

Article

# In Silico Analysis of a Highly Mutated Gene in Cancer Provides Insight into Abnormal mRNA Splicing: Splicing Factor 3B Subunit 1<sup>K700E</sup> Mutant

Asmaa Samy <sup>1</sup>, Baris Ethem Suzek <sup>2</sup>, Mehmet Kemal Ozdemir <sup>3</sup> and Ozge Sensoy <sup>3,4,\*</sup>

<sup>1</sup> The Graduate School of Engineering and Natural Science, Istanbul Medipol University, 34810 Istanbul, Turkey; amahmoud@st.medipol.edu.tr

<sup>2</sup> Department of Computer Engineering, Muğla Sıtkı Koçman University, 48000 Muğla, Turkey; barissuzek@mu.edu.tr

<sup>3</sup> The School of Engineering and Natural Science, Istanbul Medipol University, 34810 Istanbul, Turkey; mkozdemir@medipol.edu.tr

<sup>4</sup> Regenerative and Restorative Medicine Research Center (REMER), Istanbul Medipol University, 34810 Istanbul, Turkey

\* Correspondence: osensoy@medipol.edu.tr

Received: 16 January 2020; Accepted: 20 April 2020; Published: 28 April 2020



**Abstract:** Cancer is the second leading cause of death worldwide. The etiology of the disease has remained elusive, but mutations causing aberrant RNA splicing have been considered one of the significant factors in various cancer types. The association of aberrant RNA splicing with drug/therapy resistance further increases the importance of these mutations. In this work, the impact of the splicing factor 3B subunit 1 (SF3B1) K700E mutation, a highly prevalent mutation in various cancer types, is investigated through molecular dynamics simulations. Based on our results, K700E mutation increases flexibility of the mutant SF3B1. Consequently, this mutation leads to i) disruption of interaction of pre-mRNA with SF3B1 and p14, thus preventing proper alignment of mRNA and causing usage of abnormal 3' splice site, and ii) disruption of communication in critical regions participating in interactions with other proteins in pre-mRNA splicing machinery. We anticipate that this study enhances our understanding of the mechanism of functional abnormalities associated with splicing machinery, thereby, increasing possibility for designing effective therapies to combat cancer at an earlier stage.

**Keywords:** cancer; somatic mutations; aberrant RNA splicing; SF3B1; missense mutations; molecular dynamics simulations

## 1. Introduction

Cancer is considered a complex disease caused by a combination of predisposing genetic variants and environmental factors [1]. Moreover, the genetic background of the mutation and/or epigenetic alterations further increase the risk of development of the disease. The somatic mutations affecting the “pre-mRNA splicing machinery” are particularly critical as they cause aberrant splicing [2–4] that leads to formation of proteins with altered functions. Various studies have shown that these altered proteins are dominantly involved in processes such as tumor metabolism, apoptosis, cell cycle control, invasion, metastasis, and angiogenesis [5,6].

Processing of an intron-containing pre-mRNA consists of three steps: i) intron excision, ii) 3'-end processing, and iii) transportation of mature transcripts from nucleus into the cytoplasm [7]. The splicing process is carried out by a complex ribonucleoprotein (RNP) called spliceosome [8,9] which is composed of five small nuclear ribonucleoproteins (snRNPs) (U1, U2, U4, U5, and U6)

along with other related proteins [10]. SF3b complex is the major component of U2 small nuclear ribonucleoprotein (U2snRNP) and it is responsible for reinforcing the interaction between U2 snRNA and the branch site on pre-mRNA in the first transesterification reaction of splicing process [11–14].

SF3b is composed of seven subunits: SF3B1, SF3B2, SF3B3, SF3B4, SF3B5, SF3B6, and SF3B7 [15]. Among them, splicing factor 3B subunit 1 (SF3B1) is the largest component of the SF3b complex [16,17] and it is composed of N-terminal helix-loop-helix domain (NTD) and C-terminal of 20 HEAT-repeat domain (HD) [15,18]. SF3B1 interacts directly with SF3B3, SF3B7, and SF3B5, all of which facilitate its binding to pre-mRNA. In addition, SF3B1 also interacts with SF3B6 (also known as p14) [15,19,20], two of which form a complex with U2 auxiliary factor, U2AF65, to recognize and select the branch splicing site properly [12,14,20]. In particular, the superhelical heat domain of SF3B1 provides a platform for effective binding of interacting partners and/or RNAs to the complex [21].

The point mutations occurring in the heat domain of SF3B1 promote usage of alternative branch points, hence abnormal 3' splice sites. Consequently, this causes either (partial or complete) intron retention or excision of an exon in pre-mRNA. Indeed, it has been shown that SF3B1 mutations increased the number of transcripts with retained introns as seen in myelodysplastic syndrome [22] or they caused production of transcripts that included a shorter copy of the exon as seen in chronic lymphocytic leukemia [3]. In another study, it has been shown that mutations affecting SF3B1 gene caused alternative splicing of key genes such as UQCC, F8, ABCC5, and GUSBP11 in ER (estrogen receptor)-positive breast cancer [23]. The abnormal transcripts follow two possible pathways: 1) translation of mRNA into proteins with truncated or missing domain(s) which consequently leads to change, loss or gain of function of the protein [24] and 2) degradation by nonsense-mediated mRNA decay (NMD) mechanism which leads to a reduction in the proteome content of the cell [25]. Therefore, understanding the molecular mechanism behind aberrant splicing may help in preventing cancer development or progression; however, the number of related studies is scarce. Recently, Borisek et al. studied the impact of K335E, L378V, and N295D mutations on Hsh155, which is the yeast homolog of human SF3B1, using atomistic molecular dynamics to investigate the mechanism underlying constitutive/alternative/aberrant splicing [26]. In another recent study, Borisek et al. investigated the impact of splicing modulators such as pladienolides, herboxidienes, and spliceostatsins, which have been shown to exert anti-tumor activity, using mutant Y36C and R38C of PHF5A as well as R1074H of SF3B1, two of which are the components of SF3b [27].

In this study, we utilized computational biophysics tools to investigate the impact of K700E mutation on structure and dynamics of SF3B1. Comparative analysis of molecular dynamics trajectories of wild type and mutant protein revealed that the mutation i) increased the flexibility of the protein, thus preventing proper alignment of pre-mRNA within the splicing machinery, ii) caused disruption of interactions between p14 and pre-mRNA. K700E mutation is shown to be associated with various cancer types including myelodysplastic syndrome (MDS), chronic lymphocytic leukemia (CLL), uveal melanoma (UV), and pancreatic cancer [28]. Therefore, our results can provide insights into the molecular mechanism of aberrant splicing in various cancer types, highlighting the importance and general applicability of our work.

## 2. Results

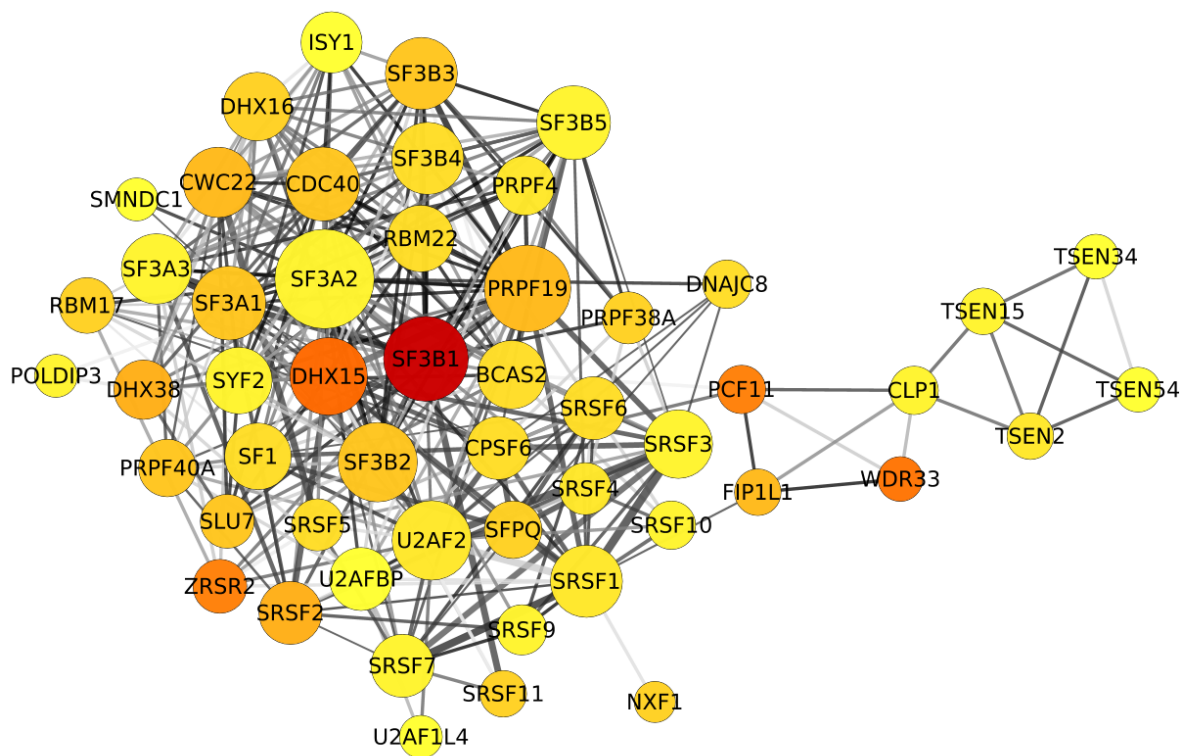
### 2.1. SF3B1 Gene Has the Highest Number of Mutations in the Pre-mRNA Splicing Network

We identified a somatic mutation to further study, namely SF3B1<sup>K700E</sup>, factoring in two aspects; the associated gene's centrality in pre-mRNA splicing machinery and the prevalence of the somatic mutations it carries. To this end, we constructed a pre-mRNA splicing network and superimposed the number of mutations obtained from latest version of COSMIC database on this network.

The network analysis revealed the top six genes with high centrality in the network which are SF3B1, splicing factor 3A subunit 2 (SF3A2), DEAH-Box helicase 15 (DHX15), U2 auxiliary factor 2 (U2AF2), pre-mRNA-processing factor 19 (PRPF19), and serine/arginine-rich splicing factor 1 (SRSF1).

SF3A2 is the only component in the SF3a complex that recognizes the U2/BS duplex [18,29]. DHX15 is required to catalyze the disassembly of spliceosome complex after the completion of splicing and releasing mature mRNA to cytoplasm [30–32]. U2AF2 recognizes and binds to the polypyrimidine (Py) tract on pre-mRNA [33]. PRPF19 (also known as PRP19) interacts with other components of the NineTeen complex (NTC) that is essential for NTC stability [34], which is responsible for strengthening interactions between mRNA and snRNAs [35]. SRSF1 (also known as SF2/ASF) regulates pathways of mRNA metabolism such as pre-mRNA splicing [36], export [37], and translation [36]. The centrality metrics of SF3B1 including closeness, radiality, and eccentricity were equal to 0.58, 0.88, and 5, respectively which are slightly lower than the other five genes. However, SF3B1 was the most frequently mutated gene (associated with 34 highly pathogenic missense mutations) among other genes included in the constructed pre-mRNA splicing network. SF3A2, DHX15, U2AF2, PRPF19, and SRSF1 had 1, 13, 2, 6, and 2 highly pathogenic missense mutations, respectively.

To maximize the general applicability of our study, we favored the prevalence of mutation over gene network centrality in deciding the mutation to further study. SF3B1 was a gene containing the highest number of mutations while having comparably high centrality in the constructed pre-mRNA splicing network to other genes as shown in Figure 1. Among the 34 highly pathogenic SF3B1 missense mutations, SF3B1 K700E mutation was the most prevalent mutation in various cancer types whose significance supported by previous experimental studies. Hence, we decided to study the impact of K700E mutation on the structure and dynamics of by means of MD simulations.

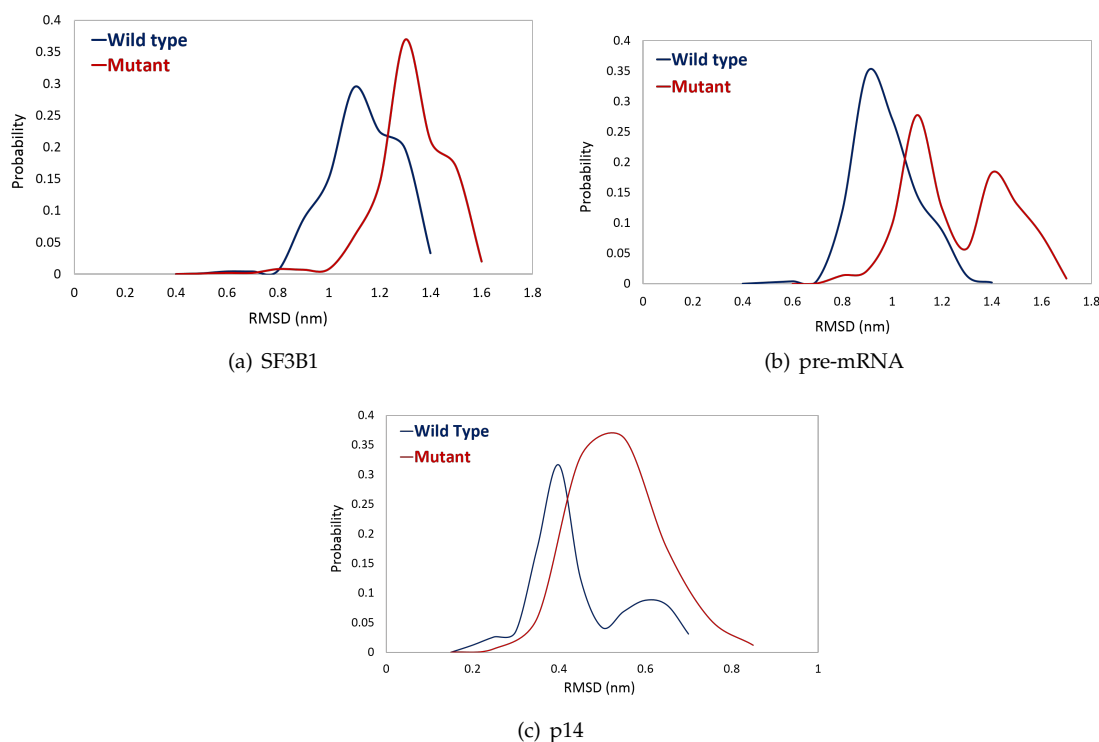


**Figure 1.** Depiction of the extracted network's component. It consists of 54 nodes and 382 edges. In this component, nodes represent genes that are connected with 382 edges depending on different experimental data. Nodes' color is mapped according to the number of associated mutations whereas nodes' size is mapped according to their degree. An edge between every two nodes corresponds to their experimental interaction as explained in the methods section.

## 2.2. K700E Mutation Affects the Stability of SF3B1, Pre-mRNA, and p14

K700E mutation is located in the heat domain of SF3B1 and is within the interaction distance from pre-mRNA (See Figure 8). The distance between the side chain of K700 and pre-mRNA is smaller than

0.20 nm. Therefore, we set out to investigate the impact of the mutation on the stability of pre-mRNA and also SF3B1 using trajectories obtained from MD simulations. The results showed that the mutation decreased the stability of pre-mRNA and SF3B1, which is evident from the right-shift in probability distributions of RMSD in the mutant system as shown in Figure 2a,b. We also did the same analysis for p14, which is shown to be involved in selection of the branch point on pre-mRNA, and observed a similar impact as well.

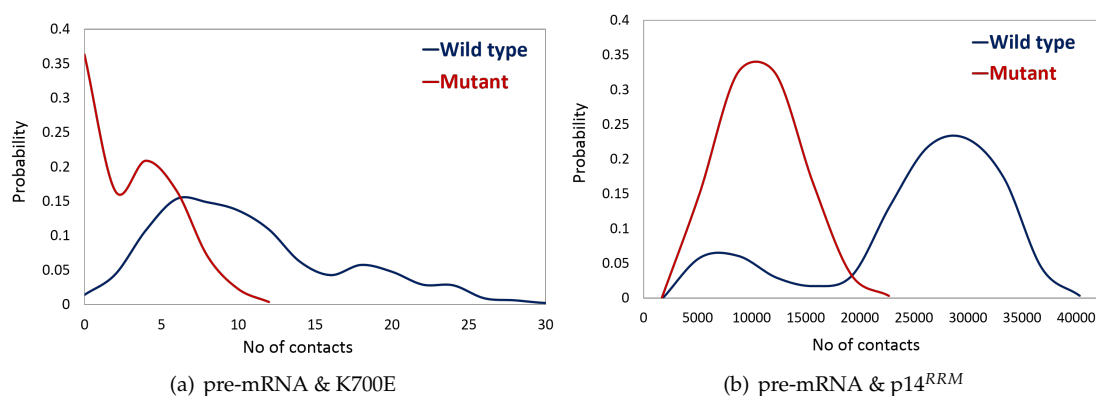


**Figure 2.** Probability plots of root mean square deviation (RMSD) distributions of (a) SF3B1<sup>wt</sup> in blue vs. SF3B1<sup>K700E</sup> in red. (b) pre-mRNA in the wild type structure in blue vs. mutant structure in red. (c) p14 in the wild type structure in blue vs. mutant structure in red.

### 2.3. K700E Mutation Weakens Interactions Formed with Pre-mRNA

As shown in the above section, K700E mutation decreases stability of both SF3B1 and pre-mRNA, which presumably impacts interactions formed between the heat domain of SF3B1 and pre-mRNA. To test this, we calculated the number of contacts between K700 and pre-mRNA. The results showed that contact number was smaller in the mutant protein than in the wild type as shown in Figure 3a. Specifically, the number of contacts reached to zero after 250ns in the mutant protein. Presumably, this drastic reduction in the number of contacts has resulted from disruption of electrostatic interactions formed between pre-mRNA and K700. In the wild type, positively charged side chain of lysine “K700” attracts negatively charged phosphate backbone atoms of pre-mRNA [38], whereas negatively charged side chain of glutamic acid “E700” causes repulsion in the mutant system.

Additionally, we also calculated the number of interactions formed between pre-mRNA and p14, in particular, side chains of residues (20–100), as they are shown to be cross-linked to adenosine nucleotide at the branch point during splicing [29,39], hence known as RNA recognition motif (RRM) [20]. The results showed that the number of contacts formed between p14 and pre-mRNA decreased in the mutant protein as shown in Figure 3b.

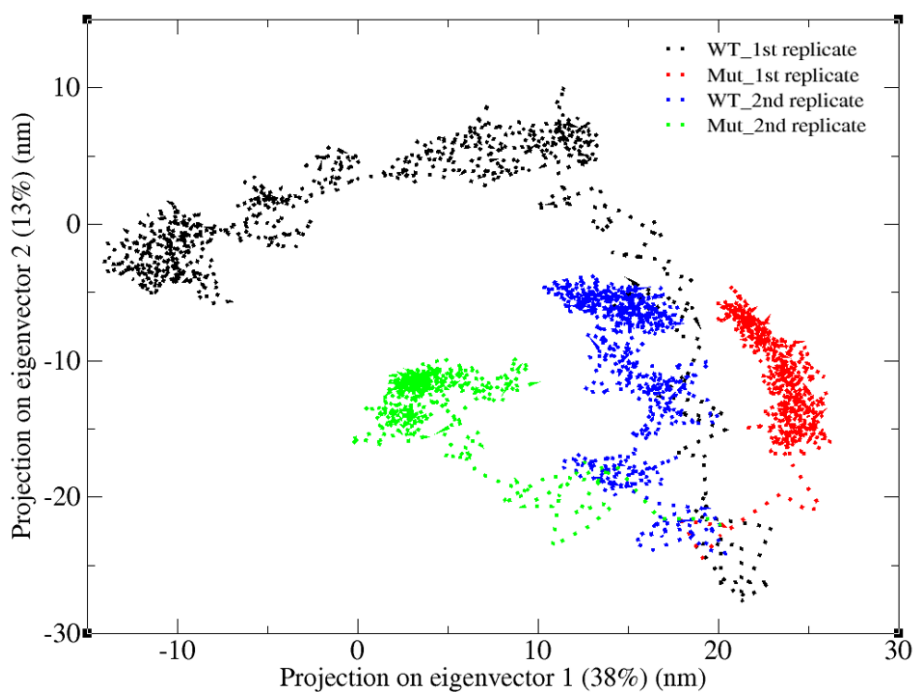


**Figure 3.** Probability plots of distributions of the number of contacts formed between (a) pre-mRNA and side chain of amino acid K700 in wild type shown in blue vs. E700 mutant shown in red, with cutoff distance  $\leq 0.3$  nm. (b) pre-mRNA and side chains of p14 residues (20–100) in wild type shown in blue vs. mutant shown in red, with cutoff distance  $\leq 3$  nm.

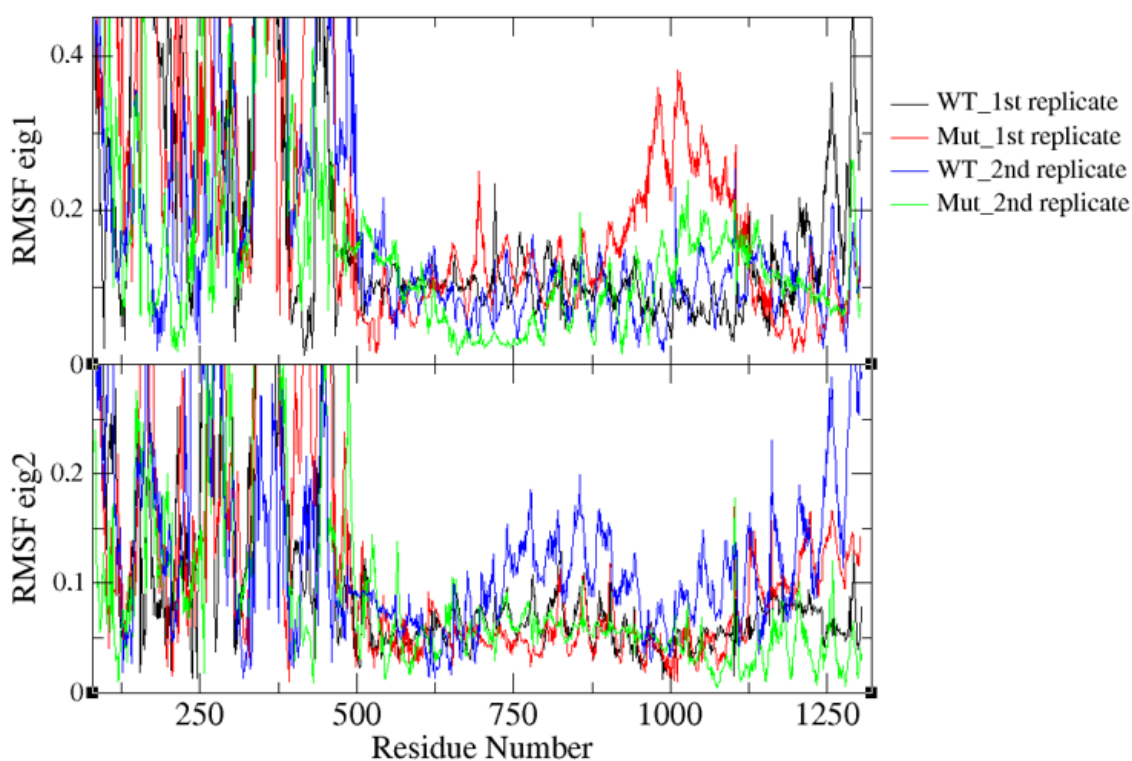
#### 2.4. K700E Mutation Impacts the Global Dynamics of SF3B1

As shown in the previous section, K700E mutation increased the flexibility of both SF3B1 and pre-mRNA. To further investigate the impact of the mutation on global dynamics of SF3B1, we performed essential dynamics analysis. To do so, we projected trajectories of both wild type and mutant SF3B1 systems along the first and the second eigenvectors as shown in Figure 4. With that, we could capture more than 50% of overall dynamics in these systems. The results showed that both replicates of the wild type system followed similar paths even though the first one sampled a wider space than the second one. Presumably, the protein in the second replicate might be trapped (See blue in Figure 4) in one of the many local energy minima present on the rugged potential energy surface. On the other hand, the two replicates followed completely different paths (See green and red color in Figure 4) in the mutant system which might be due to the increase in the protein flexibility upon mutation.

We further computed the RMSF values per  $C\alpha$  atoms of the protein along the first and the second eigenvectors of both wild type and mutant SF3B1 which cumulatively account for more than 50% of overall motion in the systems studied. The results showed that residues 900–1125 displayed relatively higher fluctuation in the first and the second replicate of the mutant system (See red and green color in the first eigenvector) as shown in Figure 5. On the other hand, we observed relatively higher fluctuation for residues 750–900 and 1125–1300 in the wild type system (See blue color in the second eigenvector); however, the contribution of these residues to the overall dynamics is smaller (15%) than did by the first eigenvector (37%).

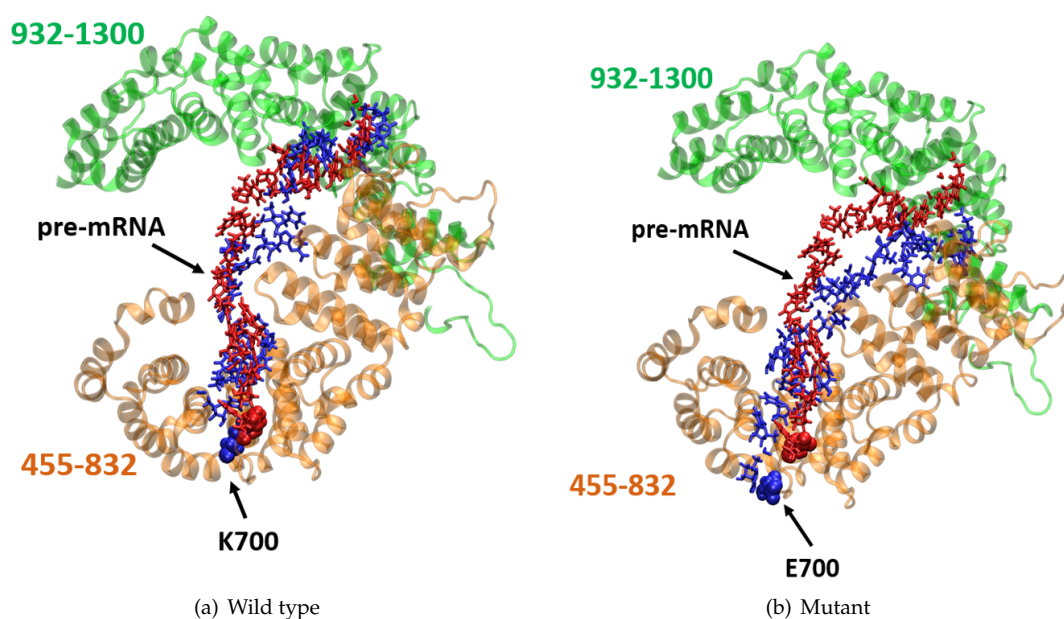


**Figure 4.** 2D projection of SF3B1 trajectories for two replicates of the wild type protein (black and blue) and two replicates of the mutant system (red and green) along the first two eigenvectors of SF3B1<sup>K700</sup> system.



**Figure 5.** Root mean square fluctuation (RMSF) per C $\alpha$  atoms of the first two eigenvectors of SF3B1 obtained for two replicates of the wild type protein (black and blue) and two replicates of the mutant system (red and green).





**Figure 7.** 3D structure of the two highlighted regions surrounding pre-mRNA in wild type system (a) and mutant system (b).  $\alpha$  helical region (455–832) in orange,  $\alpha$  helical region (932–1300) in green, both of pre-mRNA and K700E residue are shown in red representing the initial frame of simulation and blue color representing the last frame. The structural image was created using VMD [40].

### 3. Discussion

Identification of aberrant splicing, as one of the significant cancer drivers, has brought SF3B1 into prominence due to its crucial role in recognition of the proper branch sites during splicing process. The superhelical heat domain of SF3B1 is particularly critical for fulfilling its function and has been shown to be mutated in various cancer types. Therefore, there is an urgent need for understanding the impact of these mutations on structure and dynamics of splicing factor, SF3B1, which will help improve drug discovery studies that target mutant protein.

As a first step in that direction, we provided a mechanistic insight into the mechanism of altered branch point usage elicited by K700E mutation. We showed that the mutation affected interactions of pre-mRNA with SF3B1 and p14, which is in line with functional studies that demonstrate that point mutations in SF3B1 alter branch point selection [41] and also interactions between mutant SF3B1 and other spliceosome components [42,43].

Herein, it is important to emphasize that the altered interaction of p14 with pre-mRNA is novel and warrants follow-up *in vitro* models. Moreover, findings of dynamic cross correlation analysis provide an insight into the impact of mutation on interactions between SF3B1 and other proteins that are not part of the spliceosome. For instance, it has been shown that PPP1R8 interacts with SF3B1 through amino acids between 223–491 [44] and the cross correlation analysis showed that the interaction network among residues 455–832 changed upon mutation. Therefore, it is likely that the mutation distorts interaction also between SF3B1 and PPP1R8, thus leading to problems in (dis)assembly and catalytic steps of splicing. The information regarding network modulation can also be used to determine novel sites on SF3B1 to which small therapeutic molecules can bind and revert the mutant protein to its original function. Alternatively, this knowledge can also help improve existing drug discovery strategies that target mutant SF3B1 as in the case of riboside analog 8-azaguanine [45], H3B-8800 [46], and E7107 [47]. Among them, E7107 has been shown to enhance antitumor response when combined with the proteasome inhibitor, bortezomib, in triple-negative breast cancer, thus highlighting the possible usage of mutant SF3B1 inhibitors in combination therapies.

Our work, though need experimental validations, sheds light on the unknown mechanism of altered branch point usage caused by SF3B1-K700E mutation. Considering that the mutation is



frequently seen in myelodysplastic syndrome (MDS), chronic lymphocytic leukemia (CLL) and breast cancer our results can be used to study other cancer types, and mutant SF3B1 can be considered a potential drug target for cancer therapy. Consequently, this demonstrates the general applicability of the study.

#### 4. Materials and Methods

Previous studies showed that mutant genes, which are involved in cancer [48], encode abnormal proteins that interfere with various vital biological processes such as cell proliferation, transcription, splicing, immune response [49], cell communication, and apoptosis [50]. Among them, aberrant splicing is considered one of the most significant driving factors for development and progression of cancer [51], as it leads to production of proteins with altered function. Therefore, “pre-mRNA splicing” is selected as the target biological process in this study.

As mentioned above, (SF3B1)<sup>K700E</sup> mutant has been shown to be associated with various cancer types. As such, we wanted to investigate if it is among those having higher centrality and associated with the highest number of mutations in the gene network that is involved in pre-mRNA splicing before assessing the impact of the mutation *in silico*. To do so, we first identified the genes that participate in pre-mRNA splicing and then constructed the corresponding network as explained below.

##### 4.1. Identification of Genes Participating in Pre-mRNA Splicing and Constructing the Related Gene Network

To ensure the general applicability of our findings around the mutation studied and to optimize the use of our limited computing resources, we created a generic pipeline to prioritize the genes/mutations that will be studied through molecular dynamics experiments. To this end, we devised a method that factors in two aspects; (1) the prevalence of the mutation in various cancer types, and (2) the centrality of the mutant gene in the target functional context, namely pre-mRNA splicing.

We first used Reactome database [52] to get the list of genes that are involved in mRNA splicing. First, we got 3734 genes which were decreased down to 86 upon filtering based on species as “homo sapiens” and the type as “protein”. STRING database [53] was used to construct a network of genes from these 86 proteins. It consisted of five connected components; four small isolated components consist of 5, 3, 3, and 2 nodes which are connected by 10, 2, 2, and 1 edges, respectively which is shown in Figure 1. The complete network has 67 nodes connected by 397 edges and 12 disconnected nodes (Figure S1), where the nodes represent genes and edges represent experimental interactions taken from different studies such as crystallography and purification assays, cross-linking, and expression level measurements.

We used NetworkAnalyzer plugin [54] within Cytoscape software platform [55] to compute topological parameters and centrality metrics for individual genes (i.e., nodes) on the constructed pre-mRNA splicing network, including closeness, radiality, and eccentricity. Thereafter, we also included somatic mutation frequency data coming from COSMIC database (September 2019 release, the latest version) [56] in our calculations to determine the gene with the highest number of mutations. COSMIC database was filtered based on the following criteria:

1. For practical reasons, we chose breast cancer, one of the most common cancer type, with large number of samples in COSMIC.
2. Only the histological type of “Carcinoma” was used as it is the most common histological type covering the majority of the breast malignancies [57,58].
3. Only “Substitution-Missense” somatic mutations are included. According to COSMIC statistics, 82% of breast carcinoma patients have somatic mutations of type “Substitution-Missense”.
4. Only somatic mutations that are predicted strictly as pathogenic mutations are included. Functional Analysis through Hidden Markov Models (FATHMM) is a predictive tool for the mutation’s impact and classifies any mutation as one of the following classes; pathogenic, neutral or not specified. In addition, FATHMM scores are given in the form of *p*-value that ranges from 0 to

1. By default, the pathogenic mutations have scores greater than 0.5. For our purposes, we further increased the threshold for the pathogenic mutations to 0.9 to avoid false positive predictions.

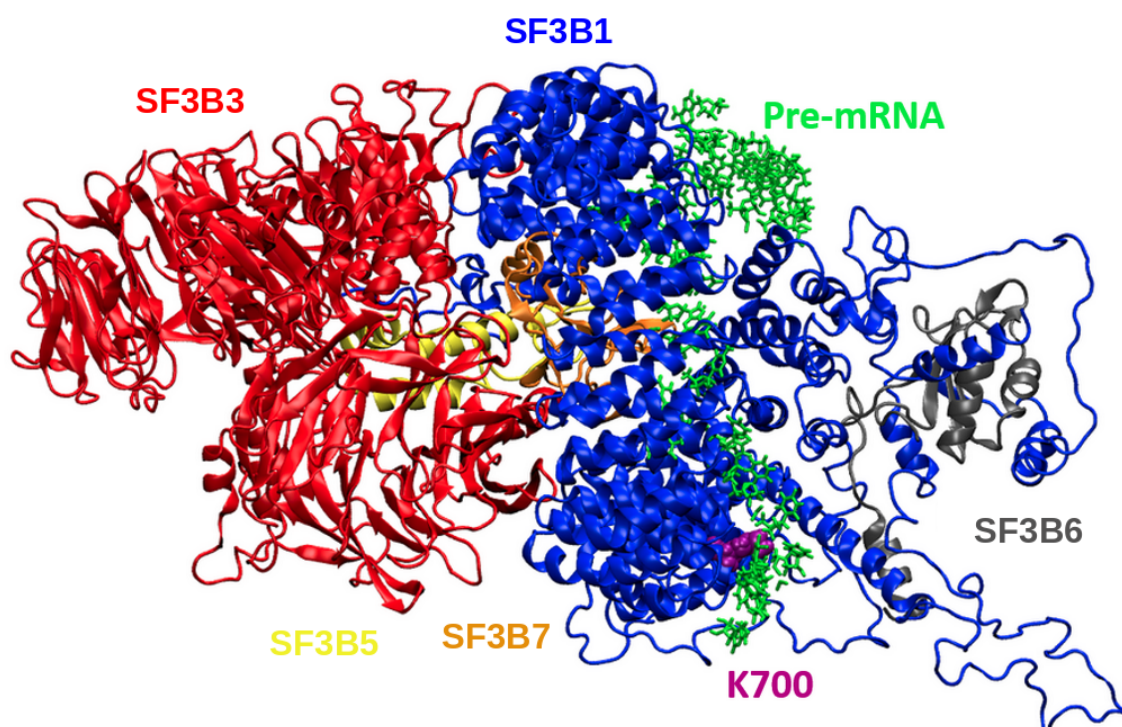
From the filtered somatic mutation dataset, we computed the number of mutations associated with each gene, and then we superimposed these data on the constructed pre-mRNA splicing. SF3B1 emerged a gene containing the highest number of mutations; K700E being the most frequent one, while having comparably high centrality in the constructed pre-mRNA splicing network to other genes as shown in Figure 1.

#### 4.2. Assessment of the Impact of the Mutation by Means of Molecular Dynamics Simulations

In 2015, Maguire et al. analyzed available exome and whole genome sequencing data that belong to 1293 tumor samples and showed that SF3B1-K700E mutation appeared in 17 of 23 mutant samples [23]. Later, in 2018, Seiler et al. reanalyzed whole-exome sequencing data, which included 33 cancer types, to identify SF3B1-related somatic mutations. Consequently, a group of mutations was identified as hotspots regarding SF3B1, namely, R625C/H, E622D/Q, N626D/H/Y, K666E/N/T, K700E, G742D, L833F, E860K, and E862K [59]. Therefore, we utilized the most recent version of COSMIC database which includes almost all studies that report on the SF3B1-K700E mutant involved in breast cancer. In this way, we could i) identify the most frequent mutation among other hotspot mutations that are associated with the target gene in breast cancer and also ii) include a larger and more recent pool of samples. The results showed that K700E was detected in 45 samples while an unknown mutation was detected in 21 samples among 115 samples studied. Moreover, each of G83=/L982= (synonymous substitution), K666Q, R625C, and T935K mutations were detected only in 2 samples, and each of the remaining 44 mutations were detected in only one sample. Therefore, K700E substitution mutation (lysine  $\Rightarrow$  glutamic acid at position 700) was found to be the one with the highest frequency (40%) among other mutations associated with SF3B1. Consequently, K700E mutation was selected to be used in subsequent molecular dynamics (MD) simulations.

##### 4.2.1. System Setup

Cryogenic electron microscopy (cryo-EM) structure of the human spliceosome (PDB ID:5Z56) at resolution 5.1 Å [18], which corresponded to B<sup>act</sup> state, was retrieved from Protein Data Bank (PDB). The complex contains 52 proteins, three snRNAs, and synthetic pre-mRNA with a molecular weight of ca. 1.8 mega-Daltons. Due to the size of the system, we did not include all of SF3b complex's components available in the cryo-EM structure, instead five chains, namely, SF3B1, SF3B3, SF3B5, SF3B6, SF3B7, and pre-mRNA (See Figure 8) were used in MD simulations since they directly interact with SF3B1. We used SWISS-MODEL web-server [60] to model the missing parts of the SF3B1 except the first 81 residues. We did not include both SF3B2 and SF3B4 in the final complex for the following reasons: i) they did not have any direct interactions with SF3B1, ii) SF3B2 has 895 residues but only 183 of them are available, and iii) SF3B4 has 424 residues, but only 78 residues are resolved in the experimental structure. We exploited CHARMM-GUI [61] to prepare both wild type and the mutant systems. CHARMM36 force field [62] was used to model protein and RNA, whereas water molecules were represented by TIP3P water model [63]. After including water molecules, wild type system has 1089028 while the mutant one has 1089281 atoms. The thickness of water layer was set to 16 Å in each direction to prevent interaction with the periodic images of the system. The system was neutralized with 0.15 M KCl.



**Figure 8.** The depiction of the 3D structure of five domains of the SF3b complex included in this study. They are taken from the cryogenic electron microscopy (cryo-EM) structure of human spliceosome complex (PDB:5z56) [18]. Each domain is shown in a different color. SF3B1, SF3B3, SF3B5, SF3B6, SF3B7, and pre-mRNA are shown in blue, red, yellow, grey, orange, and green, respectively. The image was created using VMD [40].

#### 4.2.2. Simulation Protocol

Molecular dynamics simulations were performed by using GROMINGEN MACHINE for Chemical Simulations (GROMACS) [64] package. Equilibration of the system was performed by using the NVT ensemble, whereas the NPT ensemble was used in the production step to maintain constant pressure and temperature throughout the simulation. We used Nose-Hoover and Parrinello-Rahman coupling algorithms with 1 ps and 5 ps coupling times to maintain the temperature at 310 °K and 1 bar, respectively. LINCS algorithm was used to constrain the bond lengths in hydrogen atoms [65]. The Particle Mesh Ewald (PME) method was used to compute long-range electrostatic interactions [66]. For van der Waals and short-range electrostatic interactions, we set the cut off value to (12 Å). The time step for integration was set to 2 femtoseconds. MD simulations were performed for 2 microseconds in total. Three replicates, each of which started with a different initial velocity distribution, were used for both wild type and the mutant system to check the reliability of the results. The two replicates were run for 500 nanoseconds (ns), whereas the last one was run for 100 ns. For the analyses, we used two trajectories each of which was around 500 ns.

#### 4.2.3. Simulation Analysis Tools

We used GROMACS package for MD trajectories analysis as it provides many tools and GRaphing and Advanced Computation and Exploration of data (Grace) tool for data plotting.

### Root Mean Square Deviation (RMSD)

Root mean square deviation is the most commonly used similarity measurement tool [67] which is given by the following equation:

$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^n d_i^2}, \quad (1)$$

where  $n$  represents pairs of equivalent atoms and  $d_i$  is the distance between the two atoms in the  $i$ th pair. Here, the RMSD is calculated by using the backbone atoms of the selected groups. Also, the initial structure in corresponding trajectories was used as the Reference [68].

### Calculation of Number of Contacts

The number of contacts between pre-determined groups is computed by using “gmX mindist” module implemented in GROMACS. We used 0.3 nm as the cut-off value to calculate the number of contacts between the mutated residue and pre-mRNA while 3 nm was used to calculate the number of contacts between side chain residues of p14 (residues 20–100) and pre-mRNA as p14 is defined in previous studies as a marker of branch point [12,13].

### Principal Component Analysis (PCA)

To determine the dominant motion of a structure, principal component analysis is computed by obtaining the covariance matrix of  $C\alpha$  atoms with respect to the reference structure. Covariance matrix is computed using “gmX covar” module in GROMACS.

$$C_{ij} = \langle M_{ij} \Delta r_i \Delta r_j \rangle, \quad (2)$$

where  $C_{ij}$  corresponds to covariance matrix of  $C\alpha$  atoms  $i$  and  $j$ .  $M_{ij} \Delta r_i \Delta r_j$  refers to positional change from time-averaged structure for each coordinate of all  $C\alpha$  atoms  $i$  and  $j$ . Diagonalization of covariance matrix is performed using “gmX covar” module in GROMACS. It results in a set of eigenvalues  $\delta^2$  and their corresponding eigenvectors ( $v$ ).

$$Cv = \delta^2 v. \quad (3)$$

We used “gmX ana eig” module to analyze the first two eigenvectors that represent the directions and relative magnitudes [69] of more than 50% of the dominant motion of the system. Additionally, we used “rmsf” module to plot root mean square fluctuation (RMSF) per atom of the first two eigenvectors. Therefore, we can identify a particular residue (or a set of residues) that leads to fluctuation of the overall chain. RMSF is given as RMSD from the average over time calculated per each residue and with respect to the initial frame [64]. RMSF is expressed as

$$RMSF = \sqrt{\frac{1}{N} \sum_{n=1}^N (X_i(n) - \bar{X}_i)}, \quad (4)$$

where  $N$  corresponds to the number of frames obtained from simulation,  $X_i(n)$  corresponds to the coordinates of backbone atom of  $X_i$  and  $\bar{X}_i$  is given as the average coordinate obtained from simulation.

### Dynamic Cross Correlation Map (DCCM)

Dynamic cross correlation (DCC) among  $C\alpha$  of SF3B1 residues is computed using Bio3d package for biological structural analysis written in R programming language [70]. Bio3d library contains major functions needed for structural bioinformatics. For this study, we used “dcm.xyz” and “plot.dcm”

modules to generate dynamical cross correlation matrices and to plot them, respectively. DCC is expressed as

$$DCC(i, j) = \frac{\langle \Delta r_i(t) \cdot \Delta r_j(t) \rangle_t}{\sqrt{\langle \|\Delta r_i(t)\|^2 \rangle_t} \sqrt{\langle \|\Delta r_j(t)\|^2 \rangle_t}}, \quad (5)$$

where  $r_i(t)$  and  $r_j(t)$  correspond to the coordinates of the  $i$ th and  $j$ th atoms as a function of time  $t$ ,  $\langle \cdot \rangle_t$  indicates the time ensemble average,  $\Delta r_i(t) = r_i(t) - \langle r_i(t) \rangle_t$  and  $\Delta r_j(t) = r_j(t) - \langle r_j(t) \rangle_t$  [71–73].

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2218-273X/10/5/680/s1>, Figure S1: Depiction of the components of the constructed gene network that contains genes participating in mRNA splicing mechanism.

**Author Contributions:** Conceptualization, A.S., B.E.S., M.K.O. and O.S.; methodology, A.S., B.E.S., M.K.O. and O.S.; formal analysis, A.S.; data curation, A.S.; writing—original draft preparation, A.S.; writing—review and editing, A.S., B.E.S., M.K.O. and O.S.; supervision, B.E.S., M.K.O. and O.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding

**Acknowledgments:** We thank Istanbul Medipol University for offering the High Performance Computing system to perform classical molecular dynamics simulations. We also thank Metehan Ilter for his help through MD trajectory analysis.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Hortobagyi, G.N.; de la Garza Salazar, J.; Pritchard, K.; Amadori, D.; Haidinger, R.; Hudis, C.A.; Khaled, H.; Liu, M.C.; Martin, M.; Namer, M.; et al. The global breast cancer burden: Variations in epidemiology and survival. *Clin. Breast Cancer* **2005**, *6*, 391–401. [CrossRef] [PubMed]
- Alsafadi, S.; Houy, A.; Battistella, A.; Popova, T.; Wassef, M.; Henry, E.; Tirode, F.; Constantinou, A.; Piperno-Neumann, S.; Roman-Roman, S.; et al. Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage. *Nat. Commun.* **2016**, *7*, 10615. [CrossRef] [PubMed]
- Darman, R.B.; Seiler, M.; Agrawal, A.A.; Lim, K.H.; Peng, S.; Aird, D.; Bailey, S.L.; Bhavsar, E.B.; Chan, B.; Colla, S.; et al. Cancer-associated SF3B1 hotspot mutations induce cryptic 3' splice site selection through use of a different branch point. *Cell Rep.* **2015**, *13*, 1033–1045. [CrossRef] [PubMed]
- DeBoever, C.; Ghia, E.M.; Shepard, P.J.; Rassenti, L.; Barrett, C.L.; Jepsen, K.; Jamieson, C.H.; Carson, D.; Kipps, T.J.; Frazer, K.A. Transcriptome sequencing reveals potential mechanism of cryptic 3' splice site selection in SF3B1-mutated cancers. *PLoS Comput. Biol.* **2015**, *11*, e1004105. [CrossRef]
- Venables, J.P. Unbalanced alternative splicing and its significance in cancer. *Bioessays* **2006**, *28*, 378–386. [CrossRef]
- Ghigna, C.; Valacca, C.; Biamonti, G. Alternative splicing and tumor progression. *Curr. Genom.* **2008**, *9*, 556–570. [CrossRef]
- Fabregat, A.; Sidiropoulos, K.; Viteri, G.; Forner, O.; Marin-Garcia, P.; Arnau, V.; D'Eustachio, P.; Stein, L.; Hermjakob, H. Reactome pathway analysis: A high-performance in-memory approach. *BMC Bioinform.* **2017**, *18*, 142. [CrossRef]
- Berget, S.M.; Moore, C.; Sharp, P.A. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 3171–3175. [CrossRef]
- Chow, L.T.; Gelinis, R.E.; Broker, T.R.; Roberts, R.J. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* **1977**, *12*, 1–8. [CrossRef]
- Wahl, M.C.; Will, C.L.; Lührmann, R. The spliceosome: Design principles of a dynamic RNP machine. *Cell* **2009**, *136*, 701–718. [CrossRef]
- Gozani, O.; Feld, R.; Reed, R. Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. *Genes Dev.* **1996**, *10*, 233–243. [CrossRef] [PubMed]

12. Will, C.L.; Schneider, C.; MacMillan, A.M.; Katopodis, N.F.; Neubauer, G.; Wilm, M.; Lührmann, R.; Query, C.C. A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. *EMBO J.* **2001**, *20*, 4536–4546. [[CrossRef](#)] [[PubMed](#)]
13. Query, C.C.; McCaw, P.S.; Sharp, P.A. A minimal spliceosomal complex A recognizes the branch site and polypyrimidine tract. *Mol. Cell. Biol.* **1997**, *17*, 2944–2953. [[CrossRef](#)] [[PubMed](#)]
14. Gozani, O.; Potashkin, J.; Reed, R. A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. *Mol. Cell. Biol.* **1998**, *18*, 4752–4760. [[CrossRef](#)]
15. Cretu, C.; Schmitzová, J.; Ponce-Salvatierra, A.; Dybkov, O.; De Laurentiis, E.I.; Sharma, K.; Will, C.L.; Urlaub, H.; Lührmann, R.; Pena, V. Molecular Architecture of SF3b and Structural Consequences of Its Cancer-Related Mutations. *Mol. Cell* **2016**, *64*, 307–319. [[CrossRef](#)]
16. Cass, D.M.; Berglund, J.A. The SF3b155 N-terminal domain is a scaffold important for splicing. *Biochemistry* **2006**, *45*, 10092–10101. [[CrossRef](#)]
17. Wang, C.; Chua, K.; Seghezzi, W.; Lees, E.; Gozani, O.; Reed, R. Phosphorylation of spliceosomal protein SAP 155 coupled with splicing catalysis. *Genes Dev.* **1998**, *12*, 1409–1414. [[CrossRef](#)]
18. Zhang, X.; Yan, C.; Zhan, X.; Li, L.; Lei, J.; Shi, Y. Structure of the human activated spliceosome in three conformational states. *Cell Res.* **2018**, *28*, 307. [[CrossRef](#)]
19. Schellenberg, M.J.; Edwards, R.A.; Ritchie, D.B.; Kent, O.A.; Golas, M.M.; Stark, H.; Lührmann, R.; Glover, J.M.; MacMillan, A.M. Crystal structure of a core spliceosomal protein interface. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 1266–1271. [[CrossRef](#)]
20. Spadaccini, R.; Reidt, U.; Dybkov, O.; Will, C.; Frank, R.; Stier, G.; Corsini, L.; Wahl, M.C.; Lührmann, R.; Sattler, M. Biochemical and NMR analyses of an SF3b155–p14–U2AF-RNA interaction network involved in branch point definition during pre-mRNA splicing. *RNA* **2006**, *12*, 410–425. [[CrossRef](#)]
21. Andrade, M.A.; Perez-Iratxeta, C.; Ponting, C.P. Protein repeats: Structures, functions, and evolution. *J. Struct. Biol.* **2001**, *134*, 117–131. [[CrossRef](#)] [[PubMed](#)]
22. Visconte, V.; Avishai, N.; Mahfouz, R.; Tabarrokhi, A.; Cowen, J.; Sharghi-Moshtaghin, R.; Hitomi, M.; Rogers, H.; Hasrouni, E.; Phillips, J.; et al. Distinct iron architecture in SF3B1-mutant myelodysplastic syndrome patients is linked to an SLC25A37 splice variant with a retained intron. *Leukemia* **2015**, *29*, 188–195. [[CrossRef](#)] [[PubMed](#)]
23. Maguire, S.L.; Leonidou, A.; Wai, P.; Marchiò, C.; Ng, C.K.; Sapino, A.; Salomon, A.V.; Reis-Filho, J.S.; Weigelt, B.; Natrajan, R.C. SF3B1 mutations constitute a novel therapeutic target in breast cancer. *J. Pathol.* **2015**, *235*, 571–580. [[CrossRef](#)] [[PubMed](#)]
24. Fackenthal, J.D.; Godley, L.A. Aberrant RNA splicing and its functional consequences in cancer Cells. *Dis. Model. Mech.* **2008**, *1*, 37–42. [[CrossRef](#)]
25. Dolatshad, H.; Pellagatti, A.; Liberante, F.G.; Llorian, M.; Repapi, E.; Steeples, V.; Roy, S.; Scifo, L.; Armstrong, R.N.; Shaw, J.; et al. Cryptic splicing events in the iron transporter ABCB7 and other key target genes in SF3B1-mutant myelodysplastic syndromes. *Leukemia* **2016**, *30*, 2322. [[CrossRef](#)]
26. Borišek, J.; Saltalamacchia, A.; Galli, A.; Palermo, G.; Molteni, E.; Malcovati, L.; Magistrato, A. Disclosing the Impact of Carcinogenic SF3b Mutations on Pre-mRNA Recognition Via All-Atom Simulations. *Biomolecules* **2019**, *9*, 633. [[CrossRef](#)]
27. Borišek, J.; Saltalamacchia, A.; Spinello, A.; Magistrato, A. Exploiting Cryo-EM Structural Information and All-Atom Simulations to Decrypt the Molecular Mechanism of Splicing Modulators. *J. Chem. Inf. Model.* **2019**. [[CrossRef](#)]
28. Read, A.; Natrajan, R. Splicing dysregulation as a driver of breast cancer. *Endocr.-Relat. Cancer* **2018**, *25*, R467–R478. [[CrossRef](#)]
29. MacMillan, A.M.; Query, C.C.; Allerson, C.R.; Chen, S.; Verdine, G.L.; Sharp, P.A. Dynamic association of proteins with the pre-mRNA branch region. *Genes Dev.* **1994**, *8*, 3008–3020. [[CrossRef](#)]
30. Arenas, J.E.; Abelson, J.N. Prp43: An RNA helicase-like factor involved in spliceosome disassembly. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 11798–11802. [[CrossRef](#)]
31. Fourmann, J.B.; Schmitzová, J.; Christian, H.; Urlaub, H.; Ficner, R.; Boon, K.L.; Fabrizio, P.; Lührmann, R. Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. *Genes Dev.* **2013**, *27*, 413–428. [[CrossRef](#)] [[PubMed](#)]

32. Tsai, R.T.; Fu, R.H.; Yeh, F.L.; Tseng, C.K.; Lin, Y.C.; Huang, Y.H.; Cheng, S.C. Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2. *Genes Dev.* **2005**, *19*, 2991–3003. [[CrossRef](#)] [[PubMed](#)]
33. Murray, J.I.; Voelker, R.B.; Henscheid, K.L.; Warf, M.B.; Berglund, J.A. Identification of motifs that function in the splicing of non-canonical introns. *Genome Biol.* **2008**, *9*, R97. [[CrossRef](#)] [[PubMed](#)]
34. Ohi, M.D.; Vander Kooi, C.W.; Rosenberg, J.A.; Ren, L.; Hirsch, J.P.; Chazin, W.J.; Walz, T.; Gould, K.L. Structural and functional analysis of essential pre-mRNA splicing factor Prp19p. *Mol. Cell. Biol.* **2005**, *25*, 451–460. [[CrossRef](#)]
35. Hogg, R.; McGrail, J.C.; O’Keefe, R.T. The function of the NineTeen Complex NTC in regulating spliceosome conformations and fidelity during pre-mRNA splicing. *Biochem. Soc. Trans.* **2010**, *38*, 1110–1115. [[CrossRef](#)]
36. Das, S.; Krainer, A.R. Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer. *Mol. Cancer Res.* **2014**, *12*, 1195–1204. [[CrossRef](#)]
37. Huang, Y.; Gattoni, R.; Stévenin, J.; Steitz, J.A. SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Mol. Cell* **2003**, *11*, 837–843. [[CrossRef](#)]
38. Chin, K.; Sharp, K.A.; Honig, B.; Pyle, A.M. Calculating the electrostatic properties of RNA provides new insights into molecular interactions and function. *Nat. Struct. Mol. Biol.* **1999**, *6*, 1055.
39. Query, C.C.; Strobel, S.A.; Sharp, P.A. Three recognition events at the branch-site adenine. *EMBO J.* **1996**, *15*, 1392–1402. [[CrossRef](#)]
40. Humphrey, W.; Dalke, A.; Schulten, K. VMD—Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [[CrossRef](#)]
41. Kesarwani, A.K.; Ramirez, O.; Gupta, A.K.; Yang, X.; Murthy, T.; Minella, A.C.; Pillai, M.M. Cancer-associated SF3B1 mutants recognize otherwise inaccessible cryptic 3’ splice sites within RNA secondary structures. *Oncogene* **2017**, *36*, 1123–1133. [[CrossRef](#)] [[PubMed](#)]
42. Carrocci, T.J.; Zoerner, D.M.; Paulson, J.C.; Hoskins, A.A. SF3b1 mutations associated with myelodysplastic syndromes alter the fidelity of branchsite selection in yeast. *Nucleic Acids Res.* **2017**, *45*, 4837–4852. [[CrossRef](#)] [[PubMed](#)]
43. Zhang, J.; Ali, A.M.; Lieu, Y.K.; Liu, Z.; Gao, J.; Rabadan, R.; Raza, A.; Mukherjee, S.; Manley, J.L. Disease-causing mutations in SF3B1 alter splicing by disrupting interaction with SUGP1. *Mol. Cell* **2019**, *76*, 82–95. [[CrossRef](#)] [[PubMed](#)]
44. Boudrez, A.; Beullens, M.; Waelkens, E.; Stalmans, W.; Bollen, M. Phosphorylation-dependent interaction between the splicing factors SAP155 and NIPP1. *J. Biol. Chem.* **2002**, *277*, 31834–31841. [[CrossRef](#)] [[PubMed](#)]
45. Liberante, F.G.; Lappin, K.; Barros, E.M.; Vohhodina, J.; Grebien, F.; Savage, K.I.; Mills, K.I. Altered splicing and cytoplasmic levels of tRNA synthetases in SF3B1-mutant myelodysplastic syndromes as a therapeutic vulnerability. *Sci. Rep.* **2019**, *9*, 1–13. [[CrossRef](#)] [[PubMed](#)]
46. Seiler, M.; Yoshimi, A.; Darman, R.; Chan, B.; Keaney, G.; Thomas, M.; Agrawal, A.A.; Caleb, B.; Csibi, A.; Sean, E.; et al. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. *Nat. Med.* **2018**, *24*, 497. [[CrossRef](#)]
47. Obeng, E.A.; Chappell, R.J.; Seiler, M.; Chen, M.C.; Campagna, D.R.; Schmidt, P.J.; Schneider, R.K.; Lord, A.M.; Wang, L.; Gambe, R.G.; et al. Physiologic expression of Sf3b1(K700E) causes impaired erythropoiesis, aberrant splicing, and sensitivity to therapeutic spliceosome modulation. *Cancer Cell* **2016**, *30*, 404–417. [[CrossRef](#)]
48. Di, C.; Zhang, Q.; Chen, Y.; Wang, Y.; Zhang, X.; Liu, Y.; Sun, C.; Zhang, H.; Hoheisel, J.D. Function, clinical application, and strategies of Pre-mRNA splicing in cancer. *Cell Death Differ.* **2019**, *26*, 1181–1194. [[CrossRef](#)]
49. Desmedt, C.; Haibe-Kains, B.; Wirapati, P.; Buyse, M.; Larsimont, D.; Bontempi, G.; Delorenzi, M.; Piccart, M.; Sotiriou, C. Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. *Clin. Cancer Res.* **2008**, *14*, 5158–5165. [[CrossRef](#)]
50. Bonifaci, N.; Berenguer, A.; Díez, J.; Reina, O.; Medina, I.; Dopazo, J.; Moreno, V.; Pujana, M.A. Biological processes, properties and molecular wiring diagrams of candidate low-penetrance breast cancer susceptibility genes. *BMC Med. Genom.* **2008**, *1*, 62. [[CrossRef](#)]
51. Wang, B.D.; Lee, N.H. Aberrant RNA splicing in cancer and drug resistance. *Cancers* **2018**, *10*, 458. [[CrossRef](#)] [[PubMed](#)]

52. Croft, D.; Mundo, A.F.; Haw, R.; Milacic, M.; Weiser, J.; Wu, G.; Caudy, M.; Garapati, P.; Gillespie, M.; Kamdar, M.R.; et al. The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **2013**, *42*, D472–D477. [[CrossRef](#)] [[PubMed](#)]
53. Szklarczyk, D.; Franceschini, A.; Wyder, S.; Forslund, K.; Heller, D.; Huerta-Cepas, J.; Simonovic, M.; Roth, A.; Santos, A.; Tsafou, K.P.; et al. STRING v10: Protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **2014**, *43*, D447–D452. [[CrossRef](#)] [[PubMed](#)]
54. Assenov, Y.; Ramírez, F.; Schelhorn, S.E.; Lengauer, T.; Albrecht, M. Computing topological parameters of biological networks. *Bioinformatics* **2007**, *24*, 282–284. [[CrossRef](#)] [[PubMed](#)]
55. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)] [[PubMed](#)]
56. Tate, J.G.; Bamford, S.; Jubb, H.C.; Sondka, Z.; Beare, D.M.; Bindal, N.; Boutselakis, H.; Cole, C.G.; Creatore, C.; Dawson, E.; et al. COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **2018**, *47*, D941–D947. [[CrossRef](#)]
57. Malhotra, G.K.; Zhao, X.; Band, H.; Band, V. Histological, molecular and functional subtypes of breast cancers. *Cancer Biol. Ther.* **2010**, *10*, 955–960. [[CrossRef](#)]
58. Makki, J. Diversity of breast carcinoma: Histological subtypes and clinical relevance. *Clin. Med. Insights Pathol.* **2015**, *8*, S31563. [[CrossRef](#)]
59. Seiler, M.; Peng, S.; Agrawal, A.A.; Palacino, J.; Teng, T.; Zhu, P.; Smith, P.G.; Caesar-Johnson, S.J.; Demchok, J.A.; Felau, I.; et al. Somatic mutational landscape of splicing factor genes and their functional consequences across 33 cancer types. *Cell Rep.* **2018**, *23*, 282–296. [[CrossRef](#)]
60. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. [[CrossRef](#)]
61. Lee, J.; Cheng, X.; Swails, J.M.; Yeom, M.S.; Eastman, P.K.; Lemkul, J.A.; Wei, S.; Buckner, J.; Jeong, J.C.; Qi, Y.; et al. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2015**, *12*, 405–413. [[CrossRef](#)]
62. Klauda, J.B.; Venable, R.M.; Freites, J.A.; O'Connor, J.W.; Tobias, D.J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A.D., Jr.; Pastor, R.W. Update of the CHARMM all-atom additive force field for lipids: Validation on six lipid types. *J. Phys. Chem. B* **2010**, *114*, 7830–7843. [[CrossRef](#)]
63. Mark, P.; Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* **2001**, *105*, 9954–9960. [[CrossRef](#)]
64. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*, 19–25. [[CrossRef](#)]
65. Wang, L.; Yuan, Y.; Chen, X.; Chen, J.; Guo, Y.; Li, M.; Li, C.; Pu, X. Probing the cooperative mechanism of the  $\mu$ - $\delta$  opioid receptor heterodimer by multiscale simulation. *Phys. Chem. Chem. Phys.* **2018**, *20*, 29969–29982. [[CrossRef](#)] [[PubMed](#)]
66. Yuan, S.; Filipek, S.; Palczewski, K.; Vogel, H. Activation of G-protein-coupled receptors correlates with the formation of a continuous internal water pathway. *Nat. Commun.* **2014**, *5*, 4733. [[CrossRef](#)]
67. Kufareva, I.; Abagyan, R. Methods of protein structure comparison. In *Homology Modeling*; Springer: Berlin, Germany, 2011; pp. 231–257.
68. Aier, I.; Varadwaj, P.K.; Raj, U. Structural insights into conformational stability of both wild-type and mutant EZH2 receptor. *Sci. Rep.* **2016**, *6*, 34984. [[CrossRef](#)]
69. Yang, L.; Song, G.; Carriquiry, A.; Jernigan, R.L. Close correspondence between the motions from principal component analysis of multiple HIV-1 protease structures and elastic network modes. *Structure* **2008**, *16*, 321–330. [[CrossRef](#)]
70. Grant, B.J.; Rodrigues, A.P.; ElSawy, K.M.; McCammon, J.A.; Caves, L.S. Bio3d: An R package for the comparative analysis of protein structures. *Bioinformatics* **2006**, *22*, 2695–2696. [[CrossRef](#)]
71. Mishra, S.K.; Jernigan, R.L. Protein dynamic communities from elastic network models align closely to the communities defined by molecular dynamics. *PLoS ONE* **2018**, *13*, e0199225. [[CrossRef](#)]



72. Kasahara, K.; Fukuda, I.; Nakamura, H. A novel approach of dynamic cross correlation analysis on molecular dynamics simulations and its application to Ets1 dimer–DNA complex. *PLoS ONE* **2014**, *9*, e112419. [[CrossRef](#)] [[PubMed](#)]
73. McCammon, J. Protein dynamics. *Rep. Prog. Phys.* **1984**, *47*, 1. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).