonal miCLIP analysis. 
(E) Bar plot showing the mean density of modification sites which are significant in *fio1-1* vs Col-0 comparison, but not in *fip37-4* vs Col-0, in different genic features annotated in Araport11. Error bars are bootstrapped 95% CIs for means. 
(F) Swarmplot showing the effect size measured in Earth mover distance using d3pendr (*Parker et al., 2021c*) of genes with significant changes in poly(A) site choice in either *fip37-4* or *fio1-1*, compared to Col-0 detected in analysis of nanopore direct RNA sequencing data.
Figure 2—figure supplement 2. Identification of modified bases in RNA encoding Arabidopsis S-adenosylmethionine (SAM) synthetases using nanopore direct RNA sequencing analysis. (A–D) Gene tracks showing predicted methylation sites in (A) MAT1, (B) MAT2, (C) MAT3, and (D) MAT4 predicted by miCLIP (blue), differential error rate method on vir-1 and VIRILIZER complemented line (orange) (Parker et al., 2020), and Yanocomp method on fip37-4 and Col-0 lines (green). FIP37-dependent sites, which also have significant modification rate change in fio1-1, are shown in yellow. A magnified view of the 3'UTR is shown in the right hand panel of each gene track. No FIO1-dependent modification sites or gene body methylation sites were detected in MAT1, MAT2, MAT3, or MAT4, using any of the approaches.
Figure 2—figure supplement 3. (A–B) Relative expression of U6 snRNA in (A) Col-0 and the fio1-1 mutant, and (B) Col-0, fio1-1, fio1-3, and fio1-4 mutants, measured by RT-quantitative (qPCR), compared to U2 snRNA. The means of three or four technical replicates are shown for each biological replicate. Grey bars with points represent mean and 95% CIs.
**Figure 3.** Splicing events sensitive to temperature and loss of FIO1. (A–B) Analysis of Illumina RNA-seq data presented as bar plots showing the proportion of splicing events of each class, as labelled by SUPPA, which have significantly different usage (false discovery rate [FDR] <0.05) at either (A) varying temperatures or (B) in fio1-3. In (B), events for which the response to temperature changes in fio1-3 are shown in orange. (C) Gene track of Illumina RNA-seq reads showing the change in retention of MAF2 intron 3 in fio1-3, at 20°C. Expression is normalised by the read coverage at the −1 position of the 5’ splice site (5’SS). (D) Boxplot of Illumina RNA-seq analysis showing the change in retention of MAF2 intron 3 at varying temperatures and in fio1-3.
Figure 3—figure supplement 1. Analysis of Illumina RNA-seq data. (A) Boxplot showing the reduction in overall gene expression of FIO1 in fio1-3. (B) Volcano plots showing the distribution of alternative splicing events from four classes (alternative 5’ splice sites [5’SSs], alternative 3’ splice sites [3’SSs], skipped exons, and retained introns) in the fio1-3 mutant compared to Col-0. For skipped exons and retained introns, a positive ΔPSI means that the feature is retained more in the fio1-3 mutant, whilst a negative ΔPSI means the feature is spliced out more in the fio1-3 mutant. ΔPSIs are unsigned for alternative 5’SSs and 3’SSs. (C) Histogram showing change in the difference between fio1-3 and Col-0 percent splicing index (PSI) between 4 and 28°C (ΔΔPSI). A positive ΔΔPSI indicates that there is a greater deviation from Col-0 splicing levels in fio1-3 at 28°C than at 4°C.
Figure 3—figure supplement 2. Illumina RNA-seq analysis of differential gene expression between Col-0 and fio1-3. (A–I and K–L) Boxplots showing the increase in overall gene expression of the flowering time activators (A) FT (AT1G65480), (B) SOC1 (AT1G65480), and reduction in the expression of the flowering repressors (C) FLM (AT1G77080), (D) MAF2 (AT5G65050), (E) MAF3 (AT5G65060), (F) MAF4 (AT5G65070), (G) MAF5 (AT5G65080), (H) FLC (AT5G10140), and (I) COOLAIR transcripts expressed antisense to FLC. (J) Gene track of Illumina RNA-seq reads showing the expression of FLC and COOLAIR transcripts at 20°C in Col-0 and fio1-3. The mRNA-level expression of (K) the flowering activator CO (ATSG15840) and (L) the flowering repressor SVP (AT2G22540) was also not affected.
Figure 4. Effect of fio1-3 on alternative 5' splice sites. (A–B) Sequence logos and heatmap showing the distribution of U5 snRNA and U6 snRNA interacting sequence classes for 5' splice sites (5'SSs) which (A) are sensitive to loss of FIO1 function or (B) have increased usage in fio1-3 based on the analysis of Illumina RNA-seq data. Motifs are shown for the −3 to +5 positions of the 5'SS. U5 classes are based upon the distance of the −2 to −1 positions of the 5'SS from the consensus motif AG. U6 classes are based upon the distance of the +3 to +5 positions of the 5'SS from the consensus motif.

Figure 4 continued on next page
motif RAG. (C) Gene track of Illumina RNA-seq reads showing alternative 5'SS usage at AtSAR1 intron 21 in fio1-3, at 20°C. Expression is normalised by the read coverage at the −1 position of the 5’SS. (D) Illumina RNA-seq analysis visualised with a boxplot showing the change in usage of the cryptic alternative 5'SS (Alt 5'SS 2) in AtSAR1 intron 21 at varying temperatures in Col-0 and fio1-3. (E) Gene track showing alternative 5'SS usage at AtSAR1 intron 21 in fio1-1, identified using nanopore direct RNA sequencing (DRS) read alignments. Alignments have been subsampled to a maximum of 50 per condition. (F) Contingency table showing the relationship between the nucleotide at the +4 position, and the direction of change in 5'SS usage in fio1-3, for pairs of alternative 5'SSs with significantly altered usage in fio1-3 analysed in Illumina RNA-seq data. (G) Boxplot showing effect sizes of pairs of alternative 5'SSs with significantly altered usage in fio1-3, separated by +4 position bases (A→U indicates that 5’SS with reduced usage has A+4, 5’SS with increased usage has U+4). (H) Boxplot showing effect sizes of pairs of alternative 5'SSs with significantly altered usage in fio1-3, separated by U5 classification.
Figure 4—figure supplement 1. (A) Histogram showing the distance between alternative 5' splice site (5'SS) pairs with significantly different usage in fio1-3 detected by Illumina RNA-seq analysis. Negative distances represent shifts toward greater usage of upstream 5'SSs, whilst positive distances represent shifts toward greater usage of downstream 5'SSs. (B) Sequence logo for all 5'SSs identified in the condition-specific StringTie assembly generated from Illumina RNA-seq and Nanopore direct RNA sequencing (DRS) data of fio1 mutants. Motifs are shown for the −3 to +5 positions of the 5'SS. (C–D) Sequence logos for 5'SSs which have (C) increased and (D) decreased retention in the fio1-3 mutant. Motifs are shown for the −3 to +5 positions of the 5'SS.
Figure 4—figure supplement 2. (A) Gene track of Illumina RNA-seq reads showing alternative 5' splice site (5'SS) usage at MAF2 intron 2 in fio1-3, at 20°C. Expression is normalised by the read coverage at the +1 position of the 3' splice site (3'SS). (B) Illumina RNA-seq analysis presented as a boxplot showing the change in usage of alternative 5'SS 2 of MAF2 intron 2 at varying temperatures, in Col-0 plants and in fio1-3. (C) Gene track of Illumina RNA-seq reads showing alternative 5'SS usage at FLM intron 3 in fio1-3, at 20°C. Expression is normalised by the read coverage at the +1 position of the 3'SS. (D) Illumina RNA-seq analysis presented as a boxplot showing the change in usage of alternative 5'SS 2 of MAF2 intron 2 at varying temperatures in Col-0 plants and in fio1-3. (E–G) Illumina RNA-seq analysis presented as boxplots showing the change in functional expression of (E) AtSAR1, (F) MAF2, and (G) FLM in the fio1-3 mutant, at varying temperatures, as predicted by TranSuite (Entzine et al., 2020). (H) Gene track of Illumina RNA-seq reads showing alternative 5'SS usage at MTB intron 4 in fio1-3, at 20°C. Expression is normalised by the read coverage at the +1 position of the 3'SS. (I) Illumina RNA-seq analysis presented as a boxplot showing the change in usage of alternative 5'SS 2 of MTB intron 4 at varying temperatures in Col-0 plants and in fio1-3.
Figure 4—figure supplement 3. (A) Contingency table showing the relationship between the bases at the 4+ position for fio1-3-sensitive 5’S splice sites (5’SSs) with reduced usage in fio1-3 and alternative 5’SSs with increased usage in fio1-3 revealed through Illumina RNA-seq data analysis. (B) Contingency table showing the relationship between the bases at the −1 and +5 positions for fio1-3-sensitive and alternative 5’SSs, where both 5’SSs in the pair have the same base at the +4 position. (C) Boxplot showing effect sizes of pairs of alternative 5’SSs with significantly altered usage in fio1-3, separated by +5 position bases (G→H indicates that 5’SS with reduced usage has G+4, 5’SS with increased usage has H+4).
Figure 5. Effect of flo1-3 on retained introns and exon skipping events. (A) Sequence logos and heatmaps showing the distribution of U5 snRNA and U6 snRNA interacting sequence classes for 5’ splice sites (5’SSs) which have increased (left) and decreased (right) retention in flo1-3 based on Illumina RNA-seq data analysis. Motifs are shown for the −3 to +5 positions of the 5’SS. U5 classes are based upon the distance of the −2 to −1 positions of the 5’SS from the consensus motif AG. U6 classes are based upon the distance of the +3 to +5 positions of the 5’SS from the consensus motif RAG. (B) Gene
Figure 5 continued

track of Illumina RNA-seq reads showing intron retention at WNK1 intron 5 in fio1-3, at 20°C. Expression is normalised by the read coverage at the +1 position of the 3’ splice site (3’S’SS). (C) Boxplot of Illumina RNA-seq data analysis showing the change in intron retention of WNK1 intron 5 at varying temperatures, in Col-0 and fio1-3. (D) Gene track showing intron retention at WNK1 intron 5 in fio1-1, identified using nanopore direct RNA sequencing (DRS) read alignments. Alignments have been subsampled to a maximum of 50 per condition. (E–F) Sequence logos for 5’SSs at introns upstream (left) and downstream (right) of exons with (E) increased skipping or (F) increased retention in fio1-3 based on Illumina RNA-seq data analysis.
Figure 5—figure supplement 1. (A) Heatmaps showing the distribution of U5 snRNA and U6 snRNA interacting sequence classes for 5' splice site (5'SSs) of exons (right) and upstream 5'SSs (left) which have increased (above) and decreased (below) skipping in flo1-3 as deduced from analysis of Illumina RNA-seq data. U5 classes are based upon the distance of the −2 to −1 positions of the 5'SS from the consensus motif AG. U6 classes are based upon the distance of the +3 to +5 positions of the 5'SS from the consensus motif RAG. (B) Gene track of Illumina RNA-seq data showing change in exon

Figure 5—figure supplement 1 continued on next page
skipping at MAF3 exon 2 in fio1-3, at 20°C. Expression is normalised by the coverage at the +1 position of the upstream 5’SS. (B) Boxplot visualisation of RNA-seq data showing the change in MAF2 exon 2 inclusion at varying temperatures in Col-0 and in fio1-3. (B) Gene track of Illumina RNA-seq data showing change in exon skipping at PTB1 exon 3 in fio1-3, at 20°C. Expression is normalised by the coverage at the +1 position of the upstream 5’SS. (B) Boxplot showing the change in PTB1 exon 3 inclusion at varying temperatures in Col-0 and in fio1-3.
Figure 6. Effect of fio1-3 on alternative 3’ splice site (3’SS) usage. (A) Histogram showing the distance between alternative 3’SS pairs with significantly different usage in fio1-3 revealed by Illumina RNA-seq data analysis. Negative distances represent shifts toward greater usage of upstream 3’SSs, whilst positive distances represent shifts toward greater usage of downstream 3’SSs. (B–C) Sequence logos for 5’ splice sites (5’SSs; left), upstream 3’SSs (middle), and downstream 3’SSs (right) at pairs of alternative 3’SSs with increased (B) upstream or (C) downstream usage in fio1-3 revealed by Illumina RNA-seq data analysis.
Illumina RNA-seq data analysis. 5'SS logos are for −3 to +5 positions, and 3'SS logos are for −5 to +2 positions. (D) Gene track of Illumina RNA-seq reads showing alternative 3'SS usage at LHY intron 5 in fio1-3, at 20°C. Expression is normalised by the read coverage at the −1 position of the 5'SS. (E) Boxplot of Illumina RNA-seq data analysis showing the change in usage of the upstream alternative 5'SS (Alt 5'SS 2) in LHY intron 5 at varying temperatures in Col-0 and in fio1-3. (F) Gene track showing intron retention at LHY intron 5 in fio1-1, identified using nanopore direct RNA sequencing (DRS) read alignments. Alignments have been subsampled to a maximum of 50 per condition. (G) Boxplot of Illumina RNA-seq data analysis showing the change in usage of downstream 3'SSs in alternative 3'SS pairs with different 5'SS +4 bases, separated by whether the base at the −3 position of the two alternative 3'SSs is the same (e.g. CAG\CAG\) or different (e.g. UAG\CAG\).
Figure 6—figure supplement 1. (A) Upset plot derived from Illumina RNA-seq data analysis showing the overlap of fio1-sensitive 5' splice sites (5'SSs) involved in alternative 5'SS usage, with 5'SSs at fio1-sensitive intron retention events and alternative 3' splice sites (3'SSs). (B) Gene track of Illumina RNA-seq data showing alternative 3'SS usage at MAF2 intron 4 in fio1-3, at 20°C. Expression is normalised by the coverage at the −1 position of the 5'SS. (C) Boxplot showing the change in usage of alternative 3'SS 2 of MAF2 intron 4 at varying temperatures in Col-0 and fio1-3. (D) Sequence logos for all upstream and downstream alternative 3'SS pairs identified from the RNA-seq data.
Figure 6—figure supplement 2. (A) Line-plot showing GC content of Illumina RNA-seq reads from Wang et al., 2022. One replicate of Col-0 data was discarded due to extreme negative GC-bias likely resulting from PCR overamplification. (B) Sequence logos for 5’ splice sites (5’SSs) identified from RNA-seq data taken from Wang et al., 2022, which (above) are sensitive to loss of FIO1 function or (below) have increased usage in fio1-1.
Figure 6—figure supplement 3. (A–B) Sequence logos for 5' splice sites (5'SSs) identified from Illumina RNA-seq data of (A) fio1-1 and (B) fio1-5 alleles reanalysed from Sun et al., 2022, which (left) are sensitive to loss of FIO1 function or (right) have increased usage in the respective fio1 mutants. (C) Upset plot showing the intersection of the sets of alternative 5'SS events which are identified from Illumina RNA-seq data of fio1-1 and fio1-5 alleles from Sun et al., 2022 independent RNA-seq of the fio1-3 allele from this study.
Figure 6—figure supplement 4. (A–F) Gene tracks showing (A and D) alternative 5' splice site (5'SS) selection at AtSAR1 intron 21, (B and E) intron retention at WNK1 intron 5, and (C and F) alternative 5'SS selection at MTB intron 4 in (A–C) the fio1-2 mutant, identified using nanopore direct RNA sequencing (DRS) read alignments reanalysed from Xu et al., 2022 and (D–F) fio-1 and fio1-5 mutants, identified using nanopore DRS read alignments reanalysed from Sun et al., 2022.
Figure 7. Global analysis of U5 and U6 interaction strengths reveals anticorrelation. (A) Sequence logos of annotated Arabidopsis splice sites showing base frequency probabilities at −3 to +5 positions for (above) 5’ splice sites (5’SSs) with //GURAG sequence and (below) all other 5’SSs. (B) Empirical cumulative distribution function of Arabidopsis U5 position-specific scoring matrix (PSSM) log-likelihood scores for 5’SSs with either //GURAG sequence or all other 5’SSs. U5 PSSM scores are calculated using a PSSM derived from all 5’SSs in the Araport11 reference annotation, at the −2 to −1 positions of the 5’SS, inclusive. (C) Arabidopsis sequence logos showing base frequency probabilities at the −3 to +5 positions for (above) 5’SSs with AG//GU sequence and (below) all other 5’SSs. (D) Empirical cumulative distribution function of Arabidopsis U6 PSSM log-likelihood scores for 5’SSs with different U5 classes. U6 PSSM scores are calculated using a PSSM derived from all 5’SSs in the Araport11 reference annotation, at the +3 to +5 positions of the 5’SS, inclusive. (E–F) Scatterplot showing the ratio of PSSM log-likelihoods (log-odds ratio) for U5 and U6 snRNA interacting sequences, at pairs of upstream and downstream alternative 5’SSs in (E) the Arabidopsis Araport11 reference annotation or (F) the H. sapiens GRCh38 reference annotation. A positive log-odds ratio indicates that the PSSM score of the upstream 5’SS is greater than that of the downstream 5’SS.
Figure 7—figure supplement 1. (A, C, E, and G) Sequence logos showing base frequency probabilities at −3 to +5 positions for (above) 5' splice sites (5' SSs) with //GURAG sequence and (below) all other 5' SSs, in the organisms (A) C. elegans, (C) D. melanogaster, (E) D. rerio, and (G) H. sapiens. (B, D, F, and H) Empirical cumulative distribution function of U5 snRNA (left panels) and U6 snRNA (right panels) position-specific scoring matrix (PSSM) log-likelihood scores for 5' SSs falling into different U6 (left panels) or U5 (right panels) sequence classes, for all 5' SSs in the organisms (B) C. elegans, (D) D. melanogaster, (F) D. rerio, and (H) H. sapiens.
melanogaster, (F) D. rerio, and (H) H. sapiens. PSSM scores are calculated using a PSSM derived from all 5'SSs in the corresponding genome annotation of each organism, at the −2 to −1 positions of the 5'SS for U5 and the +3 to +5 positions for U6, inclusive.
Figure 7—figure supplement 2. (A–B) Scatterplot showing the ratio of position-specific scoring matrix (PSSM) log-likelihoods (log odds ratio) for U5 snRNA and U6 snRNA interacting sequences, at pairs of upstream and downstream alternative 5′ splice sites (5′SSs) in the (A) *C. elegans* or (B) *D. rerio* reference annotation. A positive log-odds ratio indicates that the PSSM score of the upstream 5′SS is greater than that of the downstream 5′SS.
Figure 8. U6 m^6A-5'SS interactions during splicing. (A) Model depicting U5 and U6 snRNA interactions with two major classes of 5'SS: //GURAG and AG//GU 5'SSs. //GURAG 5'SSs form strong interactions with U6 snRNA ACAGA (darkly shaded) and weaker interactions with U5 snRNA loop 1 (lightly shaded). AG//GU 5'SSs form strong interactions with U5 snRNA loop 1 and weaker interactions with U6 snRNA ACAGA. (B) Cryo-electron microscopy (cryo-EM) analysis of human pre-B and B complexes with RNA interactions detailed in the expanded section (PDB 6AHD; Bertram et al., 2017) and Prp8 shown in the background as a common scaling reference. The U6 snRNA ACAGA and U5 snRNA loop 1 sequences are missing from cryo-EM structures at this stage, probably because they present as flexible loops. In B complex, C42 and 5'SS G+5 form a canonical Watson–Crick pair. m^6A43 and 5'SS A+4 form a trans Hoogsteen sugar edge interaction (Leontis et al., 2002) that caps and stabilises the U6/5'SS helix by stacking because U6 snRNA G44 and 5'SS A+3 have not yet formed a stable interaction. The 5'SS is kinked, and U5 snRNA loop 1 is docked on the upstream exon. The methyl group of U6 snRNA m^6A43 is not modelled in the structure due to lack of resolution. (C) Model depicting U6 m^6A interactions at different stages of splicing. In B complex, U6 m^6A43 stabilises the U6/5'SS helix by stacking. As the active site forms in B^act, this role becomes less important because U6

Figure 8 continued on next page
snRNA G44 interacts more stably with 5'SS₁₃ and U6 A45 stacks on the helix stabilised by R554 of SF3B2. In C complex, the U6/5'SS helix is stabilised by NS7 of HJU2. In C* complex, the U6mA43:5'SS₁₄ interaction becomes more important again because the 5'SS₁₃ pivots to a new position. The mA43 and 5'SS₁₄ pair forms part of a continuous helical stack with the docked 3' splice site (3'SS), which is capped by the interaction between 3'SS₃ and Q1522 of Prp8. For more detail, see Figure 8—figure supplement 1.
Figure 8—figure supplement 1. Tracing m^6^A modified U6 snRNA in cryo-electron microscopy (cryo-EM) structures of different spliceosomal complexes with cryo-EM reveals in pre-B complex (PDB 6QX9, Charenton et al., 2019), the ACAGA box is flexible and disordered. In B complex (PDB 6AHD, Bertram et al., 2017), m^6^A forms a trans Hoogsteen sugar edge interaction with 5’ splice site (5’SS) A+4, capping the U6/5’SS helix by stacking. A kink in the 5’SS is detectable, and U5 snRNA loop 1 (dark blue) is docked on the upstream exon (yellow). In the Bact complex (PDB 5Z56, Zhang et al., 2018), U6 G44 pairs with 5’SS A+3, capping and stabilising the U6/5’SS helix. This interaction is further stabilised by R554 of the spliceosomal protein SF3B2. Consequently, the role of m^6^A:5’SS A+4 in stabilising the helix appears less important at this stage. In C complex (PDB 6ZYM, Bertram et al., 2020), the 5’SS/U6 helix is stabilised by R3 of Prp8 and N57 of hYJU2. Rearrangements in P complex for the second splicing reaction suggest the stabilising role of U6 m^6^A43:5’SS A+3 becomes important again because 5’SS A+3 pivots away to form new stacking interactions (PDB 6QDV, Fica et al., 2019). The 3’
splice site (3'SS) G₁, stacks on U₆ snRNA A₄₅, which now forms a non-canonical base-pair with 5'SS U₂. A continuous helix stack involving the U₆/5'SS interaction between U₆ m₆A₄₃ and 5'SS A₄₄, and the docked 3'SS, is capped by interaction between 3'SS₃ and Prp8 Q₁₅₂₂. A domain of the major architectural spliceosomal protein, Prp8, is depicted in the background as a constant reference through each stage. The methyl group of U₆ snRNA m₆A₄₃ is not modelled in the structures due to lack of resolution.