Atomic-Level Mechanism of Pre-mRNA Splicing in Health and Disease

Jure Borišek, Lorenzo Casalino, Andrea Saltalamacchia, Suzanne G. Mays, Luca Malcovati, and Alessandra Magistrato*

ACCESS | Metrics & More | Article Recommendations

CONSPECTUS: Intron removal from premature-mRNA (pre-mRNA splicing) is an essential part of gene expression and regulation that is required for the production of mature, protein-coding mRNA. The spliceosome (SPL), a majestic machine composed of five small nuclear RNAs and hundreds of proteins, behaves as an eminent transcriptome tailor, efficiently performing splicing as a protein-directed metallo-ribozyme. To select and excise long and diverse intronic sequences with single-nucleotide precision, the SPL undergoes a continuous compositional and conformational remodeling, forming eight distinct complexes throughout each splicing cycle. Splicing fidelity is of paramount importance to preserve the integrity of the proteome. Mutations in splicing factors can severely compromise the accuracy of this machinery, leading to

aberrant splicing and altered gene expression. Decades of biochemical and genetic studies have provided insights into the SPL's composition and function, but its complexity and plasticity have prevented an in-depth mechanistic understanding. Single-particle cryogenic electron microscopy techniques have ushered in a new era for comprehending eukaryotic gene regulation, providing several near-atomic resolution structures of the SPL from yeast and humans. Nevertheless, these structures represent isolated snapshots of the splicing process and are insufficient to exhaustively assess the function of each SPL component and to unravel particular facets of the splicing mechanism in a dynamic environment.
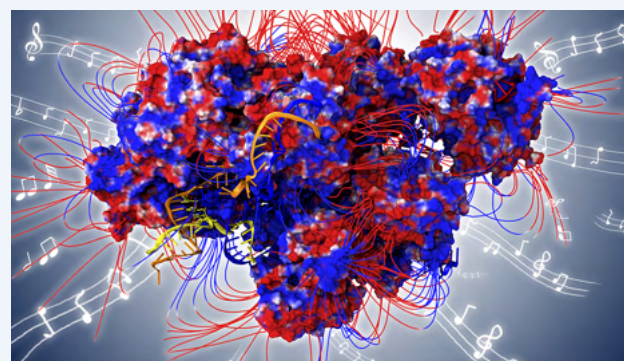
In this Account, building upon our contributions in this field, we discuss the role of biomolecular simulations in uncovering the mechanistic intricacies of eukaryotic splicing in health and disease. Specifically, we showcase previous applications to illustrate the role of atomic-level simulations in elucidating the function of specific proteins involved in the architectural reorganization of the SPL along the splicing cycle. Moreover, molecular dynamics applications have uniquely contributed to decrypting the channels of communication required for critical functional transitions of the SPL assemblies. They have also shed light on the role of carcinogenic mutations in the faithful selection of key intronic regions and the molecular mechanism of splicing modulators. Additionally, we emphasize the role of quantum-classical molecular dynamics in unraveling the chemical details of pre-mRNA cleavage in the SPL and in its evolutionary ancestors, group II intron ribozymes. We discuss methodological pitfalls of multiscale calculations currently used to dissect the splicing mechanism, presenting future challenges in this field. The results highlight how atomic-level simulations can enrich the interpretation of experimental results. We envision that the synergy between computational and experimental approaches will aid in developing innovative therapeutic strategies and revolutionary gene modulation tools to fight the over 200 human diseases associated with splicing misregulation, including cancer and neurodegeneration.

■ KEY REFERENCES

- Casalino, L.; Palermo, G.; Spinello, A.; Rothlisberger, U.; Magistrato, A. All-atom simulations disentangle the functional dynamics underlying gene maturation in the intron lariat spliceosome. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 6584−6589.[1] *This was a pioneering work for the field as it is the very first molecular dynamics simulation of the spliceosome that elucidates the atomic-level perspective of the late stage of the splicing cycle.*

- Saltalamacchia, A.; Casalino, L.; Borišek, J.; Batista, V. S.; Rivalta, I.; Magistrato, A. Decrypting the Information Exchange Pathways across the Spliceosome Machinery. *J.*
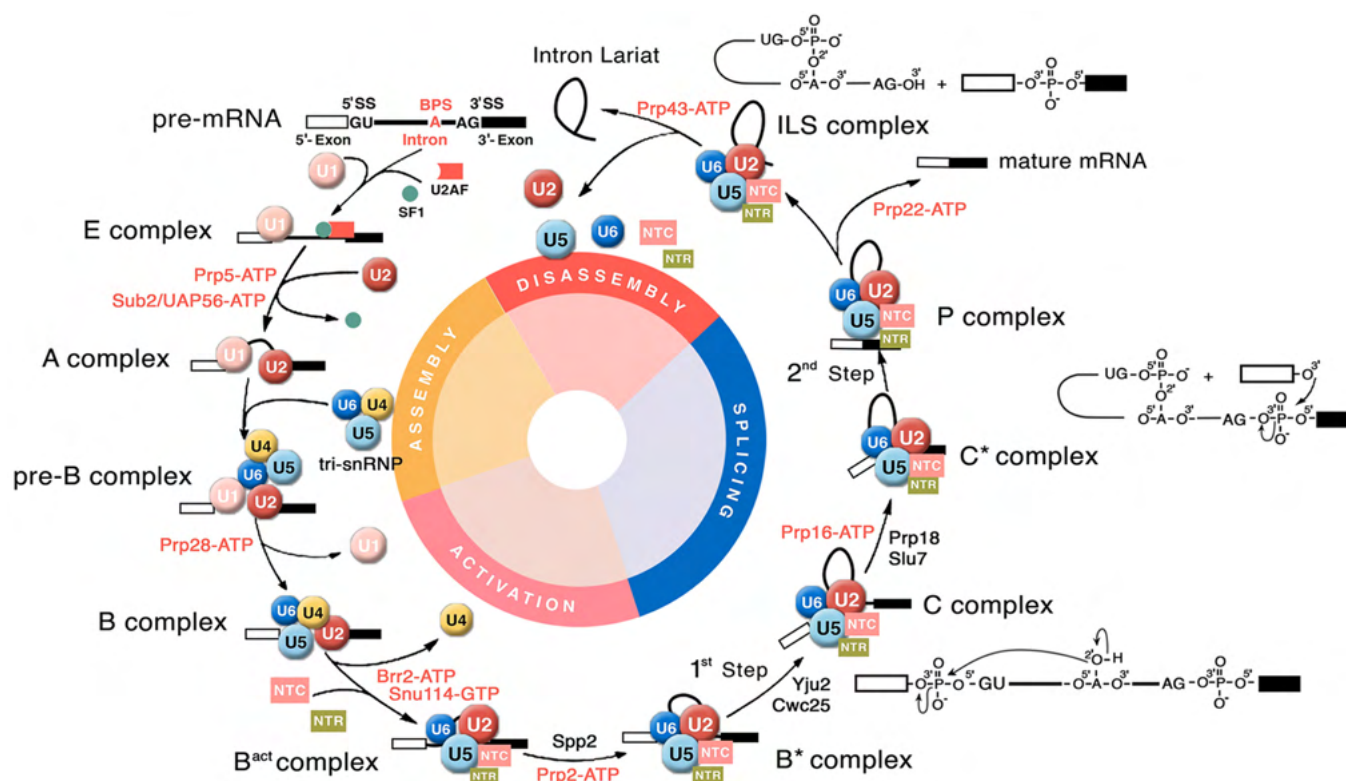
Figure 1. Splicing cycle and a schematic of the splicing reaction mechanism. The splicing factors U1, U2, U4, U5, and U6 snRNPs bind the conserved pre-mRNA splice sites (SS) and mediate the assembly of the spliceosome (E, A, pre-B, B). Subsequently, an activation phase is required to rearrange the spliceosome and move the SS into a catalytically active position (B^{act}). The catalytic reactions (B* and C*) lead to the ligation of the exons (white and black), whereas the intron-lariat is excised. Finally, the intron-lariat and the mature mRNA are released, and the spliceosome is disassembled. Helicase proteins, functional for structural rearrangements along the cycle, are shown in red. The branch point adenosine (BPA) in pre-mRNA is colored in red on the mRNA filament.

Am. Chem. Soc. **2020**, *142*, 8403−8411.[2] *The elucidation of the complexity of the spliceosome C complex through all-atom molecular dynamics simulations and community network analysis enables the identification of the key channels of information transfer over long distances that separate important protein components.*

- Borišek, J.; Magistrato, A. All-Atom Simulations Decrypt the Molecular Terms of RNA Catalysis in the Exon-Ligation Step of the Spliceosome. *ACS Catal.* **2020**, *10*, 5328−5334.[3] *Hybrid quantum−classical (QM/MM) molecular dynamics simulations elucidate that the exon-ligation step in the spliceosome entails an associative two-Mg^{2+}-ion mechanism, catalyzed exclusively by RNA, where scissile phosphate aids a proton transfer from the nucleophile to the leaving group.*

- 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.[4] *Molecular dynamics simulations of the spliceosome B^{act} complex show that selected carcinogenic isoforms undermine intron binding and/or alter the functional dynamics of Hsh155/SF3B1 protein only when binding nonconsensus branch point sites.*

INTRODUCTION

The coding regions of nascent gene transcripts, the exons, must be excised from the intervening noncoding regions, the introns,

to produce mature transcripts used for protein synthesis. Since the discovery of introns in 1977,[5] the inherent complexity and the sophisticated nature of precursor-mRNA (pre-mRNA) splicing has attracted the interest and stimulated the imagination of scientists working in different research areas. It is now well established that the organization of introns and exons in newly transcribed pre-mRNAs is a hallmark of the genomes of all organisms, therefore making the splicing process a pivotal step of gene expression, regulation, and diversification across all domains of life.[6]

THE SPLICEOSOME MACHINERY

Intron removal can rely on self-catalytic RNAs, as in prokaryotes and some organellar genomes, or depend on a sophisticated, multi-megadalton and exceptionally dynamic ribonucleoprotein machinery, the spliceosome (SPL), which catalyzes pre-mRNA splicing in eukaryotes.[7]

The SPL comprises approximately 150 proteins and five small nuclear RNAs (snRNAs), namely U1, U2, U4, U5, and U6 snRNAs, which form distinct small ribonucleoprotein complexes (snRNPs) through an intricate network of interactions. After assembling de novo at each splicing cycle, the SPL, fueled by the hydrolysis of ATP, undergoes a series of conformational and compositional transformations to form eight major complexes (A, B, B^{act}, B*, C, C*, P, and ILS) (Figure 1). These intermediate states shape the long and diverse pre-mRNA sequences into a conformation that facilitates splicing catalysis.[6]

The SPL ensures splicing fidelity by recognizing specific sequence signals in the intron, namely, the 5′ and 3′ splice sites
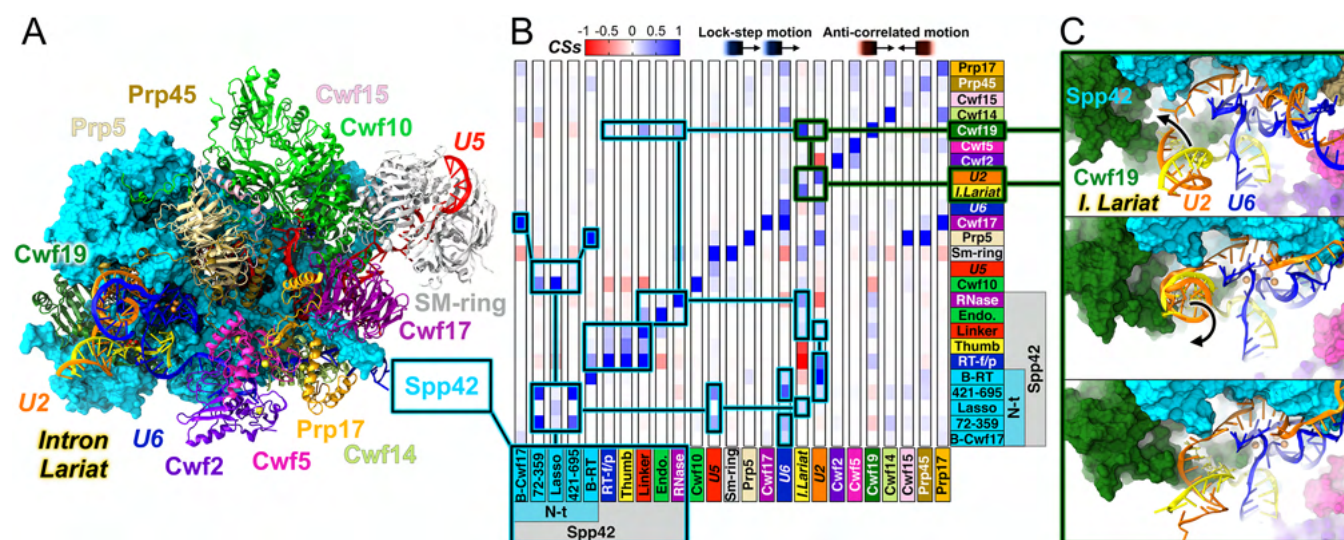
Figure 2. Cooperative motions underlying the functional dynamics in the spliceosome. Image adapted with permission from ref 1. Copyright (2018) National Academy of Science USA. (A) Model of the intron-lariat spliceosome (ILS) complex based on the *S. pombe* ILS cryo-EM structure (PDB entry 3JB9). (B) Histogram reporting the (per-column) normalized correlation scores (CSs) calculated for each pair of spliceosome components by accumulating all the Pearson's correlation coefficients between residues of the respective components. Spp42 domains and N-terminal motifs are distinctly considered. CS value ranges from −1 (anticorrelated motions) to +1 (lock-step motions). The most relevant correlations are highlighted using colored boxes. (C) Gradual displacement of the intron-lariat (IL)/U2 helix. IL/U2 is first engaged by Cwf19 and Spp42, which cooperatively promote its unrolling and final displacement.

(5′SS and 3′SS, respectively). The 5′SS, characterized by a conserved GU sequence at the 5′ end of the intron, is recognized by U1 snRNP during E complex assembly, whereas the 3′SS is recognized by U2 snRNP during A-complex formation.[8] Sequence elements involved in 3′SS selection include the branch point sequence (BPS), which base pairs with a portion of the U2 snRNA, a stretch of polypyrimidines (the polypyrimidine tract, present in higher eukaryotes and recognized by U2AF2), and an AG dinucleotide at the very end of the intron, recognized by U2AF1.[9−11]

After identifying the intron/exon boundaries, the SPL catalyzes the chemical "snip-and-stitch" splicing process in two sequential transesterification reactions that require Mg^{2+} ions.[6] In the first (branching) step, carried out by the B* complex, a conserved bulging adenosine located within the BPS, the branch point adenosine (BPA), attacks the nucleotide at the 5′-end of the intron, forming a lasso-shaped (lariat) intron and a free 5′-exon.[8] In the second (exon-ligation) step, catalyzed by the C* complex, the free 3′ OH of the 5′-exon attacks the 5′-end of the 3′-exon, resulting in a functional protein-coding mRNA and an intron-lariat (Figure 1).[6]

The SPL's complexity, large size, and extraordinary plasticity has prevented the attainment of accurate structural information for decades. Groundbreaking advancements in detector technology and software algorithms for cryogenic electron microscopy (cryo-EM)[12] made accessible, in 2015, the first near-atomic resolution structure of the intron-lariat SPL (ILS) complex from the yeast *Schizosaccharomyces pombe*.[13] After the accomplishment of this milestone, a wealth of structural information from yeast and human became rapidly available, ushering in a new era for understanding the molecular basis of pre-mRNA splicing.[8] Computational biology approaches have been used to shed light on new structural information, illuminating fundamental mechanistic facets of this fascinating machinery. Specifically, molecular simulations have contributed to elucidating the function of proteins involved in the SPL's

protein/RNA assembly,[1,2] to understanding the mechanism of faithful recognition of critical regions in pre-mRNA transcripts,[4] and to uncovering mechanistic details of splicing reactions.[3,14] In addition, they have enabled the investigation of the impact of carcinogenic mutations on SPL dynamics[4,15] and the molecular basis of splicing modulation by small molecules as a potential treatment for diseases linked to mis-splicing events.[16,17]

In this Account, by providing an overview of the most recent applications, we illustrate how molecular simulations are contributing to the understanding of pre-mRNA splicing in health and disease.

FUNCTIONAL ROLES AND PATHWAYS OF INFORMATION TRANSFER AMONG SPLICEOSOME PROTEINS

The structures provided by cryo-EM and macromolecular crystallography represent isolated snapshots of the SPL. However, biomolecules are in constant motion and exert their function via precise dynamic transitions. By integrating Newton's equations of motion, all-atom molecular dynamics (MD) simulations can describe the motion of each atom of a biomolecular system as a function of time, providing an atomic-level moving picture of the system within the simulated time frame.[18] The forces acting on the atoms are calculated as derivatives of potentials, referred to as molecular mechanics' force field (FF), which account in a simplified, yet effective manner for bond and angle vibrations and for electrostatic and van der Waals interactions between atoms.[18]

Principal component analysis (PCA) and cross-correlation matrices can provide insights into channels of information transfer among components of a complex.[19]

For example, the presence of correlated or anticorrelated motions of residues or domains occurring over long distances can indicate communication between critical distal sites in the complex.
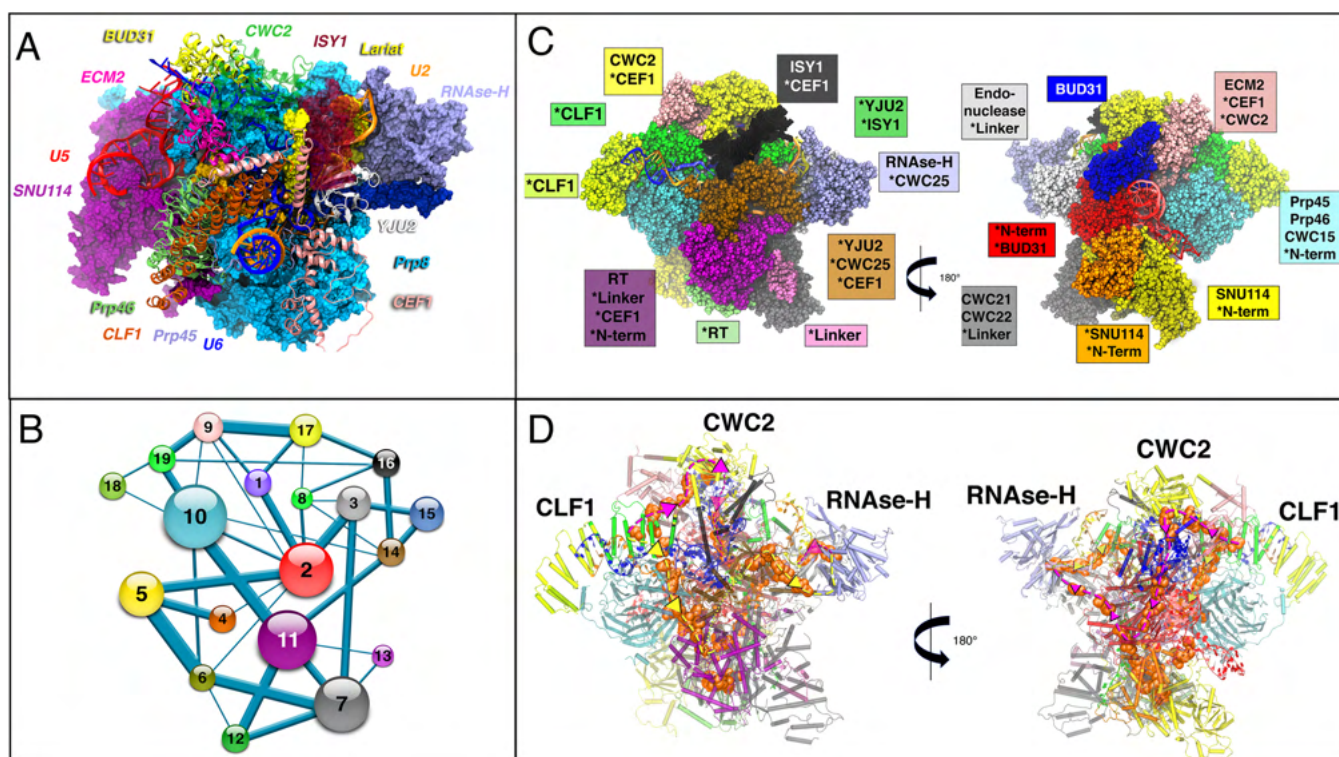
Figure 3. Community network analysis of the spliceosome. Images reproduced with permission from ref 2. Copyright 2020 American Chemical Society. (A) Model of the C complex spliceosome from the yeast *S. cerevisiae* cryo-EM structure (PDB entry: 5LJ3). (B) 2D graph of the community network where each community is represented by a circle proportional to the number of residues in the community. The connections' width accounts for the intercommunity communication flow (i.e., the totality of communication between residues belonging to the two communities). (C) 3D-structure (front and back) of the community network, with the communities depicted with the same color-code of (B). (D) Spliceosome communication routes (front and back). In orange are displayed the highest betweenness residues composing the two principal routes (path I and path II, highlighted with magenta and yellow arrows, respectively) for signal exchange between Clf1 and RNase-H and vice versa.

The first MD study of the SPL was based on the cryo-EM structure of the ILS complex from yeast *S. pombe*.[20] This computational work consisted of multireplica all-atom simulations of an explicitly solvated model of the ILS complex, approaching one million atoms.[1] In this study, MD simulations enabled the sampling of the conformational space of the system, while statistical analysis tools served to extrapolate large-scale motions of the SPL. Among the statistical instruments, calculation of cross-correlation matrices (or normalized covariance matrices), based on Pearson's coefficients, and PCA provided valuable information about the SPL's functional dynamics. Whereas the former allows for a qualitative identification of the coupled motions taking place between pairs of residues along the MD trajectory, the latter discloses the so-called "essential dynamics" of the biomolecule highlighting the motions with the largest variance through a dimensionality reduction (Figure 2).

Altogether, the analyses applied to MD simulations of the ILS complex unveiled the leading role of Spp42 (Prp8 in humans). This protein was shown to direct the SPL maturation by governing the motions of the distinct SPL components owing to its unique multipronged architecture. Among the protein components, the tight correlation coupling of Cwf19 and Spp42 suggested a cooperative function of the two proteins and the IL/U2 branch helix. Consistently, the essential dynamics revealed that Cwf19 and Spp42 promote an electrostatically driven displacement and an unrolling of the IL/U2 branch helix (Figure 2). This finding allowed for the functional annotation of Cwf19 in the IL/U2 branch helix unwinding. Besides being fully

consistent with the late stage of the splicing cycle taken on by the ILS complex, the role of Cwf19 was thereafter corroborated by a cryo-EM study of the human SPL.[1,21]

A remarkable endeavor to identify the communication channels underlying the SPL functional transition was successfully carried out in another, more recent, MD study of the SPL C complex from yeast *Saccharomyces cerevisiae*.[2] The essential dynamics unveiled a hammerlike motion, cooperatively promoted by the Clf1 protein and the RNase-H domain of Prp8, which triggers the twisting and repositioning of the IL/U2 branch helix, promoting the transition toward the C* complex. Moreover, community network analysis (CNA),[22] successfully applied to other protein/RNA machines,[23] enabled the identification of the signaling routes (i.e., the pivotal amino acids) responsible for the communication between these two critical regions of the C complex. CNA relies on a correlation-based weighted network, in which the nodes, corresponding to the Cα atoms of the amino acids, are connected by edges that reflect the amount of correlations for each pair of residues. Subdividing this representation into groups of strongly correlated amino acids (communities) and connecting them by edges that account for the strength of the communication flow between different communities, (i.e., obtained as sum of all intercommunity edge betweenness),[24] the CNA unravels the most relevant communication pathways occurring between regions involved in critical functional movements (Figure 3).

In this study, the CNA further affirmed the dominant role of Prp8 in directing the motion of the C complex. Two of the most important communication routes (Path I and Path II in Figure
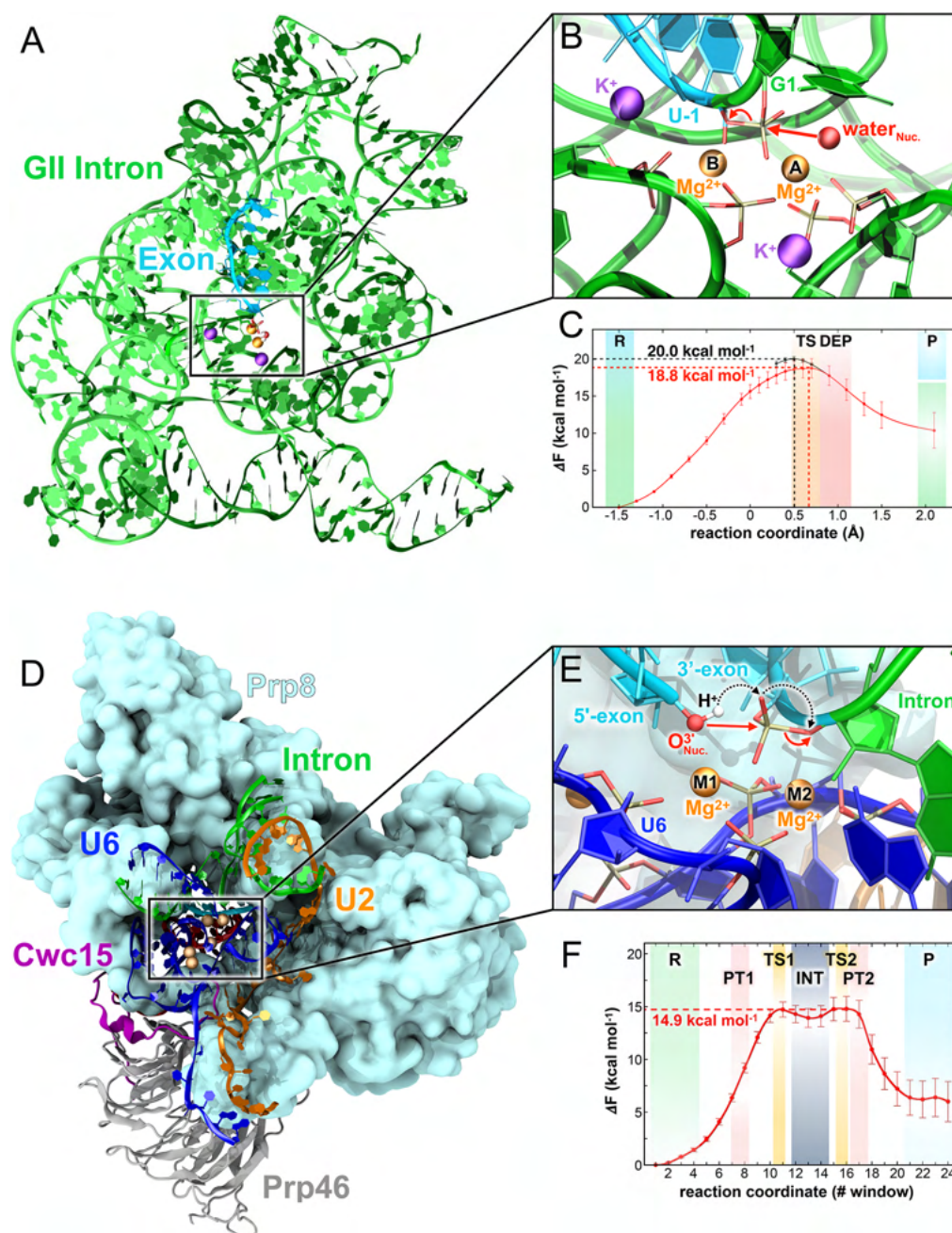
Figure 4. Group II Intron (GR2I) and spliceosome (SPL)-P complex models deployed in QM/MM simulation studies of catalytic mechanism for the first and second splicing step. Molecular representations of the GR2I and the SPL-P models are shown in panels (A) and (D), respectively. Magnified view of the catalytic site in the reacting conformation with arrows depicting the reaction path, and corresponding free energy profile are displayed in panels (B) and (C), respectively, for GR2I, and in panels (E) and (F), respectively, for SPL-P. Panels (A)−(C) are reproduced with permission from ref14. Copyright 2016 American Chemical Society.

3) connect Clf1 with the RNase-H domain (communities 19 and 15, respectively, in Figure 3) over a distance of 160 Å. Notably, the identified communication routes reveal a potential binding site along path II that could be used as a target in virtual-screening studies aimed at drugging the SPL with small-molecules potentially able to interfere with this information exchange.[2]

Taken together, these results show that Prp8, the central hub of the SPL assembly, plays a critical role in modulating the SPL functional dynamics by mediating the signal exchange between functionally critical proteins/regions.[1,2,4]

CHEMICAL MECHANISM OF SPLICING

The conformational changes occurring along the splicing cycle are instrumental to place the introns into the optimal conformation to efficiently and accurately promote pre-mRNA splicing. Akin to various other nucleic acid processing enzymes, the SPL and its self-splicing ancestors, group II intron (GR2I) ribozymes, catalyze intron removal in two transesterification steps coordinated by two Mg^{2+} ions.[25,26]

In 1993, Steitz and Steitz proposed a general two-Mg^{2+}-ion mechanism for DNA/RNA cleavage/ligation.[27,3,14,27−29] According to their proposal, the first metal activates the nucleophile by acting as a Lewis acid, the second stabilizes the

leaving group, while both ions cooperatively support the transition state structure.[27] In pre-mRNA splicing, where two phosphoryl-transfer reactions occur consecutively, the role of the two Mg^{2+} ions is reversed in the two steps. Additional metal ions (three Mg^{2+} in SPL, and two K^+ in GR2I), cradled within an intricate RNA network of interactions, stabilize the catalytic core by minimizing the electrostatic repulsion of the phosphate backbone.

With this respect, computer simulations can be exploited to gain unique, atomic-level insights into the reaction mechanism of splicing. However, they require the use of accurate and computationally expensive quantum mechanical (QM) approaches to account for the breaking and formation of chemical bonds, unlike FF-based MD simulations. Hybrid quantum mechanics/molecular mechanics (QM/MM) methods allow simulating the whole biological system in a computationally affordable manner, treating only the reacting moieties at QM level, whereas the remainder of the biomolecule and the solvent are described with a classical FF.[30]

Nevertheless, the costly computation of the QM region severely limits the time scales accessible to this approach. As a result, appropriate statistical mechanics methods including free energy or enhanced sampling approaches, are employed to observe the chemical reactions of interest, while simultaneously estimating the associated kinetic and thermodynamic properties (reaction free energies).[30]

The first computational attempt to study the mechanism of pre-mRNA splicing revolved around the investigation of the first splicing step of GR2I from *Oceanobacillus iheyensis*, which, unlike the first reaction catalyzed by the SPL, can be mediated by a water nucleophile.[14] QM/MM MD simulations, based on density functional theory in combination with thermodynamic integration, disclosed a dissociative mechanism, in which the catalytic water releases its proton to the bulk water during the nucleophilic attack on the scissile phosphate.[14] The reaction was predicted to occur in one step with a free energy cost of 18.8 ± 1.5 kcal/mol, in line with an experimental catalytic rate of 0.011 min^{-1} (corresponding to a free energy barrier of ~23 kcal/mol) accounting also for the metal-dependent folding of the ribozyme.[31] In this context, instead of priming the nucleophile for attack as proposed in the general two-metal-ion mechanism, the Mg^{2+} ion activates the scissile phosphate group (i.e., the electrophile), leading to a dissociative (metaphosphate-like) transition state, where the bond rupture is more advanced than bond formation.[14]

Following the mechanistic study on GR2I, computational efforts have been directed toward the SPL. Although none of the available cryo-EM structures has trapped the catalytic site configuration completely suitable for studying the catalytic mechanism, a substantial computational remodeling of the Mg^{2+} ion positions and of the pre-mRNA reactant allowed the investigation of the first step of splicing catalyzed by the SPL, leading to the formation of the intron-lariat.[32] The calculations were based on a SPL structure from the yeast *S. cerevisiae* trapping the C complex immediately after the branching step. As observed for GR2I, this study confirmed a dissociative pathway for the intron-lariat formation, where the nucleophilic proton is transferred from the $O^{2'}$ of the BPA to the $O^{3'}$ atom at the 3′-end of the 5′-exon through the scissile phosphate, after overcoming a free energy barrier of 21.5 kcal/mol.[32]

Another step forward in understanding the catalytic mechanism of SPL-directed pre-mRNA splicing was done through the study of the second transesterification reaction leading to exon-ligation,[3] which was based on a cryo-EM structure of the P complex from *S. cerevisiae*.[33] This structure shared high similarity with that of a G2IR trapped before exon-ligation,[33] showing a favorable catalytic site in terms of Mg^{2+} ions positions and their coordination with phosphates of the catalytic core.[3,33] QM/MM MD simulations, encompassing a QM/MM system of ~400 000 atoms, disclosed that the second step of catalysis (exon-ligation) takes place via an associative two-Mg^{2+}-ion mechanism, featuring a scissile-phosphate-mediated proton transfer from the nucleophile to the leaving group. The mechanism followed a stepwise pathway via the formation of a metastable phosphorane-like intermediate.[3] Therefore, whereas the first splicing step, in both GR2I and SPL, follows a dissociative mechanism, the second step agrees with the general associative two-metal-ion mechanism proposed by Steiz and Steiz.[27] Remarkably, in both SPL studies, only the RNA strands directly promote catalysis, albeit with critical support from Mg^{2+} ions to activate the nucleophile (in SPL) or the scissile phosphate (in GR2I), thus stabilizing the transition state and the leaving group. This endorses the general view of the SPL as a protein-directed metallo-ribozyme (Figure 4).

ALTERNATIVE AND ABERRANT SPLICING AND ITS IMPLICATION FOR SPLICEOSOME MUTANT CANCERS

The crucial importance of splicing fidelity is noted by the fact that up to one-third of all disease-causing mutations are associated with splicing defects.[34] To ensure faithful splicing, the SPL must identify exons interspersed between long introns containing degenerate splice site sequences. The situation is even more complex considering that pre-mRNAs can undergo alternative splicing patterns, whereby some exons can be skipped or some introns can be retained.[35] This process, named alternative splicing (AS), is estimated to occur in 95% of the pre-RNA multiexon transcripts in eukaryotes.[36] AS contributes to the diversification of the eukaryotic proteome by enabling production of multiple proteins from a single pre-mRNA substrate.

Deregulation of AS is strongly associated with distinct pathological states including several cancer types, neurodegeneration, and retinitis pigmentosa.[37] Mutations to core components of the SPL, splicing factors, or pre-mRNA can lead to aberrant splicing, often eliciting pathologic consequences. Several disease-associated mutations reside within regulatory sequences for splicing or within genes encoding for critical splicing factors. As an example, mutations of U1 snRNA influence the selection/recognition of 5′SS,[38] while mutations in the splicing factors U2AF1, ZRSR2, SRSF2, and SF3B1 alter the recognition of the 3′SS, affecting AS of particular substrates.[39] Mutations in the SPL and splicing factors are recurrent in myelodysplastic syndromes, chronic lymphocytic leukemia, and in solid tumors, including breast cancer and uveal melanoma.[40−42]

The splicing factor SF3B1, part of the SF3b complex in the U2 snRNP, is the most frequently mutated component of the SPL in cancers.[40] SF3B1 is involved in 3′SS recognition and stabilizes the interaction between the BPS and the U2 snRNA. SF3B1 mutations have been suggested to induce the activation of cryptic (erroneous) 3′SS.[10,17,43−45] Introduction of mutations associated with cancer into the yeast orthologue of SF3B1, Hsh15, also causes aberrant splicing, altering the selection of 3′SS containing certain nonconsensus BPS without affecting those exhibiting consensus BPS.[10] An MD simulation study of
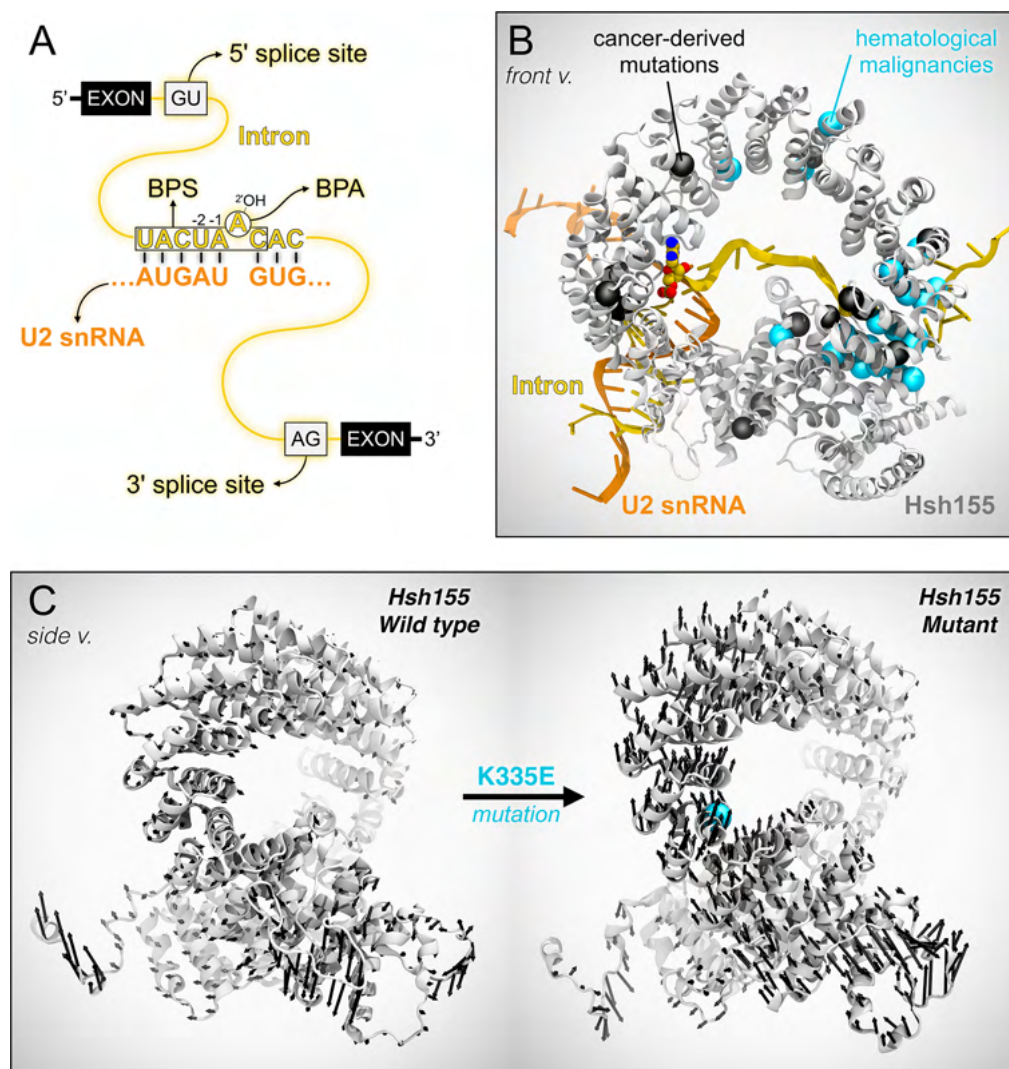
Figure 5. Intron and Hsh155 protein mutations and their influence on essential dynamics. (A) Key intron recognition site with branch point adenosine (BPA) at the branch site encircled in black. Alterations A-1U or U-2C weaken the intron-U2 base-pairing, which results in a nonconsensus branch point sequence (BPS). The intronic sequence at the 5′ splice site (SS) is marked by the highly conserved GU nucleotides and ends at the 3′SS by the conserved AG nucleotides. (B) Sites of mutations associated with hematological malignancies (cyan spheres) and other cancers (black spheres) are mapped on the Hsh155 protein (yeast ortholog of SF3B1). (C) Essential dynamics as revealed by principal component analysis (PCA) for the wild type (left) and mutant K335E (right) Hsh155 protein. The K355E mutation is highlighted with a cyan sphere. The black arrows indicate the direction of motion given by the first principal component, and the length of the arrows is proportional to the magnitude of the motion. The differences in the length indicate the distinct amplitude of the motion in the wild type and mutant system.

the B^{act} complex from the yeast *S. cerevisiae* was used to model the molecular mechanism underlying the selection of distinct BPS caused by SF3B1 mutations.[4] Mutations to the Hsh155 protein were introduced in silico and compared to B^{act} complexes containing wild type (wt) Hsh155. In particular, we considered two frequently observed mutations that are pathogenic in the human orthologue: K335E (K666E in human) and N295D (N626D in human), as well as the nonpathogenic L378V (I709V in human) variant, used as a control. In these simulations, wt and mutant Hsh155 were either binding a consensus or a nonconsensus BPSs (i.e., A > U or U > C in RNA positions −1 and −2 from the BPA, respectively). Hsh155 has a disordered N-terminus and a super helical spiral shape C-terminus, composed of 20 HEAT-repeats.[46] The BPA binding pocket is nestled between the H15−H16 HEAT-repeats. Peculiarly, most mutations implicated in hematologic malignancies in SF3B1 cluster at H5−H7, lying at a distance of

40 Å from the BPA binding pocket. Cross-correlation and PCA analyses of wt and mutant Hsh155 disclosed a springlike motion of the protein, with two hinge points located at the BPA binding pocket and at the region where most mutations cluster. This was also observed in simulations with human SF3B1.[17] Interestingly, when a pathogenic variant is introduced in Hsh155, intron binding is destabilized and/or the functional dynamics of Hsh155 are altered by a change in the position of these hinges (Figure 5). The carcinogenic mutations (N295D and K335E) induce these changes only upon binding of a nonconsensus BPS, in line with experimental evidence.[10,44] Conversely, in the presence of the nonpathogenic L378V variant, no intron destabilization nor variation of the Hsh155 functional dynamics was observed.[4] Although additional work is needed to determine whether these results extend to human SF3B1, this study identified an intriguing cross-talk between the distant Hsh155 mutation and BPA recognition sites, suggesting that the
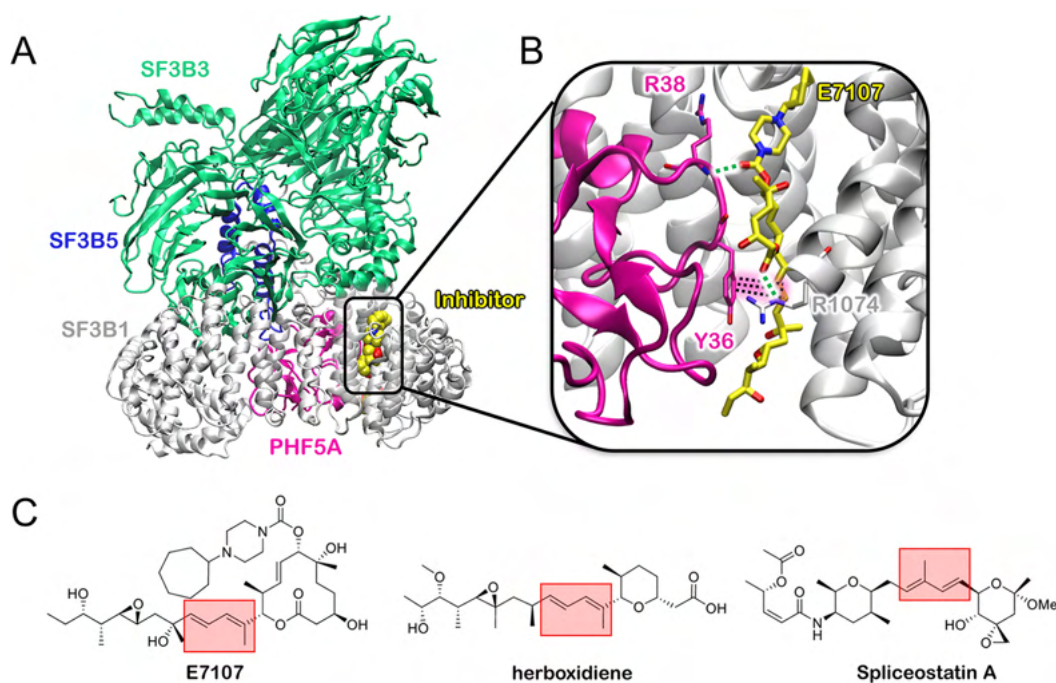
Figure 6. Splicing modulators and their binding site at the interface of PHF5A and SF3B1 proteins. (A) Model from the cryo-EM structure of the human SF3b complex bound to the E7107 inhibitor (PDB ID: 5ZYA). (B) Close view of the E7107 binding cavity at the interface of the proteins SF3B1 (gray) and PHF5A (magenta). Hydrogen bonds (dashed green lines), aromatic−diene π-interactions (dashed black lines highlighted in magenta) and residues that, when mutated, confer resistance to splicing modulators are shown. (C) Sketch of the splicing modulators E7107, herboxidiene, and spliceostatin A, wherein the conjugated diene motif is highlighted by a red rectangle.

pathogenic mutants may weaken nonconsensus BPS binding. Disruption of this cross-talk by cancer-associated mutations could lead to erroneous recruitment of a cryptic 3′SS.

SPLICING MODULATORS AS ANTICANCER AGENTS

Aberrant splicing is prevalent in cancer, affecting virtually every cancer hallmark. Splicing patterns causing adverse effects like proliferation, invasion, and angiogenesis can arise via mutations to splicing factors, as mentioned above, or through a dysregulation of splicing factor expression levels.[9] Moreover, there is often an unusual burden on the spliceosomal machinery in cancer, since tumor cells are rapidly dividing and highly anabolic. This has been described as a type of "splicing addiction," and has been especially linked to MYC-driven cancers.[47] Thus, splicing dysregulation creates a vulnerability in cancer cells, with or without mutations to the SPL or splicing factors, that can impart sensitivity to splicing inhibitors. Therefore, small-molecule splicing modulators (SMs) supply a potential selective therapeutic approach to counteract cancers.[44,48−50]

Interestingly, besides hosting frequent cancer-driving mutations, SF3B1 is also a target of SMs. These molecules mostly belong to the spliceostatin, herboxidiene, and pladienolide families and share a tripartite chemical structure with a central diene motif (Figure 6). The therapeutic potential of SMs has been demonstrated in many cancer models,[50] and the pladienolide derivative H3B-8800 is currently in a phase I clinical trial based on its potent and preferential ability to kill epithelial and hematologic tumor cells.[51] Since mutations to SF3B1 are linked to distinct tumor types,[52] SMs offer the appealing perspective of novel and personalized approaches for tissue-agnostic cancer treatment.

Cryo-EM and crystallographic information revealed that pladienolide B and E7107, a pladienolide derivative, bind at the BPA binding pocket, which is nestled between the proteins PHF5A and the H15−H16 HEAT-repeats of human SF3B1.[11,53] Using this structural information, models of the whole SF3b complex bound to either E7107, spliceostatin A, or herboxidiene were built to probe their inhibition mechanism. MD simulations were used as a tool to assess the impact of selected SM on the dynamical behavior of the SF3B1 protein, since, as detailed above, its function revolves around an opening/closing springlike motion.[53] Remarkably, each SM was found to differently affect the structural plasticity of SF3B1, locking its conformational equilibrium into the open state, which prevents binding of pre-mRNA, as suggested from crystallographic studies.[11] These results support and expand experimental evidence of SM action being not merely limited to a competitive inhibition, but rather affecting SF3B1's internal dynamics.[53] Overall, these findings provide potential valuable information that could be used for rational-based discovery of drugs to fight cancers associated with dysregulated splicing (Figure 6).[17]

METHODOLOGICAL PITFALLS ENCOUNTERED BY MULTISCALE SIMULATIONS

Despite the growing number of successful applications, all-atom simulations based on cryo-EM structural data of large macromolecular machinery, such as the SPL, face some methodological pitfalls. Often, these structures reach resolutions of 3−4 Å in the central body, while the peripheral components are either missing or poorly resolved and thus not being completely amenable to reliable molecular modeling.[54−56] Additionally, finding a compromise between system size and accuracy is in many cases not straightforward and remains arbitrary. In the

process of limiting the size of the system and the consequent computational cost of the simulations, a fundamental requirement is the selection of an insightful model of the SPL being computationally feasible without compromising the accuracy of the simulated biological system. However, even after a wise selection, the SPL model rapidly approaches, and sometimes even surpasses, one million atoms. In spite of the burgeoning development of computer hardware and software allowing for increasingly longer, brute force unbiased MD simulations, the direct sampling of biologically relevant events taking place within the SPL machinery (i.e., large conformational changes, drug-binding/dissociation, RNA binding and remodeling, splicing reactions) still remains out of reach. The bottleneck is mostly represented by the extreme computational cost ascribable to the large size of the systems and the complexity of underlying biological processes. Although a short conformational sampling (i.e., the statistics generated by MD simulations) could be partially addressed by comparing and combining different replicas of the same system, enhanced sampling methods still represent the first choice in order to force the event of interest within a computationally affordable time frame. These techniques have been successfully adopted to elucidate the chemical mechanism of splicing.[3,14,32] However, other events like RNA chaperoning and sculpting promoted by substantial conformational changes of the SPL as well as binding/dissociation of drugs, splicing factors, and long RNA strands pose many challenges for these methods, calling for novel or improved algorithms. Finally, MD simulations rely on FFs, which are inherently approximated. Their limitations, and associated uncertainties, particularly affect metal-dependent RNA or protein/RNA machineries, owing to the well-known flaws of RNA[57] and Mg^{2+} ion[58] parameters. When modeling such systems, these critical aspects have to be handled with care, requiring benchmark studies to identify the most appropriate computational protocol, and a strict experimental validation in order to obtain trustworthy and meaningful results.[3] The reliability of these approaches has also been established by the successful characterization of allostery,[23] selectivity,[59] and catalysis[60] of other protein/RNA machineries such as CRISPR-Cas9.

■ SUMMARY AND PERSPECTIVES

In spite of the challenges and limitations listed above, atomic-level computer simulations have significantly advanced the comprehension of the splicing mechanism in health and disease states and have provided the conceptual basis for the development of novel splicing modulators that selectively target cancer cells. New methodological advances in computer simulations, modeling, and analysis techniques will foster longer and increasingly accurate atomic-level studies of splicing to address unsolved mechanistic and biological questions. As an alternative, the use of coarse grain (CG) models, in which multiple atoms are described as single beads, may enable the exploration of systems of significantly larger size and for a longer time scale, bridging the gap between experiments and all-atom simulations.[57]

In conclusion, the showcased applications exemplify how atomic-level computer simulations can contribute to a deeper comprehension of RNA splicing, a critical step of gene expression. Molecular simulations clarify, support, and enrich the available structural, biochemical, and genetic data. We expect that the synergy between computer simulations and experiments will lead to a thorough understanding of the complex nature of the SPL machinery and will contribute to the development of splicing-based therapeutic strategies and revolutionary gene modulation tools for treating the many diseases associated with splicing mis-regulation.

■ AUTHOR INFORMATION

Corresponding Author

Alessandra Magistrato − *CNR-IOM-Democritos National Simulation Center c/o SISSA, Trieste 34136, Italy;* orcid.org/0000-0002-2003-1985; Email: alessandra.magistrato@sissa.it

Authors

Jure Borišek − *Theory Department, National Institute of Chemistry, Ljubljana 1001, Slovenia;* orcid.org/0000-0003-3417-0940

Lorenzo Casalino − *Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093, United States;* orcid.org/0000-0003-3581-1148

Andrea Saltalamacchia − *International School for Advanced Studies (SISSA/ISAS), Trieste 34136, Italy;* orcid.org/0000-0003-1174-9271

Suzanne G. Mays − *Centre for Genomic Regulation (CRG), Barcelona 08003, Spain*

Luca Malcovati − *Department of Molecular Medicine, University of Pavia, Pavia 27100, Italy; Department of Hematology, IRCCS S. Matteo Hospital Foundation, Pavia 27100, Italy*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.accounts.0c00578

Author Contributions

J.B. and L.C. equally contributed to this work. The manuscript has been written through contributions of all authors and all authors have approved the final version of the manuscript.

Notes

The authors declare no competing financial interest.

Biographies

Jure Borišek is an independent researcher at the National Institute of Chemistry Ljubljana, Slovenia. He is a former member of Dr. Magistrato's group and applies molecular simulations to disclose fundamental mechanistic aspects of splicing and of splicing modulators.

Lorenzo Casalino is a postdoctoral associate at the University of California at San Diego, USA. He is a former member of the Magistrato lab, and his work has pioneered the application of molecular simulations for the mechanistic understanding of splicing.

Andrea Saltalamacchia, a member of Dr. Magistrato's group, applies and develops sophisticated statistical analysis tools to study allosteric signaling of SPL.

Suzanne Mays is a postdoctoral researcher in Juan Valcarcel's group at the Centre for Genomic Regulation in Barcelona, Spain. Her research focuses on elucidating mechanisms of action of small-molecule modulators that bind the spliceosome, using a combination of structural biology, molecular biology, and bioinformatics.

Luca Malcovati leads the laboratory of Hematology Oncologyat University of Pavia & IRCCS Policlinico San Matteo Foundation, a reference center for myeloid neoplasms. He has provided important

contributions in the discovery of somatic mutation of critical splicing factors implicated in hematologic cancers.

Alessandra Magistrato leads the Laboratory of Computational Biology and Biochemistry, and she is deputy director of the National Research Council of Italy (CNR-IOM) c/o SISSA in Trieste, Italy. Her group has provided important contribution to understand the splicing mechanism at atomic-level.

REFERENCES

(1) Casalino, L.; Palermo, G.; Spinello, A.; Rothlisberger, U.; Magistrato, A. All-atom simulations disentangle the functional dynamics underlying gene maturation in the intron lariat spliceosome. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 6584−6589.

(2) Saltalamacchia, A.; Casalino, L.; Borišek, J.; Batista, V. S.; Rivalta, I.; Magistrato, A. Decrypting the Information Exchange Pathways across the Spliceosome Machinery. *J. Am. Chem. Soc.* **2020**, *142*, 8403−8411.

(3) Borišek, J.; Magistrato, A. All-Atom Simulations Decrypt the Molecular Terms of RNA Catalysis in the Exon-Ligation Step of the Spliceosome. *ACS Catal.* **2020**, *10*, 5328−5334.

(4) 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.

(5) Chow, L. T.; Gelinas, 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.

(6) Papasaikas, P.; Valcarcel, J. The Spliceosome: The Ultimate RNA Chaperone and Sculptor. *Trends Biochem. Sci.* **2016**, *41*, 33−45.

(7) Wan, R.; Bai, R.; Shi, Y. Molecular choreography of pre-mRNA splicing by the spliceosome. *Curr. Opin. Struct. Biol.* **2019**, *59*, 124−133.

(8) Wilkinson, M. E.; Charenton, C.; Nagai, K. RNA Splicing by the Spliceosome. *Annu. Rev. Biochem.* **2020**, *89*, 359−388.

(9) Bonnal, S. C.; López-Oreja, I.; Valcárcel, J. Roles and mechanisms of alternative splicing in cancer - implications for care. *Nat. Rev. Clin. Oncol.* **2020**, *17*, 457−474.

(10) 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.

(11) Cretu, C.; Agrawal, A. A.; Cook, A.; Will, C. L.; Fekkes, P.; Smith, P. G.; Luhrmann, R.; Larsen, N.; Buonamici, S.; Pena, V. Structural Basis of Splicing Modulation by Antitumor Macrolide Compounds. *Mol. Cell* **2018**, *70*, 265.

(12) Cheng, Y.; Grigorieff, N.; Penczek, P. A.; Walz, T. A primer to single-particle cryo-electron microscopy. *Cell* **2015**, *161*, 438−449.

(13) Wan, R. X.; Yan, C. Y.; Bai, R.; Lei, J. L.; Shi, Y. G. Structure of an Intron Lariat Spliceosome from Saccharomyces cerevisiae. *Cell* **2017**, *171*, 120−132.

(14) Casalino, L.; Palermo, G.; Rothlisberger, U.; Magistrato, A. Who Activates the Nucleophile in Ribozyme Catalysis? An Answer from the Splicing Mechanism of Group II Introns. *J. Am. Chem. Soc.* **2016**, *138*, 10374−10377.

(15) Samy, A.; Suzek, B. E.; Ozdemir, M. K.; Sensoy, O. In Silico Analysis of a Highly Mutated Gene in Cancer Provides Insight into Abnormal mRNA Splicing: Splicing Factor 3B Subunit 1(K700E) Mutant. *Biomolecules* **2020**, *10*, 680.

(16) Casalino, L.; Magistrato, A. Unraveling the Molecular Mechanism of Pre-mRNA Splicing From Multi-Scale Simulations. *Front. Mol. Biosci.* **2019**, *6*, 62−62.

(17) 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.* **2020**, *60*, 2510−2521.

(18) Hollingsworth, S. A.; Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **2018**, *99*, 1129−1143.

(19) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. Essential Dynamics of Proteins. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 412−425.

(20) Yan, C. Y.; Hang, J.; Wan, R. X.; Huang, M.; Wong, C. C. L.; Shi, Y. G. Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* **2015**, *349*, 1182−1191.

(21) Zhang, X.; Zhan, X.; Yan, C.; Zhang, W.; Liu, D.; Lei, J.; Shi, Y. Structures of the human spliceosomes before and after release of the ligated exon. *Cell Res.* **2019**, *29*, 274−285.

(22) Lange, O. F.; Grubmuller, H. Generalized correlation for biomolecular dynamics. *Proteins: Struct., Funct., Genet.* **2006**, *62*, 1053−1061.

(23) Palermo, G.; Ricci, C. G.; Fernando, A.; Basak, R.; Jinek, M.; Rivalta, I.; Batista, V. S.; McCammon, J. A. Protospacer Adjacent Motif-Induced Allostery Activates CRISPR-Cas9. *J. Am. Chem. Soc.* **2017**, *139*, 16028−16031.

(24) Rivalta, I.; Sultan, M. M.; Lee, N. S.; Manley, G. A.; Loria, J. P.; Batista, V. S. Allosteric pathways in imidazole glycerol phosphate synthase. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, E1428−E1436.

(25) Sgrignani, J.; Magistrato, A. QM/MM MD Simulations on the Enzymatic Pathway of the Human Flap Endonuclease (hFEN1) Elucidating Common Cleavage Pathways to RNase H Enzymes. *ACS Catal.* **2015**, *5*, 3864−3875.

(26) Sgrignani, J.; Magistrato, A. The Structural Role of Mg2+ Ions in a Class I RNA Polymerase Ribozyme: A Molecular Simulation Study. *J. Phys. Chem. B* **2012**, *116*, 2259−2268.

(27) Steitz, T. A.; Steitz, J. A. A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 6498−6502.

(28) Peebles, C. L.; Perlman, P. S.; Mecklenburg, K. L.; Petrillo, M. L.; Tabor, J. H.; Jarrell, K. A.; Cheng, H. L. A self-splicing RNA excises an intron lariat. *Cell* **1986**, *44*, 213−223.

(29) Schmelzer, C.; Schweyen, R. J. Self-splicing of group II introns in vitro: mapping of the branch point and mutational inhibition of lariat formation. *Cell* **1986**, *46*, 557−565.

(30) Brunk, E.; Rothlisberger, U. Mixed Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulations of Biological Systems in Ground and Electronically Excited States. *Chem. Rev.* **2015**, *115*, 6217−6263.

(31) Marcia, M.; Pyle, A. M. Visualizing Group II Intron Catalysis through the Stages of Splicing. *Cell* **2012**, *151*, 497−507.

(32) Huang, W. T.; Huang, Y.; Xu, J.; Liao, J. L. How Does the Spliceosome Catalyze Intron Lariat Formation? Insights from Quantum Mechanics/Molecular Mechanics Free-Energy Simulations. *J. Phys. Chem. B* **2019**, *123*, 6049−6055.

(33) Bai, R.; Yan, C.; Wan, R. X.; Lei, J. L.; Shi, Y. G. Structure of the Post-catalytic Spliceosome from Saccharomyces cerevisiae. *Cell* **2017**, *171*, 1589−1598.

(34) Lim, K. H.; Ferraris, L.; Filloux, M. E.; Raphael, B. J.; Fairbrother, W. G. Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 11093−11098.

(35) Bonnal, S. C.; López-Oreja, I.; Valcárcel, J. Roles and mechanisms of alternative splicing in cancer - implications for care. *Nat. Rev. Clin. Oncol.* **2020**, *17*, 457−474.

(36) Chen, M.; Manley, J. L. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 741−754.

(37) El Marabti, E.; Younis, I. The Cancer Spliceome: Reprograming of Alternative Splicing in Cancer. *Front. Mol. Biosci.* **2018**, *5*, 80.

(38) Shuai, S. M.; Suzuki, H.; Diaz-Navarro, A.; Nadeu, F.; Kumar, S. A.; Gutierrez-Fernandez, A.; Delgado, J.; Pinyol, M.; Lopez-Otin, C.; Puente, X. S.; Taylor, M. D.; Campo, E.; Stein, L. D. The U1 spliceosomal RNA is recurrently mutated in multiple cancers. *Nature* **2019**, *574*, 712−716.

(39) Visconte, V.; Nakashima, M. O.; Rogers, H. J. Mutations in Splicing Factor Genes in Myeloid Malignancies: Significance and Impact on Clinical Features. *Cancers* **2019**, *11*, 1844.

(40) Inoue, D.; Abdel-Wahab, O. Modeling SF3B1Mutations in Cancer: Advances, Challenges, and Opportunities. *Cancer Cell* **2016**, *30*, 371−373.

(41) Yoshida, K.; Sanada, M.; Shiraishi, Y.; Nowak, D.; Nagata, Y.; Yamamoto, R.; Sato, Y.; Sato-Otsubo, A.; Kon, A.; Nagasaki, M.; Chalkidis, G.; Suzuki, Y.; Otsu, M.; Obara, N.; Sakata-Yanagimoto, M.; Ishiyama, K.; Mori, H.; Nolte, F.; Hofmann, W. K.; Miyawaki, S.; Sugano, S.; Haferlach, C.; Koeffler, H. P.; Shih, L. Y.; Haferlach, T.; Chiba, S.; Nakauchi, H.; Miyano, S.; Ogawa, S. Frequent Pathway Mutations of Splicing Machinery in Myelodysplasia. *Blood* **2011**, *118*, 458.

(42) Papaemmanuil, E.; Cazzola, M.; Boultwood, J.; Malcovati, L.; Vyas, P.; Bowen, D.; Pellagatti, A.; Wainscoat, J. S.; Hellstrom-Lindberg, E.; Gambacorti-Passerini, C.; Godfrey, A. L.; Rapado, I.; Cvejic, A.; Rance, R.; McGee, C.; Ellis, P.; Mudie, L. J.; Stephens, P. J.; McLaren, S.; Massie, C. E.; Tarpey, P. S.; Varela, I.; Nik-Zainal, S.; Davies, H. R.; Shlien, A.; Jones, D.; Raine, K.; Hinton, J.; Butler, A. P.; Teague, J. W.; Baxter, E. J.; Score, J.; Galli, A.; Della Porta, M. G.; Travaglino, E.; Groves, M.; Tauro, S.; Munshi, N. C.; Anderson, K. C.; El-Naggar, A.; Fischer, A.; Mustonen, V.; Warren, A. J.; Cross, N. C. P.; Green, A. R.; Futreal, P. A.; Stratton, M. R.; Campbell, P. J.; Consortium, I. C. G. Somatic SF3B1Mutation in Myelodysplasia with Ring Sideroblasts. *N. Engl. J. Med.* **2011**, *365*, 1384−1395.

(43) Lieu, Y. K.; Liu, Z.; Ali, A. M.; Wei, X.; Penson, A.; Zhang, J.; An, X.; Rabadan, R.; Raza, A.; Manley, J. L.; Mukherjee, S. SF3B1 mutant-induced missplicing of MAP3K7 causes anemia in myelodysplastic syndromes. *bioRxiv*, June 26, 2020, ver. 1. DOI: 10.1101/2020.06.24.169052.

(44) Cretu, C.; Schmitzova, J.; Ponce-Salvatierra, A.; Dybkov, O.; De Laurentiis, E. I.; Sharma, K.; Will, C. L.; Urlaub, H.; Luhrmann, R.; Pena, V. Molecular Architecture of SF3b and Structural Consequences of Its Cancer-Related Mutations. *Mol. Cell* **2016**, *64*, 307−319.

(45) 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.

(46) Yan, C. Y.; Wan, R. X.; Bai, R.; Huang, G. X. Y.; Shi, Y. G. Structure of a yeast activated spliceosome at 3.5 angstrom resolution. *Science* **2016**, *353*, 904−911.

(47) Hsu, T. Y. T.; Simon, L. M.; Neill, N. J.; Marcotte, R.; Sayad, A.; Bland, C. S.; Echeverria, G. V.; Sun, T. T.; Kurley, S. J.; Tyagi, S.; Karlin, K. L.; Dominguez-Vidana, R.; Hartman, J. D.; Renwick, A.; Scorsone, K.; Bernardi, R. J.; Skinner, S. O.; Jain, A.; Orellana, M.; Lagisetti, C.; Golding, I.; Jung, S. Y.; Neilson, J. R.; Zhang, X. H. F.; Cooper, T. A.; Webb, T. R.; Neel, B. G.; Shaw, C. A.; Westbrook, T. F. The spliceosome is a therapeutic vulnerability in MYC-driven cancer. *Nature* **2015**, *525*, 384−388.

(48) Patnaik, M. M.; Lasho, T. L.; Finke, C. M.; Hanson, C. A.; Hodnefield, J. M.; Knudson, R. A.; Ketterling, R. P.; Pardanani, A.; Tefferi, A. Spliceosome mutations involving SRSF2, SF3B1, and U2AF35 in chronic myelomonocytic leukemia: Prevalence, clinical correlates, and prognostic relevance. *Am. J. Hematol.* **2013**, *88*, 201−206.

(49) Landau, D. A.; Carter, S. L.; Stojanov, P.; McKenna, A.; Stevenson, K.; Lawrence, M. S.; Sougnez, C.; Stewart, C.; Sivachenko, A.; Wang, L. L.; Wan, Y. Z.; Zhang, W. D.; Shukla, S. A.; Vartanov, A.;

Fernandes, S. M.; Saksena, G.; Cibulskis, K.; Tesar, B.; Gabriel, S.; Hacohen, N.; Meyerson, M.; Lander, E. S.; Neuberg, D.; Brown, J. R.; Getz, G.; Wu, C. J. Evolution and Impact of Subclonal Mutations in Chronic Lymphocytic Leukemia. *Cell* **2013**, *152*, 714−726.

(50) Effenberger, K. A.; Urabe, V. K.; Prichard, B. E.; Ghosh, A. K.; Jurica, M. S. Interchangeable SF3B1 inhibitors interfere with pre-mRNA splicing at multiple stages. *RNA* **2016**, *22*, 350−359.

(51) Seiler, M.; Yoshimi, A.; Darman, R.; Chan, B.; Keaney, G.; Thomas, M.; Agrawal, A. A.; Caleb, B.; Csibi, A.; Sean, E.; Fekkes, P.; Karr, C.; Klimek, V.; Lai, G.; Lee, L.; Kumar, P.; Lee, S. C. W.; Liu, X.; Mackenzie, C.; Meeske, C.; Mizui, Y.; Padron, E.; Park, E.; Pazolli, E.; Peng, S. Y.; Prajapati, S.; Taylor, J.; Teng, T.; Wang, J.; Warmuth, M.; Yao, H. L.; Yu, L. H.; Zhu, P.; Abdel-Wahab, O.; Smith, P. G.; Buonamici, S. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. *Nat. Med.* **2018**, *24*, 497−504.

(52) Yoshida, K.; Ogawa, S. Splicing factor mutations and cancer. *Wires RNA* **2014**, *5*, 445−459.

(53) Finci, L. I.; Zhang, X. F.; Huang, X. L.; Zhou, Q.; Tsai, J.; Teng, T.; Agrawal, A.; Chan, B.; Irwin, S.; Karr, C.; Cook, A.; Zhu, P.; Reynolds, D.; Smith, P. G.; Fekkes, P.; Buonamici, S.; Larsen, N. A. The cryo-EM structure of the SF3b spliceosome complex bound to a splicing modulator reveals a pre-mRNA substrate competitive mechanism of action. *Genes Dev.* **2018**, *32*, 309−320.

(54) Kastner, B.; Will, C. L.; Stark, H.; Lührmann, R. Structural Insights into Nuclear pre-mRNA Splicing in Higher Eukaryotes. *Cold Spring Harbor Perspect. Biol.* **2019**, *11*, No. a032417.

(55) Plaschka, C.; Newman, A. J.; Nagai, K. Structural Basis of Nuclear pre-mRNA Splicing: Lessons from Yeast. *Cold Spring Harbor Perspect. Biol.* **2019**, *11*, No. a032391.

(56) Yan, C.; Wan, R.; Shi, Y. Molecular Mechanisms of pre-mRNA Splicing through Structural Biology of the Spliceosome. *Cold Spring Harbor Perspect. Biol.* **2019**, *11*, No. a032409.

(57) Sponer, J.; Bussi, G.; Krepl, M.; Banas, P.; Bottaro, S.; Cunha, R. A.; Gil-Ley, A.; Pinamonti, G.; Poblete, S.; Jureacka, P.; Walter, N. G.; Otyepka, M. RNA Structural Dynamics As Captured by Molecular Simulations: A Comprehensive Overview. *Chem. Rev.* **2018**, *118*, 4177−4338.

(58) Casalino, L.; Palermo, G.; Abdurakhmonova, N.; Rothlisberger, U.; Magistrato, A. Development of Site-Specific Mg2+-RNA Force Field Parameters: A Dream or Reality? Guidelines from Combined Molecular Dynamics and Quantum Mechanics Simulations. *J. Chem. Theory Comput.* **2017**, *13*, 340−352.

(59) Ricci, C. G.; Chen, J. S.; Miao, Y.; Jinek, M.; Doudna, J. A.; McCammon, J. A.; Palermo, G. Deciphering Off-Target Effects in CRISPR-Cas9 through Accelerated Molecular Dynamics. *ACS Cent. Sci.* **2019**, *5*, 651−662.

(60) Casalino, L.; Nierzwicki, Ł.; Jinek, M.; Palermo, G. Catalytic Mechanism of Non-Target DNA Cleavage in CRISPR-Cas9 Revealed by Ab Initio Molecular Dynamics. *ACS Catal.* **2020**, *10*, 13596−13605.