Tracing Allostery in the Spliceosome Ski2-like RNA Helicase Brr2

Francesco Guidarelli Mattioli,[†] Andrea Saltalamacchia,[‡] and Alessandra Magistrato[*,¶]

[†]*Sapienza University of Rome, Piazzale Aldo Moro 2, 00185 Rome, Italy*

[‡]*International School for Advanced Studies (SISSA/ISAS), via Bonomea 265, 34136 Trieste, Italy*

[¶]*National Research Council of Italy, Institute of Material Foundry at International School for Advanced Studies (SISSA/ISAS), via Bonomea 265, 34136 Trieste, Italy*

E-mail: alessandra.magistrato@sissa.it

Abstract

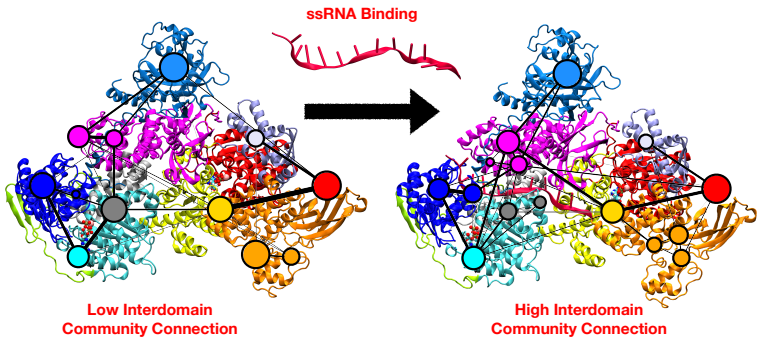
RNA ATPases/helicases remodel substrate RNA–protein complexes in distinct ways. The different RNA ATPases/helicases, taking part to the spliceosome complex, reshape the RNA/RNA-protein contacts to enable premature-messenger RNA splicing. Among them, the Bad Response to Refrigeration 2 (Brr2) helicase promotes U4/U6 small nuclear (sn)RNA unwinding via ATP-driven translocation of the U4 snRNA strand, thus playing a pivotal role during the activation, catalytic and disassembly phases of splicing. The plastic Brr2 architecture consists of an enzymatically active N-terminal cassette (N-cassette) and a structurally similar, but inactive, C-terminal cassette (C-cassette). The C-cassette, along with other allosteric effectors and regulators, tightly and timely controls Brr2's function via an elusive mechanism.

Here, microsecond-long molecular dynamics simulations, dynamical network theory and community network analysis are combined to elucidate how allosteric effectors/regulators modulate the Brr2 function. We also unexpectedly reveal that U4 snRNA itself acts as an allosteric regulator, amplifying the cross-talk of distal Brr2 domains and triggering a conformational reorganization of the protein. Our findings offer fundamental understanding into Brr2's mechanism of action, and broaden

our knowledge on the sophisticated regulatory mechanisms by which spliceosome ATPases/helicases control gene expression. This includes the allosteric regulation exerted by client RNA strands, a mechanism that may be broadly applicable to other RNA-dependent ATPases/helicases.

TOC Graphic

Here TOC Graphic.



TOC Graphic

Present in all domains of life, RNA-dependent ATPases/helicases take part to many key aspects of RNA metabolism, from transcription and processing to translation and decay. At the functional level, RNA helicases promote rearrangements of RNAs and RNP particles by coupling nucleic acid binding and release with ATP hydrolysis. Structural and biochemical studies pinpointed the plasticity of RNA ATPases/helicases as a fundamental and instrumental property to control their timely activation in the cell.[1] Among the key cellular processes in which RNA ATPases/helicases play key role is precursor messenger RNA (pre-mRNA) splicing, a process which produces functional protein-coding mRNAs and long non-coding RNA.[2,3] The spliceosome, a complex ribonuleoprotein (RNP) machinery, composed by five small nuclear RNAs (snRNAs) and hundreds of proteins, catalyzes pre-mRNA splicing.[5] To perform this process with high precision the spliceosome assembles de novo at each cycle and undergoes substantial conformational and compositional changes, forming different intermediates (A, B, Bact, B*, C, C*, P, ILS) (Figure S1A).[6] Its progression through these states is fueled by eight distinct RNA-dependent ATPases/RNA helicases.[7] Among them, Brr2 contributes to spliceosome activation by promoting U4/U6 small nuclear (sn)RNA unwinding. Namely, Brr2 couples ATP hydrolysis to U4 snRNA translocation, thus allowing U6 snRNA to form the RNA-based catalytic site of the spliceosome.[5] Thus, Brr2 plays a key role in splicing metabolism by timing the RNAs and protein/RNA structural changes and by proofreading the RNA strands (Figure S1A). Besides tacking part to pre-mRNA splicing, Brr2 is involved in other cellular processes such as RNA degradation, viral defence, regulation of circadian rhythm, DNA repair.[13] Belonging to the Ski2-like subfamily of RNA-dependent SF2 ATPase/helicase, Brr2's has an almost duplicated architecture as compared to the other spliceosome DExH/DExD-box helicases,[8,9] being composed by two structurally similar N-termimal and C-terminal cassettes (referred hereafter N-cassette and C-cassette, respectively). Each cassette is made of two RecA-like domains (RecA1 and RecA2), a widget helix (WH) and a Sec63 homology unit, which contains a helix bundle (HB), a helix-loop-helix (HLH), and an immunoglobulin (IG) domain (all domains belonging to N-Cassette and C-cassette are hereafter referred with an initial N or C letter, respectively) (Figure 1).

Its unique and striking architectural complexity makes Brr2 function amenable to be controlled via intra- and inter-molecular regulatory mechanisms. Allostery is a fundamental concept in biochemistry describing the communication between distant sites in a protein, which leads to a functional response.[10–12] This phenomenon allows proteins to adapt to their environment and carry out complex biological processes under the action of ligand-effectors and protein-regulators. Allosteric regulation is a pervasive principle in RNA helicases, being paramount to Brr2 function.[13] Indeed, in Brr2 only the N-cassette exhibits AT-

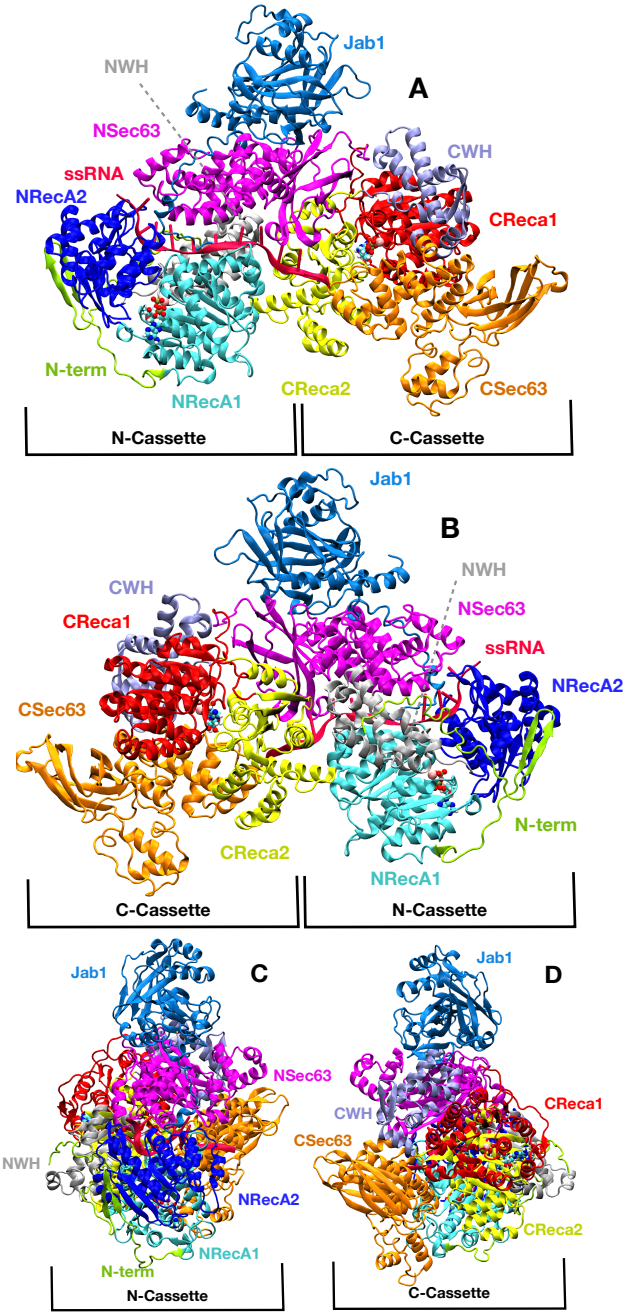


Figure 1: **(A-D)** Structure of Brr2 ATPase/helicase (PDB code 4KIT and 5O9Z for Brr2 and RNA coordinates, respectively) in frontal **(A)**, back **(B)** and two sides view **(C-D)**. In N-terminal-cassette N-loop, NRecA1, NRecA2, NWH, NSec63 domains are depicted in light-green, cyan, blue, grey, magenta new-cartoons, respectively. In C-terminal-cassette CRecA1, CrecA2, CWH, CSec63 domains are shown in red, yellow, lilac and orange new-cartoons, respectively. The Jab1 domain of Prp8 protein and U4 snRNA are shown in light-blue and red, respectively. The Mg^{2+} ion and the ATP molecule are shown in balls and colored by atom name. Each cassette is indicated by the black box.

Pase/RNA unwinding activity, while the C- cassette is enzymatically inactive, but it al-

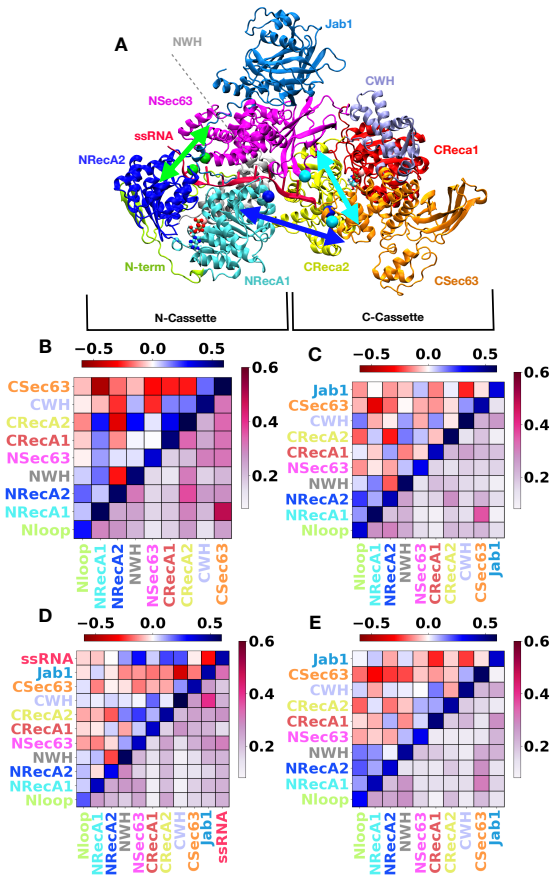


Figure 2: **(A)** View of Brr2 ATPase/helicase with highlighted the distances selected to monitor the Longitudal Opening-closing motion (LOC),Tranversal OC motion (TOC) and the TOC motion within the N-cassette (NcTOC), respectively D1 (distance between Glu578@C_α@NRecA1 and Glu1864@C_α@CSec63, blue arrow), D2 (distance between Lys1244@C_α@NSec63 and Arg1869@C_α@CSec63, cyan arrow) and D3 (distance between Gly788@C_α@NRecA2 and Met1078@C_α@NSec63, green arrow). **(B-E)** Simplified cross-correlation matrices bases on the Pearson correlation coefficients. Density Correlation Scores (DCSs) (top half-matrix) and Absolute Density Correlation Scores (ADCSs) (bottom half-matrix) are shown for Brr2 in complex with **(B)** two ATP molecules ($Brr2_{ATP}$), **(C)** two ATPs and Prp8-Jab1 domain ($^{Jab1}Brr2_{ATP}$), **(D)** two ATPs, Prp8-Jab1 and U4 snRNA ($^{U4/Jab1}Brr2_{ATP}$), **(E)** two ADPs and Prp8-Jab1 ($^{Jab1}Brr2_{ADP}$). Correlation-scores are reported in the range -0.6 to 0.6 and the absolute CSs in the range from 0.1 to 0.6 for clarity reasons. Each domain is labelled with the same color of Fig 1.

losterically stimulates the binding of ATP and the catalytic activity of N-cassette.[13] Additionally, it acts as binding platform of distinct protein-regulators.[7,8] Among the proteins regulating Brr2 function is Prp8. Its Jab1 domain[14] can in fact inhibit Brr2 function, by inserting its C-term tail into the Brr2 RNA binding tunnel or enhance Brr2 function via the binding of globular part to Brr2.[15] Moreover, U4 snRNA binding was shown to increase the interactions between the two cassettes (Figure S1B) and to acti-

vate ATP hydrolysis, similarly to other RNA-dependent ATPase/helicases.[8,16]

The pivotal importance of the Brr2 allosteric regulation for proper function and for splicing fidelity maintenance is remarked by the fact that Brr2 mutations placed along at the RNA binding channel and mutations of the Prp8-Jab1 domain tail, which, in the spliceosome complex, faces the N-cassette, are associated to RP33[17] and RP13[18] forms of Retinitis pigmentosa, respectively, an inherited disease associated with progressive degeneration of the retina.[19]

Despite its importance for splicing, the diverse molecular mechanisms underlying Brr2 allosteric modulation remain elusive. Brr2, therefore, appears as an appealing target system to extricate the hierarchy of sophisticated regulatory mechanisms underlying gene expression and regulation.

To this end we performed μs-long molecular dynamics (MD) simulations combined with Network Theory (NWA) and Community Network Analysis (CNA). These techniques, applied to the analysis of MD simulation trajectories, allow to trace signaling-paths taking place within a protein and to identify densely connected sub-networks, representing functionally-related groups of atoms that act in synergy to perform specific biological functions. These methods have been successfully applied to other protein/RNA machines.[20–22]

Briefly, to extricate the interconnection of Brr2 domains under the stimulus of different effectors/regulators we build different model systems: Brr2 in complex with: (i) two ATP molecules, hereafter named as $\mathrm{Brr2_{ATP}}$; (ii) two ATPs and Prp8-Jab1 domain, hereafter named as $^{\mathrm{Jab1}}\mathrm{Brr2_{ATP}}$; (iii) Prp8-Jab1, two ATPs and U4 snRNA, hereafter named as $^{\mathrm{U4/Jab1}}\mathrm{Brr2_{ATP}}$; (iv) Prp8-Jab1 and two ADPs, hereafter named as $^{\mathrm{Jab1}}\mathrm{Brr2_{ADP}}$. Finally, to understand the possible influence of the catalytically-inactive C-cassette on N-cassette, which has ATPse function, we also considered the N-cassette alone in the presence of ATP, hereafter named as $\mathrm{NCBrr2_{ATP}}$ (See Method section in the Supporting Information (SI) for details).

MD simulations of all explicitly-solvated models were done with the Amber ff14sb[23] force field and the OL3 correction for RNA.[24,25] All systems reached structural stability within 1 μs (Figure S2-S3), but ($\mathrm{Brr2_{ATP}}$), whose MD simulation was extended due to larger fluctuations of the two cassettes (Figure S4). Further analysis assessing the dependence of the calculated properties on the length of the MD simulations are shown in Figures S5-S7. To inspect the dynamic coupling of Brr2 domains we initially computed the cross-correlation matrix (CCM) of all systems in its original (Figure S8) and simplified variants (Figure 2 B-E, upper diagonal). The latter is calculated by summing the cross-correlation scores (CCs) of each couple of Brr2 domains and by averaging it by the number of residues belonging to each couple of domains. The resulting density correlation scores (DCS) provide a first simplified glimpse on the dynamical coupling tacking place in the system.[4] Addi-

tionally, we also computed the absolute DCS by summing of the absolute values of the CCs of each couple of BBr2 domain to monitor the effect of the coarsening procedure (Figure 2 B-E, lower diagonal). Although the simplified CCM provides a coarser picture of the internal dynamical coupling, here no major differences were observed with respect to its original counterpart.

Similarly to a previous study,[13] we observed an intra-cassette negative-correlation, involving NWH and NRecA2. This dynamical coupling is common to all ATP-containing systems, but it vanishes after ATP hydrolysis. An inter-cassette negative-correlation between NRecA1 and CSec63 domains and between NRecA2 and CRecA2 domains was instead shared by all systems. Additionally we complemented this analysis by calculating the energetic coupling between different domains (see Supporting Result section of SI). This supplies a picture of the interaction energy strength between the couple of domains and between the each domain and the effectors, which is compatible with that obtained from the CCM (Figure S9). Notably, the U4 snRNA is establishing strong hydrophobic and electrostatic interactions with Nsec63 domain of Brr2 and the Jab1 domain of Prp8. As well, strong interactions are present between Nsec63 and Jab1 domains.

We next performed the Principal Components Analysis (PCA)(Figure S10), which allows to gather the most relevant functional motion hidden within a MD simulation trajectory. In $Brr2_{ATP}$ the essential dynamics under the stimulus of ATP effectors corresponds to (i) a longitudinal opening-closing (hereafter referred as LOC) movement between NRecA1/2 and CSec63, (Figure 3A, PC1); (ii) a transversal opening-closing (hereafter referred as TOC) motion of NSec63 and CSec63 (Figure 3B, PC2); (iii) an intra-cassette opening-closing motion of NRecA1 and NSec63 (hereafter referred as NcTOC, PC3) (Figure 3C). While the TOC and LOC may be functional to RNA translocation, the NcTOC motion underlies an opening-closing of the RNA binding tunnel and may be functional to RNA loading. Instead, the PCA of $NCBrr2_{ATP}$ reveals that PC1 corresponds to the NcTOC (Figure 3D) motion, while PC2 describes a type-writer-like rotatory motion between NSec63 and NRecA1-2 domains (Figure 3F). This movement was observed and suggested to promote RNA translocation in Prp2 ATPase/helicase.[26–28] Remarkably, in the complete Brr2 structure, the C-cassette hinders this movement and it is only the coordinated effort of the N- and C-cassettes that drives RNA translocation (Figure 3E).

Notably, all Brr2 models have qualitatively the same internal motions. However, to better inspect how these movements are influenced by different cofactors/effectors, we monitored the distribution of selected distances (Figure 2A) accounting for the amplitudes of each of these movements.

We initially analyzed the distance (D1) between Glu578@C_α@NRecA1 and Glu1864@C_α@CSec63 (Figures 3G and S6), which accounts for the LOC movement (Figure 3A). In $Brr2_{ATP}$ D1 shows a bimodal

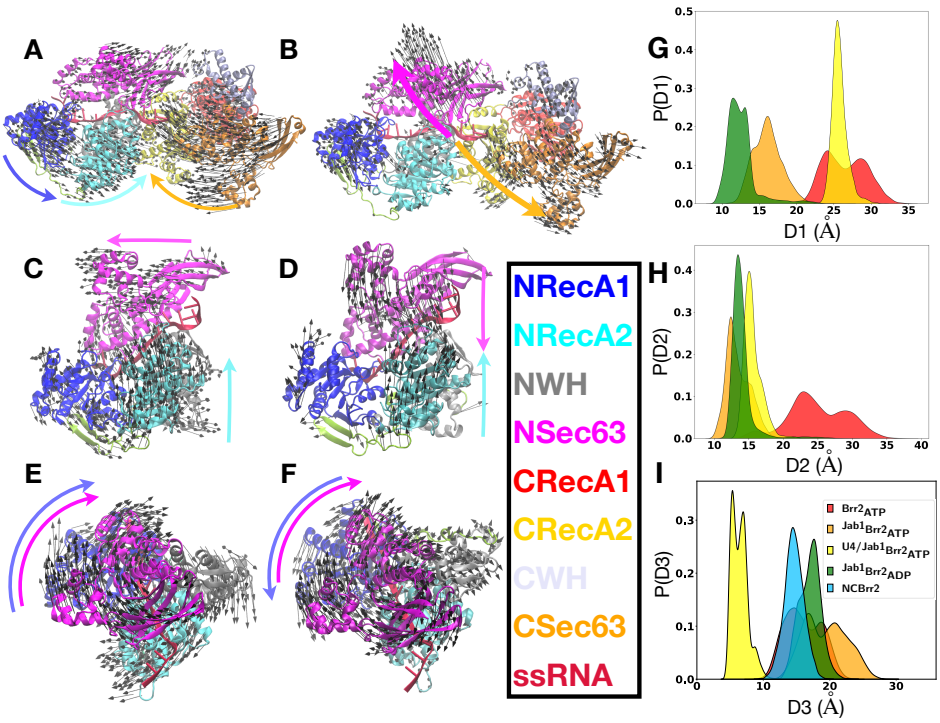


Figure 3: **(A)**-**(F)** Protein motion along Principal Components (PC) **(A)** along PC1 for $Brr2_{ATP}$; **(B)** along the PC2 for $Brr2_{ATP}$; **(C)** along the PC3 for the $Brr2_{ATP}$; **(D)** along the PC1 for the $NCBrr2_{ATP}$, **(E)** along the PC3 for the $Brr2_{ATP}$ and **(F)** along the PC2 for the $NCBrr2_{ATP}$. **(G)**-**(I)** Histograms of selected distances. **(G)** D1: distance between Glu578@C_α@NRecA1 and Glu1864@C_α@CSec63 accounting for the longitudinal opening-closing (LOC) movement, **(H)** D2: distance between Lys1244@C_α@NSec63 and Arg1869@C_α@CSec63 accounting for transversal opening-closing (TOC) movement and **(I)** D3: distance between Gly788@C_α@NRecA2 and Met1078@C_α@NSec63 accounting for transversal opening-closing of the N-cassette (NcTOC) movement.

distribution peaked at $20 - 30$ Å . This persists, although with reduced amplitude, when Jab1 binds, suggesting that Jab1 increases Brr2 compactness. A shift at larger values is, instead, observed in the ADP-containing model, suggesting that ATP hydrolysis also regulates this movement. Conversely, when the U4 snRNA is present, the LOC motion is damped since the RNA strand physically connects the two cassettes. For the TOC motion we instead inspected the distance (D2) between Lys1244@C_α@NSec63 and Arg1869@C_α@CSec63. We observed a bimodal distribution of D2 in $Brr2_{ATP}$, which becomes monomodal and shifts at lower values in all other systems containing Jab1 (Figures 3H and S6). This suggests that Jab1 markedly influences the TOC movement. For the NcTOC motion we analyzed the distance (D3) between Gly788@C_α@NRecA2 and Met1078@C_α@NSec63 (Figures 3I and S6). We observed that D3 has again bimodal distribution in $Brr2_{ATP}$. This shifts to smaller values in $NCBrr2_{ATP}$, while shifts to larger

values in $^{Jab1}Brr2_{ATP}$. Finally, after ATP hydrolysis becomes monomodal and reduced ($^{Jab1}Brr2_{ADP}$). Hence, the NcTOC motion appears to be stimulated by C-cassette, Jab1 and ATP, while being damped upon the RNA binding (Figure 3I).

Next, aimed at tracing the cross-communication paths underlying the observed functional motions, we performed NWA (see the Supplementary Methods and Results in the SI for a detailed description). In NWA the protein is represented as correlation-based weighted network, in which nodes (corresponding to residues' center of mass), are connected by edges. Each edge is then associated with a weight, which accounts for the amount of correlation for each pair of residues. Namely, small/large weights indicate highly/poorly correlated motions. By computing cross-correlations between residues along a MD trajectory, NWA finds the optimal and sub-optimal signaling-paths, through these weighted-edges which connects a user-selected source and a sink residue. The resulting paths have lengths, defined in network-space, which are inversely proportional to the amount of correlation existing among the constituent nodes. Firstly, we inspected the inter-cassette signaling-paths, underlying the LOC motion, by selecting as source and sink the Lys509@NRecA1 and Lys1356@CRecA1 residues, which hydrogen-bond to ATP (or ADP) in the N-cassette and C-cassette, respectively. We observed that four domains (NRecA1, NWH, CRecA1, CRecA2) are crossed by these communication-routes with the inter-cassette bridging-point being located at the NRecA1 and CRecA2 interface (Figure 4A) in all systems except for $^{U4/Jab1}Brr2_{ATP}$. Interestingly, the RNA diverts inter-cassette signal-exchange point to the NWH/CRecA2 interface (Figure 4A).

Indeed, U4 snRNA binding remodels hydrogen-bond pattern, reducing their formation at NRecA1/CRecA2 interface (Table S4), while inducing a persistent hydrogen-bond at the NWH/CRecA2 interface (i.e. Val906@NWH-Ser1541@CRecA2) (Tables S6-S9).

We next inspected the degeneracy (i.e. the number of times a residue is present in the signaling-paths) of the residues taking part to these communication-routes. Indeed, residues with high-degeneracy score are essential conveyors of information-exchange along the optimal/sub-optimal paths and it has been shown they most often correspond to experimentally validated residues for protein function.[29–31] Consistently, among the highly-degenerate residues we identified Arg603-Arg637@NRecA1 and His1548@CRecA2 (Figure S11A), which were demonstrated to decrease the Brr2 RNA unwinding activity in mutagenesis studies, thus, supporting our results.[8] We also explored optimal and suboptimal paths along NcTOC and TOC, choosing as source residues in NSec63 the most conserved highly-degenerated residues over different selected paths from Jab1 to the respective ATP site of NC and CC (Figure S13-S14)(see Supporting Result section of SI). To identify the communication-routes

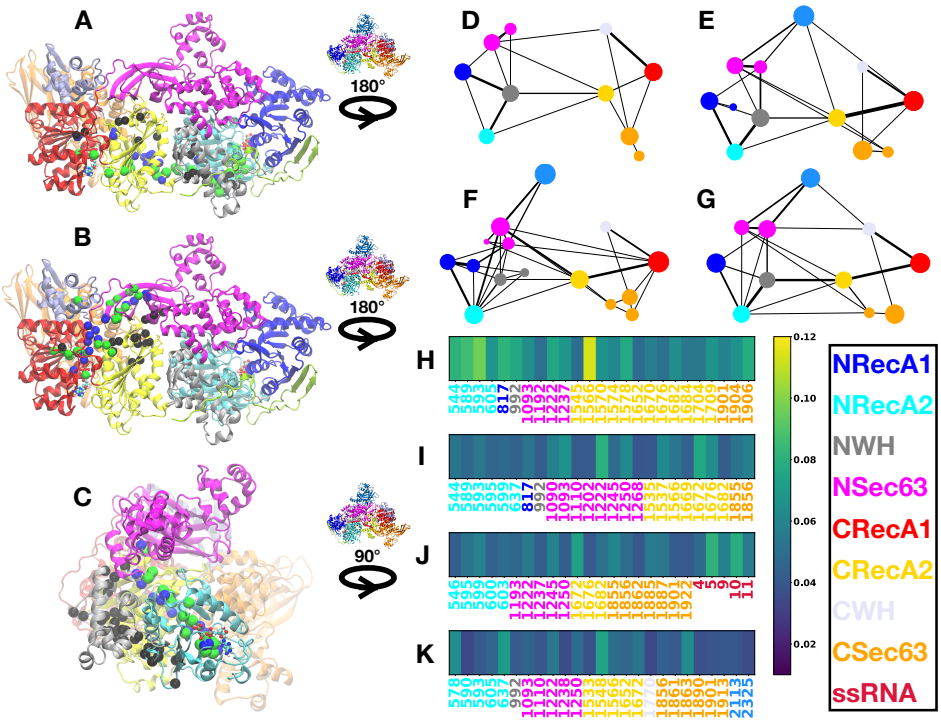


Figure 4: **(A)-(C)** Signaling-paths underlying Brr2 essential dynamics. Paths are traced by considering the residues having node-degeneracy larger than 0.20. The cross-communication paths underlying the LOC, TOC and NcTOC functional motions are depicted in **(A)**, **(B)**, **(C)**, respectively. In **(A)**, **(B)** and **(C)** a back and side-view of Brr2 are shown, respectively. The paths belonging to $Brr2_{ATP}$, $^{Jab1}Brr2_{ATP}$ and $^{U4/Jab1}Brr2_{ATP}$ are shown as green, blue and black spheres, respectively. **(D)-(G)** Protein Community Graph for (D) $Brr2_{ATP}$ (E) $^{Jab1}Brr2_{ATP}$ (F) $^{U4/Jab1}Brr2_{ATP}$ and (G) $^{Jab1}Brr2_{ADP}$. Circles represent communities, whose size accounts for the number of residues belonging to the community, and whose edge-thickness accounts for communication-flow strength occurring between the communities. The communities have the same color of the Brr2 domain to which have largest similarity score . **(H)-(K)** Vertex Betweeness of the first 25 residues of (H) $Brr2_{ATP}$ (I) $^{Jab1}Brr2_{ATP}$ (J) $^{U4/Jab1}Brr2_{ATP}$ and (K) $^{Jab1}Brr2_{ADP}$.

underlying the TOC (PC2) motion we, instead, traced the signalling-paths between Pro1267@NSec63 (source-residue) and Lys1356@CRecA1 (sink-residue). Here, the routes convey through NSec63, CRecA1, CRecA2 and CWH domains (Figure 4B) and, again, U4 snRNA binding redirects these paths. Indeed, NSec63 interacts mainly with CRecA1 through the link-loop when U4 snRNA is absent. However, upon U4 snRNA binding, this route is diverted (Figure 4B) due to an increased number of contacts between residues of CRecA2 and NSec63 (Figure S12), as revealed by the node-degeneracy matrix (Figure 4B), and to an increased number of hydrogen-bonds (Table S10-S13). Among the highly-degenerated nodes of this signaling-path are Ile1193@NSec63 and

Thr1194@NSec63 (Figure S11B), which lie in the vicinity of Arg1195@NSec63, a residue implicated by mutagenesis studies as key in modulating Brr2 helicase activity.[8] Finally, we analysed the signaling-paths underlying the NcTOC (PC3) motion by selecting Lys509@NRecA1 (source-residue) and Glu1097@NSec63 (sink-residue) (Figure 4C). We found that in all ATP-containing systems the communication-routes travel on NRecA1, but, when snRNA or ADP are present, the routes divert on NWH (Figure S11C). Hence, in the ADP-containing system NSec63 is not directly connected to NRecA1 and its signaling to the catalytic site travels a longer path. Remarkably, besides identifying nodes that have been already pinpointed by mutagenesis studies, this analysis also supplies information for future experimental studies as it identifies experimentally unexplored residues (such as Pro1192-Thr1194, Lys1199-Val1200) as high degeneracy nodes. Mutagenesis studies of these residues can further support to our findings.

Ultimately, to gather a more complete picture of the Brr2 allosteric-signaling, we also performed Community Network Analysis (CNA). This allows identifying groups of residues ("communities", i.e. cohesive structural units with synchronized dynamics), which establish a dynamic cross-talk with each other, and allows computing the strength of correlation between them using the 'edge-betweenness' (EB) measure. Differently from optimal/suboptimal paths identified before, CNA is a coarse graining procedure of the protein and the resulting community-graph allows to understand how a protein responds to the stimulus of different effectors/regulators.[32]

The CNA of Brr2 reveals a similar size of the communities in all systems, but the number of the communnities and the thickness of their connecting edges change in the different models. In short, Jab1 binding splits NRecA2 in two communities, while reinforcing a communication short-cut via NSec63 and CRecA2, thus enhancing their cross-talk. The communities fragmentation further raises upon U4 snRNA binding, which, nevertheless, increases the communication strength between NSec63, NRecA2 and CRecA2. As such, RNA binding reinforces the allosteric cross-talk between distal Brr2 functional sites, similarly to what observed in other spliceosome proteins.[26,33] Hence, upon RNA/Jab1 binding the intra-cassette communication is weakened, but a stronger communication-channel is established between N- and C-cassette (Figure 4D-G). In all systems the links to the community located on CSec63 are weak, indicating this domain is not strongly linked to the rest of Brr2 and, therefore, that this cassette may act as landing-pad mediating the binding of other spliceosome proteins, which regulate Brr2 function.

We finally calculated the Vertex-Betweeness (VB) for each residue, defined as the percentage of shortest (i.e. most-relevant) signaling-paths traveling through that residue, regardless of the sink and source residues. Hence, high VB score allows identifying residues that act as bridge

in the communication-flow between any pair of nodes of the network (see Supporting Information for details) (Figure 4H-K). This information is complementary to NWA, which instead privileges continuous optimal/suboptimal paths between a user-selected source and sink residue, even when they cross low-VB-score residues. In the VB analysis we observed that most of the high-VB residues identified by CNA are placed on NSec63, CRecA2 and NRecA2. Among them are His1548@CRecA2, Arg637@NRecA2, Arg603@NRecA2, similarly to what observed from the node-degeneracy of NWA detailed above, and Arg1090@NSec63. These residues were found to be critical for Brr2 function by mutagenesis studies and to be involved in Retinitis Pigmentosa.[17] These findings support the functional significance of these hot-spots and further underline the relevance of functional movements and the cross-talk of the two cassette for Brr2 function.

Collectively, our findings offer a fundamental advance in understanding the molecular terms of Brr2 mechanism and its control by allosteric effectors/regulators. We reveal that the C-cassette alters the intrinsic essential dynamics of the isolated N-cassette and that the two cassettes cooperatively act to promote RNA binding and translocation. Surprisingly, U4 snRNA exerts the largest impact on conformational dynamics and allosteric cross-talk of Brr2, by creating a stronger communication-channels between the two cassettes. This knowledge broadens our understanding of the subtle regulatory mechanisms underlying eukaryotic splicing, disclosing that even RNA molecules can act as allosteric regulators of spliceosomal proteins to control the time, amount, degree, and rate of gene expression. We also supply information for future mutagenesis studies to establish their relevance in allosteric signaling in Brr2 thus further supporting and complementing results. We remark that this study does not allow us to formulate hypothesis on Brr2 translocation mechanism due to the lack of structural information on the snRNA binding to the C-cassette.[27] This lack of knowledge also prevents to unravel the molecular basis of RNA-driven allosteric regulation of Brr2 ATPase activity and compare it with that observed in other DExH helicases.[16] RNA-dependent helicases, which play key roles in distinct aspects of RNA metabolism. Allosteric regulation is widespread phenomenon in RNA-dependent NTPase/helicases. The architectural sophistication of Brr2 makes it suitable to distinct and subtle intra- and inter-molecular allosteric regulatory mechanisms, which may be operative in related dual-cassette helicases involved in many other key cellular processes.

ASSOCIATED CONTENT

Supporting Information. The Supporting Information is available free of charge. Details of the computational methods and analysis are reported: RMSD, radius of gyration, cross-correlation matrices, structural data, inter-domain energy, figures and movies of the principal movements along PCs. Detailed Hydrogen-bond, contact analysis and tables

of degeneracy of the allosteric paths are reported. SI Movie 1 (S1.mp4), SI Movie 2 (S2.mp4), SI Movie 3 (S3.mp4), SI Movie 4 (S4.mp4), SI Movie 5 (S5.mp4), SI Movie 6 (S6.mp4)

AUTHOR INFORMATION

Corresponding Author
Alessandra Magistrato - CNR-IOM at SISSA, 34136 Trieste, Italy;
orcid.org/0000-0002-2003-1985;
Email: alessandra.magistrato@sissa.it
Author
Francesco Guidarelli Mattioli - Sapienza University of Rome, Piazzale Aldo Moro 2, 00185 Rome, Italy;
orcid.org/0000-0003-3216-9102
Andrea Saltalamacchia- nternational School for Advanced Studies (SISSA/ISAS), via Bonomea 265, 34136 Trieste, Italy
orcid.org/0000-0003-1174-9271

Acknowledgments

AM thanks the Italian Association for Cancer Research (project AIRC IG 24514) and PNRR: National Center for Gene Therapy and Drugs based on RNA Technology CUP $B83C22002860006CN_0000004$. F.G.M thanks CINECA grant ISCRAB NNPROT.

References

(1) Ozgur, S.; Buchwald, G.; Falk, S.; Chakrabarti, S.; Prabu, J. R.; Conti, E. The conformational plasticity of eukaryotic RNA-dependent ATP ases. *The FEBS journal* **2015**, *282*, 850–863.

(2) Tholen, J.; Galej, W. Structural Studies of the Spliceosome: Bridging the Gaps. *Curr. Opin. Struct. Biol.* **2022**, *77*, 102461.

(3) Borišek, J.; Aupič, J.; Magistrato, A. Establishing the Catalytic and Regulatory Mechanism of RNA-based Machineries. *WIREs Comput Mol Sci* **2023**, *13*, e1643.

(4) Borišek, J.; Saltalamacchia, A.; Gallì, 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) Borisek, J.; Casalino, L.; Saltalamacchia, A.; Mays, S. G.; Malcovati, L.; Magistrato, A. Atomic-level Mechanism of Pre-mRNA Splicing in Health and Disease. *Acc. Chem. Res.* **2020**, *54*, 144–154.

(6) Borisek, J.; Magistrato, A. All-atom simulations decrypt the molecular terms of RNA

catalysis in the exon-ligation step of the spliceosome. *ACS Catalysis* **2020**, *10*, 5328–5334.

(7) Absmeier, E.; Santos, K. F.; Wahl, M. C. Functions and Regulation of the Brr2 RNA Helicase During Splicing. *Cell Cycle* **2016**, *15*, 3362–3377.

(8) Santos, K. F.; Jovin, S. M.; Weber, G.; Pena, V.; Lührmann, R.; Wahl, M. C. Structural Basis for Functional Cooperation between Tandem Helicase Cassettes in Brr2-mediated Remodeling of the Spliceosome. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 17418–17423.

(9) De Bortoli, F.; Espinosa, S.; Zhao, R. DEAH-box RNA Helicases in Pre-mRNA Splicing. *Trends Biochem. Sci.* **2021**, *46*, 225–238.

(10) Nussinov, R.; Tsai, C.-J. Allostery in Disease and in Drug Discovery. *Cell* **2013**, *153*, 293–305.

(11) Tsai, C.-J.; Nussinov, R. A Unified View of "How Allostery Works". *PLoS Comput. Biol.* **2014**, *10*, e1003394.

(12) Nussinov, R.; Tsai, C.-J.; Liu, J. Principles of Allosteric Interactions in Cell Signaling. *J. Am. Chem. Soc.* **2014**, *136*, 17692–17701.

(13) Absmeier, E.; Vester, K.; Ghane, T.; Burakovskiy, D.; Milon, P.; Imhof, P.; Rodnina, M. V.; Santos, K. F.; Wahl, M. C. Long-range Allostery Mediates Cooperative Adenine Nucleotide Binding by the Ski2-like RNA Helicase Brr2. *J. Biol. Chem.* **2021**, *297*.

(14) 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.

(15) Mozaffari-Jovin, S.; Wandersleben, T.; Santos, K. F.; Will, C. L.; Lührmann, R.; Wahl, M. C. Inhibition of RNA Helicase Brr2 by the C-terminal Tail of the Spliceosomal Protein Prp8. *Science* **2013**, *341*, 80–84.

(16) Movilla, S.; Roca, M.; Moliner, V.; Magistrato, A. Molecular Basis of RNA-Driven ATP Hydrolysis in DExH-Box Helicases. *J. Am. Chem. Soc.* **2023**, *145*, 6691–6701.

(17) Cvačková, Z.; Matěj, D.; Staněk, D. Retinitis Pigmentosa Mutations of SNRNP 200 Enhance Cryptic Splice-Site Recognition. *Hum. Mutat.* **2014**, *35*, 308–317.

(18) Zhang, L.; Shen, J.; Guarnieri, M. T.; Heroux, A.; Yang, K.; Zhao, R. Crystal Structure of the C-terminal Domain of Splicing Factor Prp8 Carrying Retinitis Pigmentosa Mutants. *Protein Sci.* **2007**, *16*, 1024–1031.

(19) Rozza, R.; Janoš, P.; Spinello, A.; Magistrato, A. Role of computational and structural biology in the development of small-molecule modulators of the spliceosome. *Expert Opin. Drug Discovery* **2022**, *17*, 1095–1109.

(20) Eargle, J.; Luthey-Schulten, Z. NetworkView: 3D display and analysis of protein· RNA interaction networks. *Bioinformatics* **2012**, *28*, 3000–3001.

(21) Saltalamacchia, A.; Casalino, L.; Borisek, J.; Batista, V. S.; Rivalta, I.; Magistrato, A. Decrypting the information exchange pathways across the spliceosome machinery. *Journal of the American Chemical Society* **2020**, *142*, 8403–8411.

(22) Molina Vargas, A. M.; Osborn, R. M.; Sinha, S.; Arantes, P. R.; Patel, A.; Dewhurst, S.; Palermo, G.; O'Connell, M. R. New design strategies for ultra-specific CRISPR-Cas13a-based RNA-diagnostic tools with single-nucleotide mismatch sensitivity. *bioRxiv* **2023**, 2023–07.

(23) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

(24) Banás, P.; Hollas, D.; Zgarbová, M.; Jurecka, P.; Orozco, M.; Cheatham III, T. E.; Sponer, J.; Otyepka, M. Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins. *J. Chem. Theory Comput.* **2010**, *6*, 3836–3849.

(25) Zgarbová, M.; Otyepka, M.; Sponer, J.; Mladek, A.; Banas, P.; Cheatham III, T. E.; Jurecka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.* **2011**, *7*, 2886–2902.

(26) Rozza, R.; Saltalamacchia, A.; Orrico, C.; Janos, P.; Magistrato, A. All-atom simulations elucidate the impact of U2AF2 cancer-associated mutations on pre-mRNA recognition. *Journal of Chemical Information and Modeling* **2022**, *62*, 6691–6703.

(27) Agrò, S. N.; Rozza, R.; Movilla, S.; Aupic, J.; Magistrato, A. Molecular Dynamics Simulations Elucidate the Molecular Basis of Pre-mRNA Translocation by the Prp2 Spliceosomal Helicase. *J. Chem. Inf. Model.* **2023**,

(28) Sinha, S.; Molina Vargas, A. M.; Arantes, P. R.; Patel, A.; O'Connell, M. R.; Palermo, G. Unveiling the RNA-mediated Allosteric Activation Discloses Functional Hotspots in CRISPR-Cas13a. *Nucleic Acids Research* **2023**, gkad1127.

(29) Sethi, A.; Eargle, J.; Black, A. A.; Luthey-Schulten, Z. Dynamical Networks in tRNA: Protein Complexes. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6620–6625.

(30) 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.

(31) Van Wart, A. T.; Durrant, J.; Votapka, L.; Amaro, R. E. Weighted Implementation of Suboptimal Paths (WISP): an Optimized Algorithm and Tool for Dynamical Network Analysis. *J. Chem. Theory Comput.* **2014**, *10*, 511–517.

(32) Newman, M. E.; Girvan, M. Finding and Evaluating Community Structure in Networks. *Physical review E* **2004**, *69*, 026113.

(33) Spinello, A.; Janos, P.; Rozza, R.; Magistrato, A. Cancer-Related Mutations Alter RNA-Driven Functional Cross-Talk Underlying Premature-Messenger RNA Recognition by Splicing Factor SF3b. *J. Phys. Chem. Lett.* **2023**, *14*, 6263–6269.