Supplementary Figures

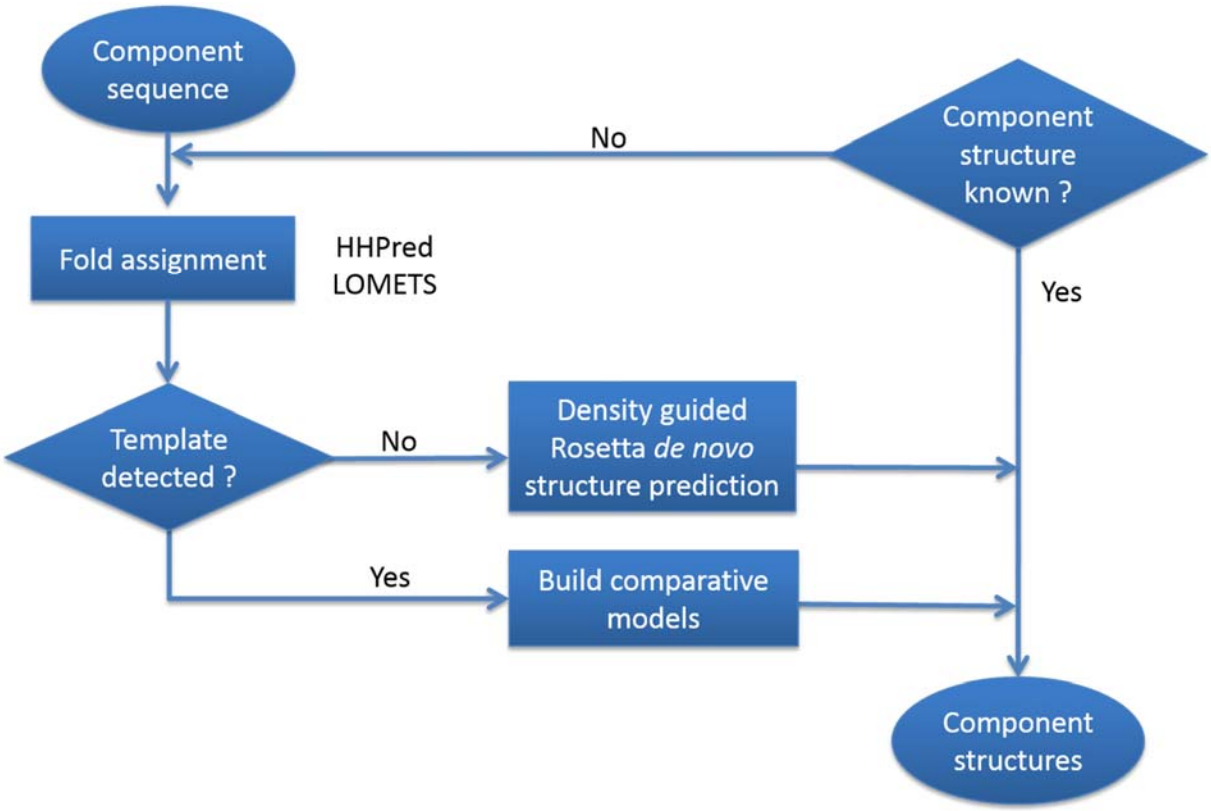


Figure S1. Hierarchical modeling protocol for the SF3b components.

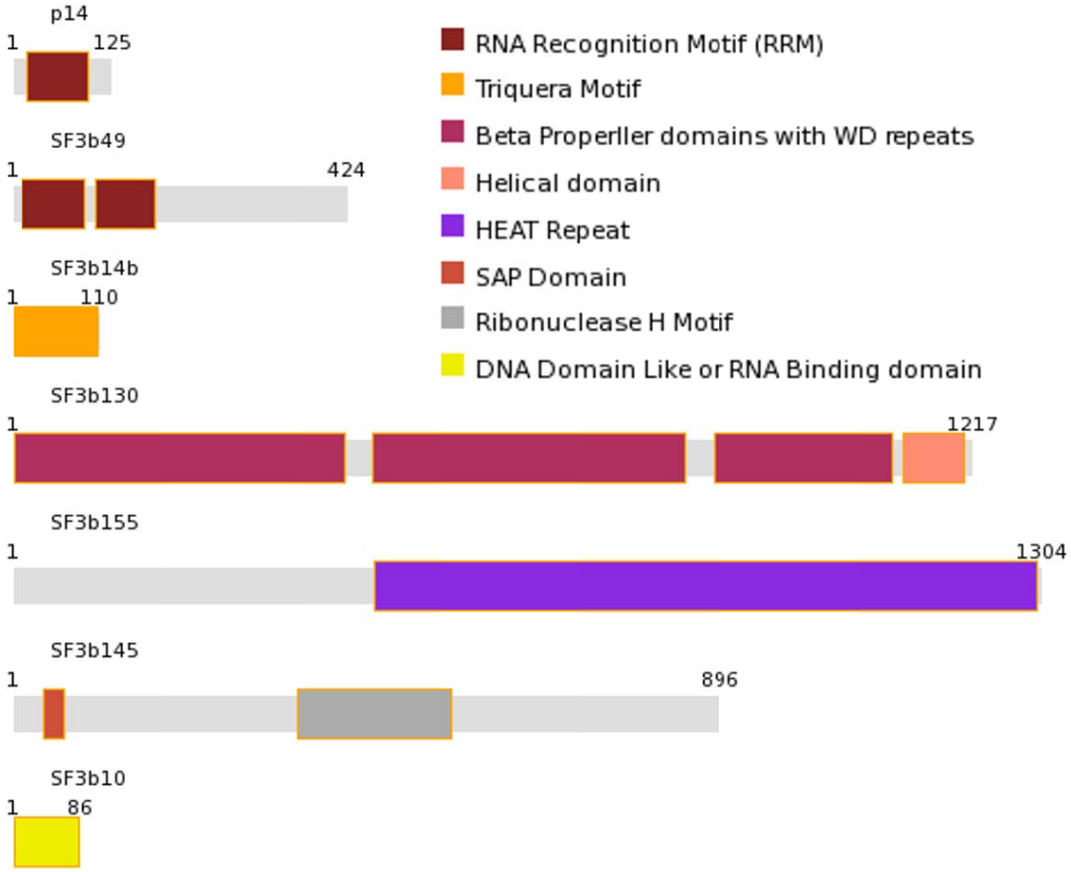


Figure S2. Domain architectures of seven components of SF3b complex. The domains assigned to each SF3b component based on the Integrative Structure Modeling protocol. The figure was generated using DomainDraw

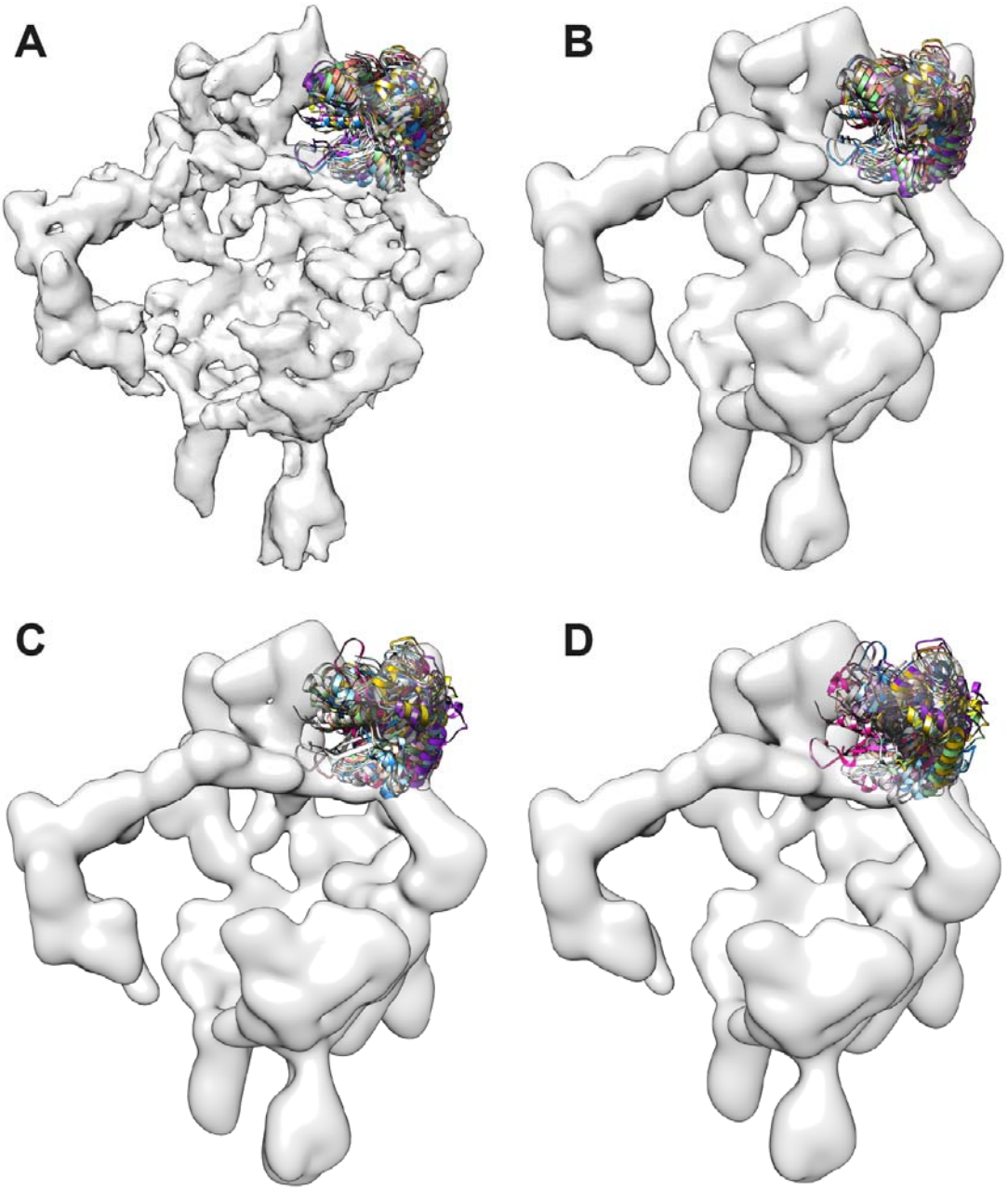


Figure S3. Location of the SF3b145 RNaseH like domain in SF3b density obtained from a global search. We performed a global search for the RNase-H like domain of SF3b145 in SF3b cryo-EM density (lacking densities for SF3b49, p14 and SF3b155) using a fine angular sampling of 3° using the *colores* program in SITUS package. The location of the top 10 solutions based on cross-correlation scores are shown here. Across the different resolutions **(A)** 9.7Å, **(B)** 12Å, **(C)** 15Å and **(D)** 18Å we obtained similar results. SF3b145 is also known to interact with both SF3b49 and SF3b155 based on the SF3b protein-protein interaction network validating our localization.

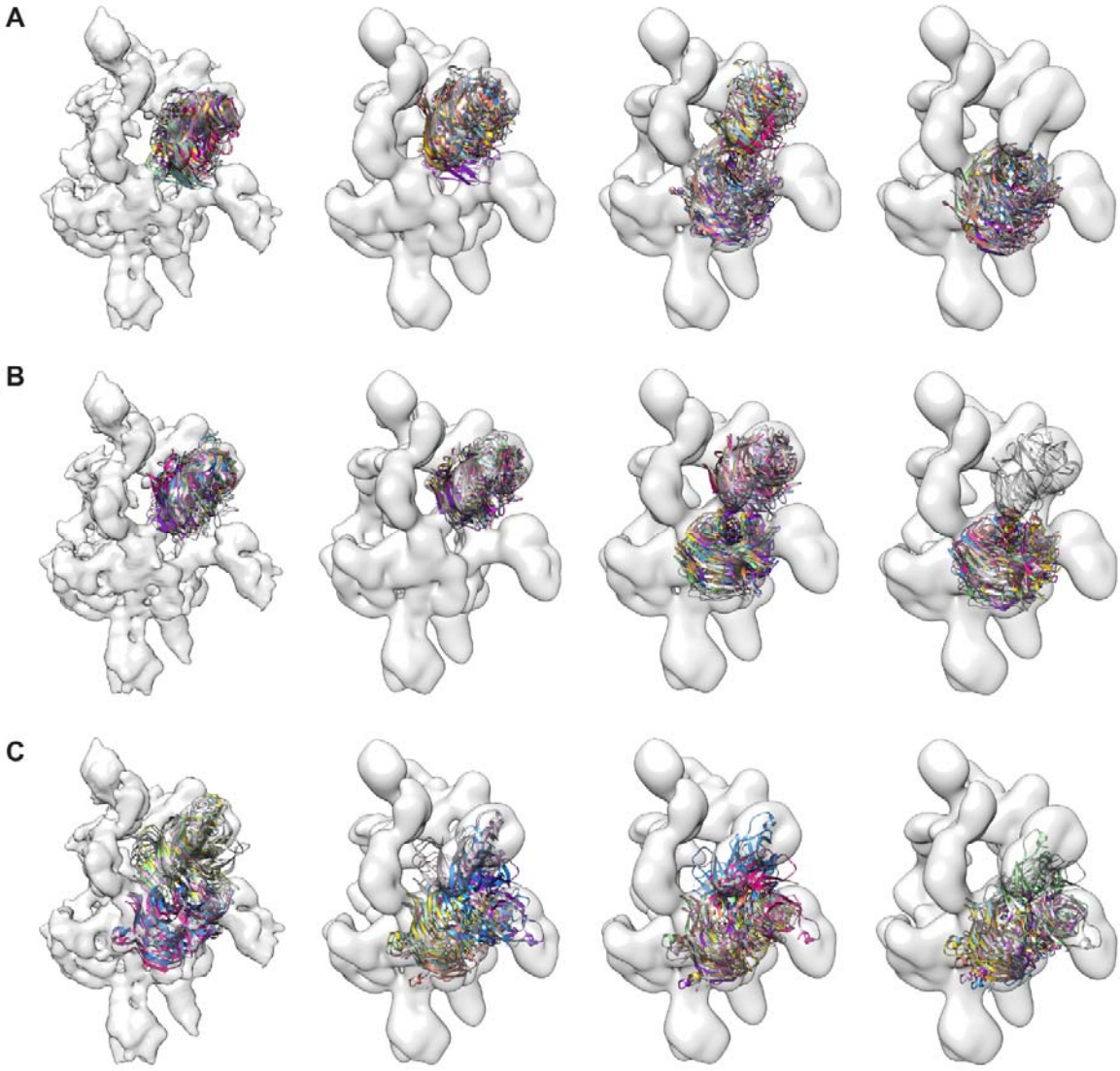


Figure S4. Location of the SF3b130 β-propeller domains in SF3b density obtained from a global search. We performed a global search for the β-propeller domains of SF3b130 in SF3b cryo-EM density (lacking densities for SF3b145, SF3b49, p14 and SF3b155) using a fine angular sampling of 3° using the *colores* program in SITUS package across the different resolutions (9.7Å, 12Å, 15Å, 18Å - left to right in each row). The location of the top 10 solutions based on cross-correlation scores are shown here. In the case of **(A)** SF3b130 β-propeller1 and **(B)** β-propeller2 domains all the top 10 solutions clustered around a flat density region searches using a 9.7 and 12Å filtered map. In the low resolution (15 and 18Å) searches, most of the solutions clustered near the flat density region. In the case of **(C)** β-propeller3 as the resolutions became worse most of the top 10 solutions started clustering away from the flat density region but still clustered to a location nearby. SF3b130 is known to interact with both SF3b145 and SF3b155, hence the density region around the top solution clusters could be reliably assigned to SF3b130.

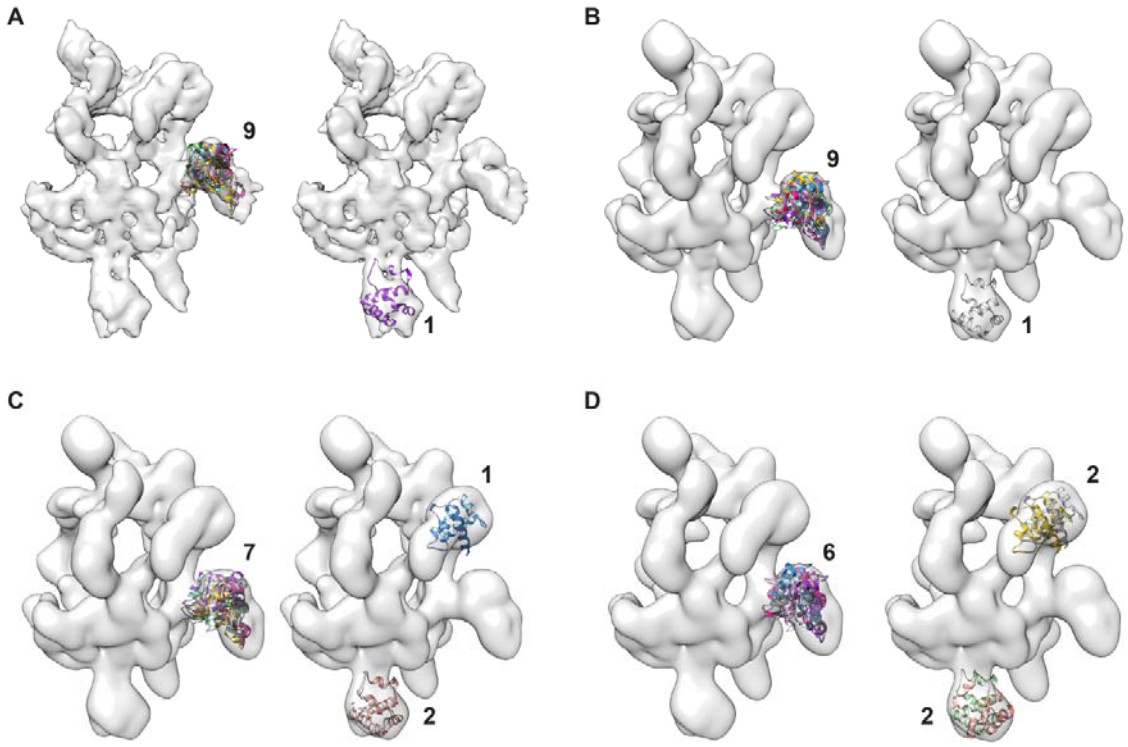


Figure S5. Location of the SF3b130 helical domains in SF3b density obtained from a global search. For SF3b130-helical domain most of the top 10 solutions clustered around a density segment which is also near the flat density region where the other two β-propeller domains clustered. Only the number of solutions clustered around this density segment changed across the different resolutions **(A)** 9.7Å, **(B)** 12Å, **(C)** 15Å and **(D)** 18Å. SF3b130 is known to interact with both SF3b145 and SF3b155 (Figure 1D and Table S2), hence the density region around the top solution clusters could be reliably assigned to SF3b130. The left side figure in panel shows the density location for helical domain where maximum number of solutions (numbers indicated) clustered and the alternate locations are shown on the right.

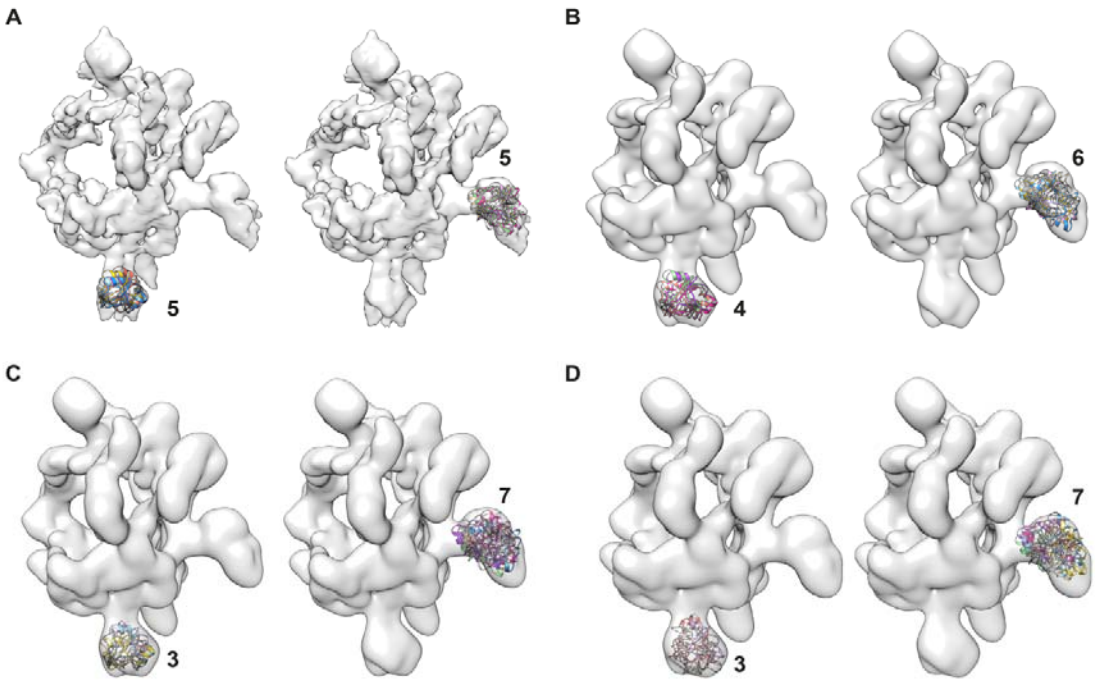


Figure S6. Location of the SF3b14b in SF3b density obtained from a global search. Searches for the location to fit SF3b14b resulted in most of the top 10 solutions clustering in locations proximal to both SF3b130 helical domain and SF3b155. Again only the number of solutions clustered around these two regions varied across the different resolutions **(A)** 9.7 Å, **(B)** 12 Å, **(C)** 15 Å and **(D)** 18 Å. The solutions for this protein in the SF3b130 helical region density segment can be treated as a false positive as it is known from the SF3b interaction network that SF3b155 and SF3b14b interact (Figure1D and Table S2). Therefore although there are more than one location obtained by fitting, only one of them was identified to be correct based on available experimental information. The left side figure in panel shows the solutions (numbers indicated) clustered in the density location for SF3b14b and the alternate locations are shown on the right.

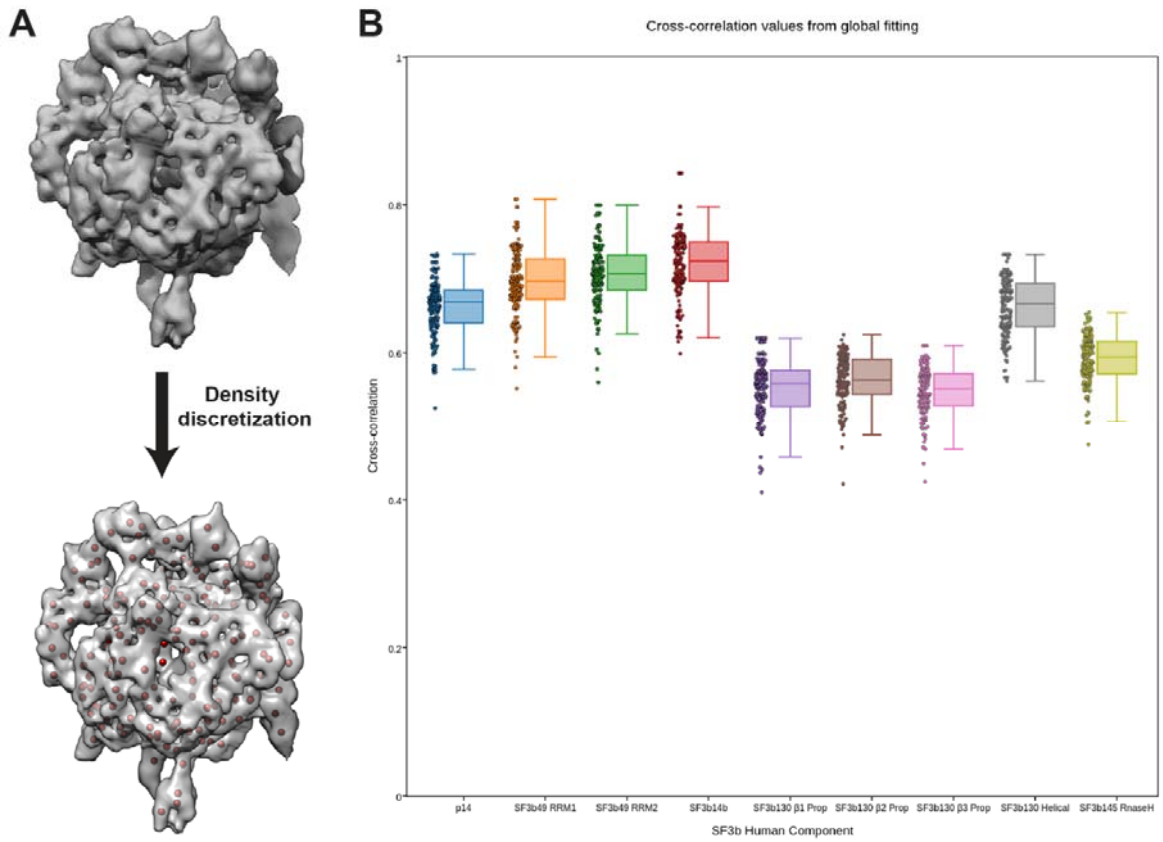


Figure S7. Density discretization and distribution of cross-correlation values from the global fitting. (A) Discretization of the SF3b density map into a set of 166 points (red spheres) and this was performed using a modified density segmentation method (see Supplementary materials and methods). **(B)** The box plot shows the cross-correlation values obtained based on the global fitting of the SF3b components using the discretized density map using UCSF Chimera.

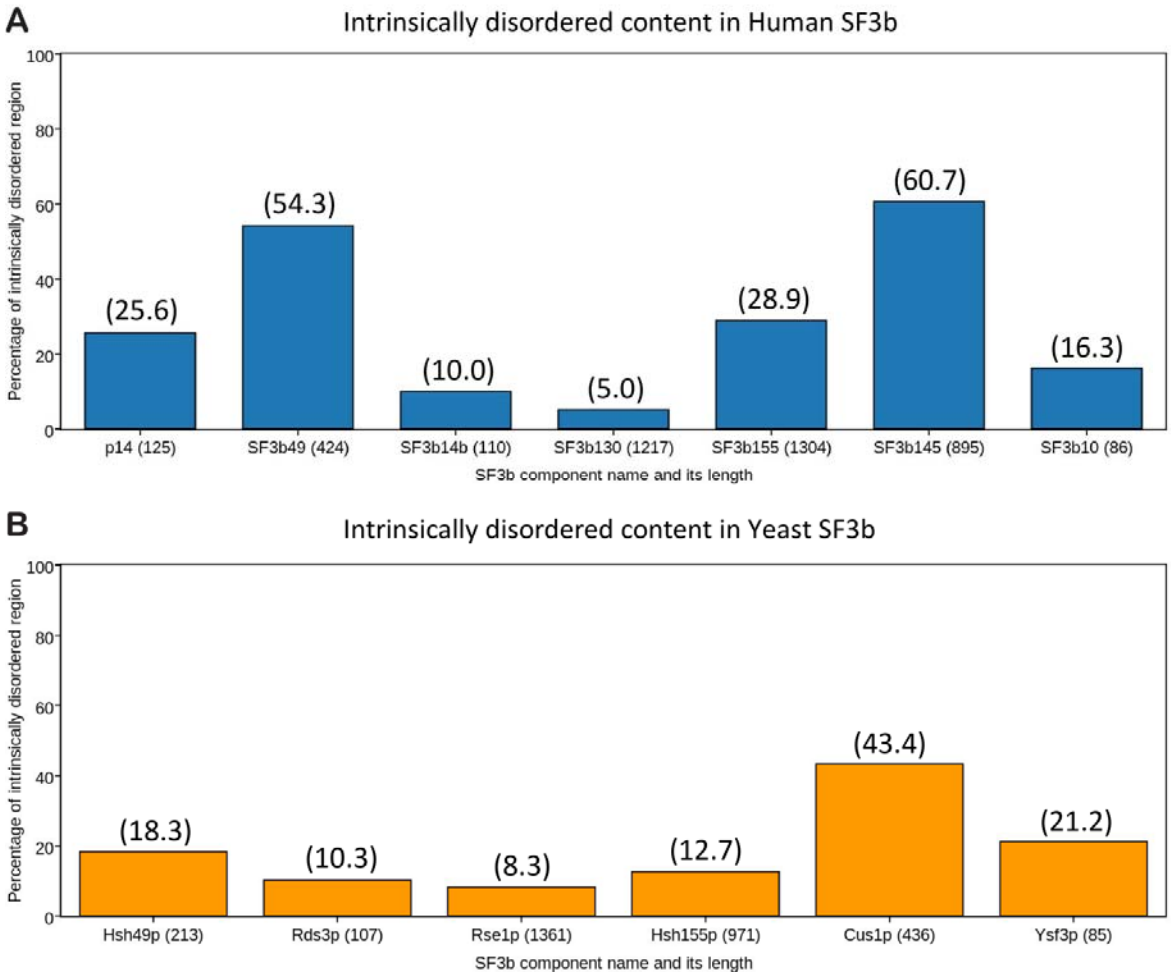


Figure S8. Intrinsically disordered content in SF3b complex. The intrinsically disordered regions were predicted using a metapredictor "metaPrDos" in **(A)** Human and **(B)** Yeast**.** The percentages were calculated for each protein relative to their total length provided in the brackets. Moreover the total number of residues in human SF3b (4161 amino acids) is substantially higher when compared to yeast SF3b (3173) both due to the lack of p14 orthologue and lesser intrinsically disordered regions in yeast.

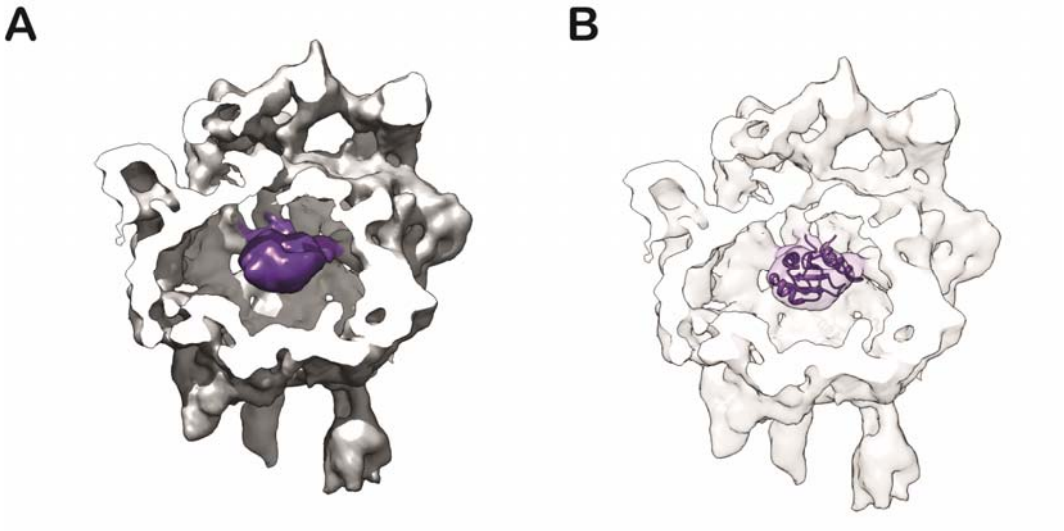


Figure S9. p14 modeling. (A) p14 localization was obtained from the previous work by Golas et al (Golas et al, 2003). **(B)** The fitting of the available X-ray structure was performed by Situs followed by local refinement.

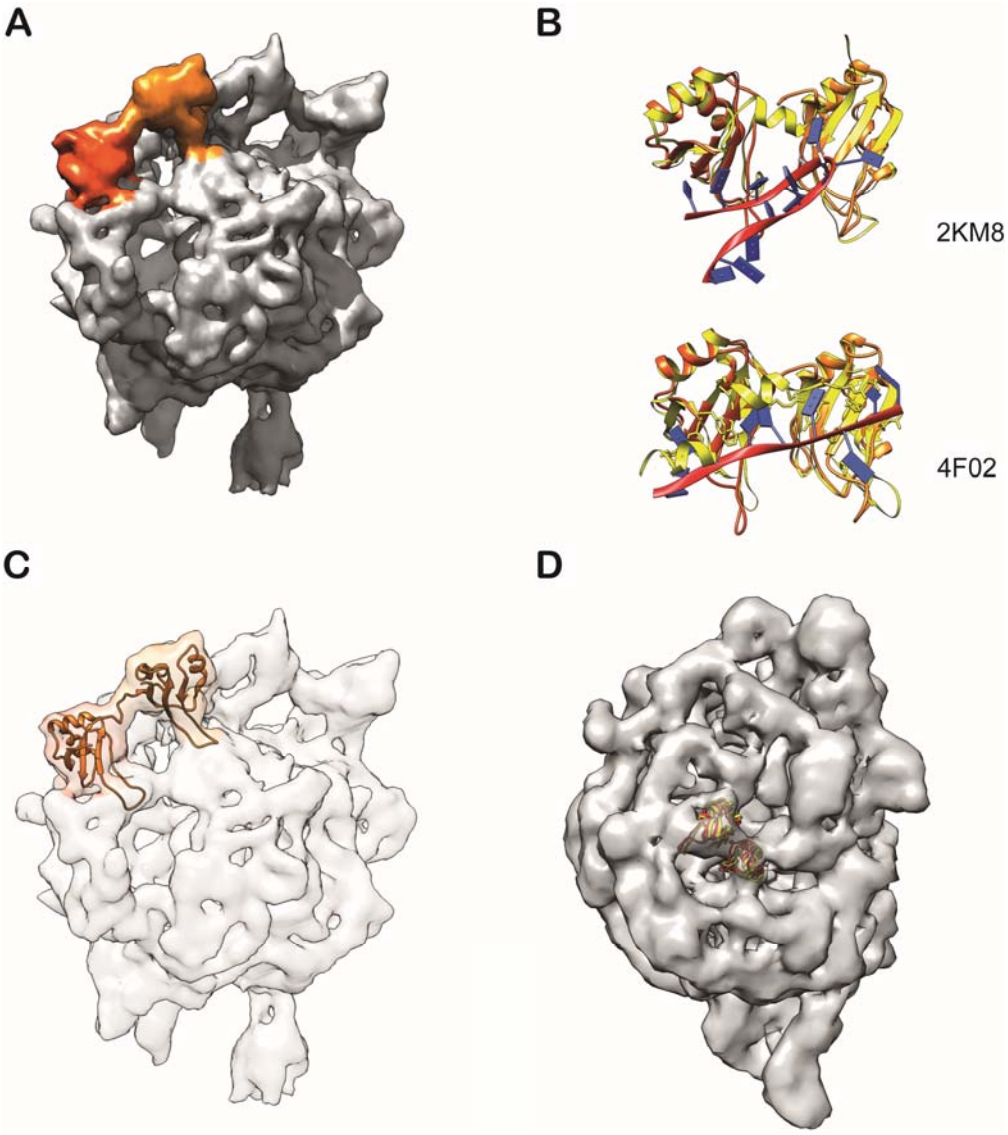


Figure S10. SF3b49 modeling. (A) The localization was obtained from the previous work by Golas et al (Golas et al, 2003) and our global fitting. **(B)** The structural alignment of RRM domains (*orange* and *orange red*) of SF3b49 with the homologues of known structures (*yellow*) to fix their relative orientation. (**C)** The fitting of the two RRM domain NMR models into the density was performed by Situs followed by local refinement and the inter-domain linker was modeled using MODELLER 9v11. **(D)** The fitting of the SF3b49 structures with their RRM domain orientation fixed according to the homologous structures 2KM8 (yellow) and 4F02 (red) in the U11/U12 cryo-EM map. SF3b49 binds to snRNA in U11/U12 di-snRNP, this also validates the orientation of SF3b49 RRM domains in SF3b closed form for RNA binding. The N-terminal RRM domain was assigned to the region near SF3b145 localized density based on the earlier interaction data showing N-terminal RRM of Hsh49 is required for interaction with Cus1. Hence, we could unambiguously assign the corresponding densities for the two SF3b49 RRMs.

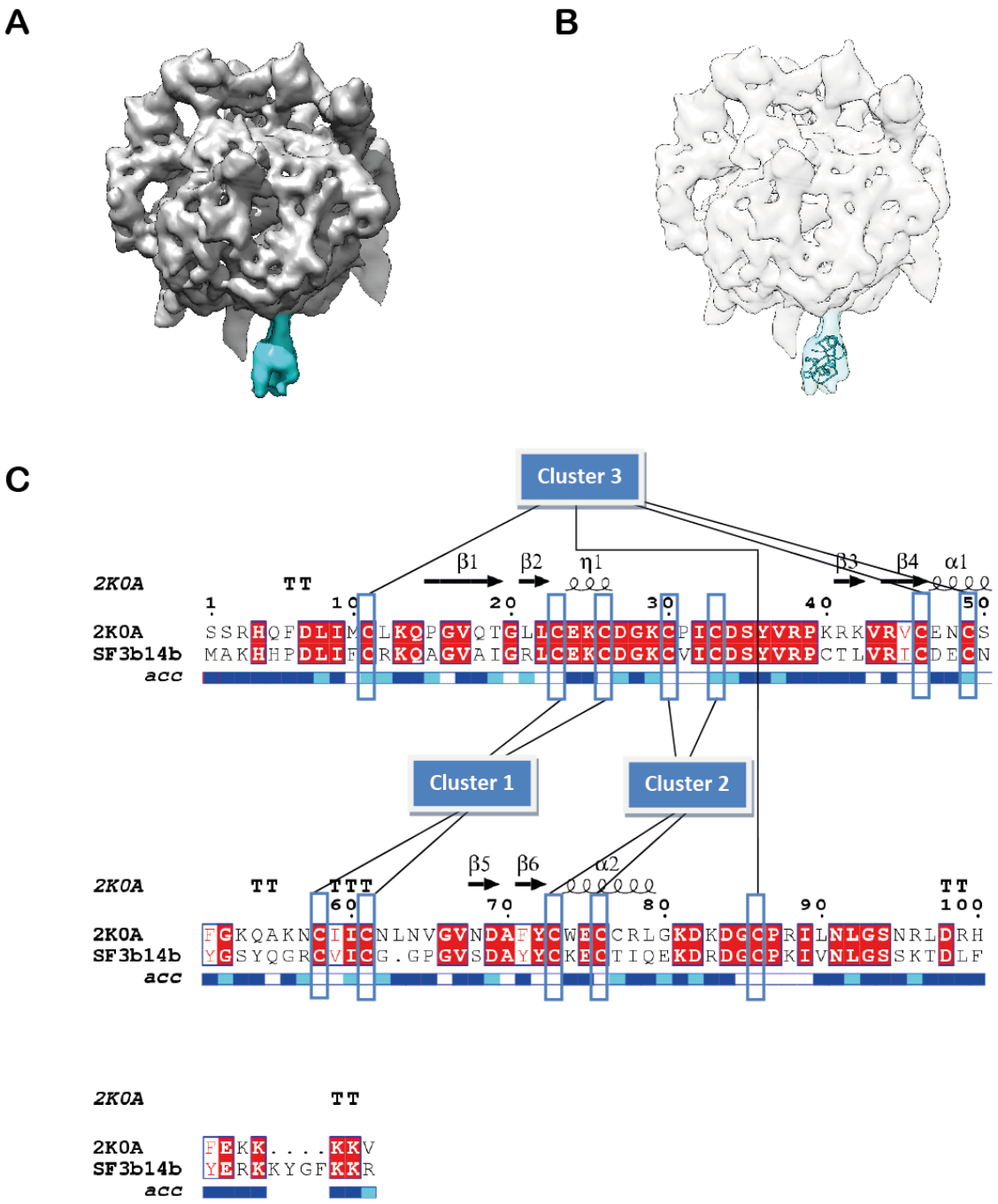


Figure S11. SF3b14b modeling. (A) The localized region for SF3b14b based on global search and interaction map **(C)** The fitting of the comparative model into the density was performed by Situs followed by local refinement using UCSF Chimera. **(D)** Sequence-structure alignment between SF3b14b and its yeast homolog Rds3p (PDB ID: 2K0A) showing the conserved zinc binding clusters represented using ESPript (Gouet et al, 1999)

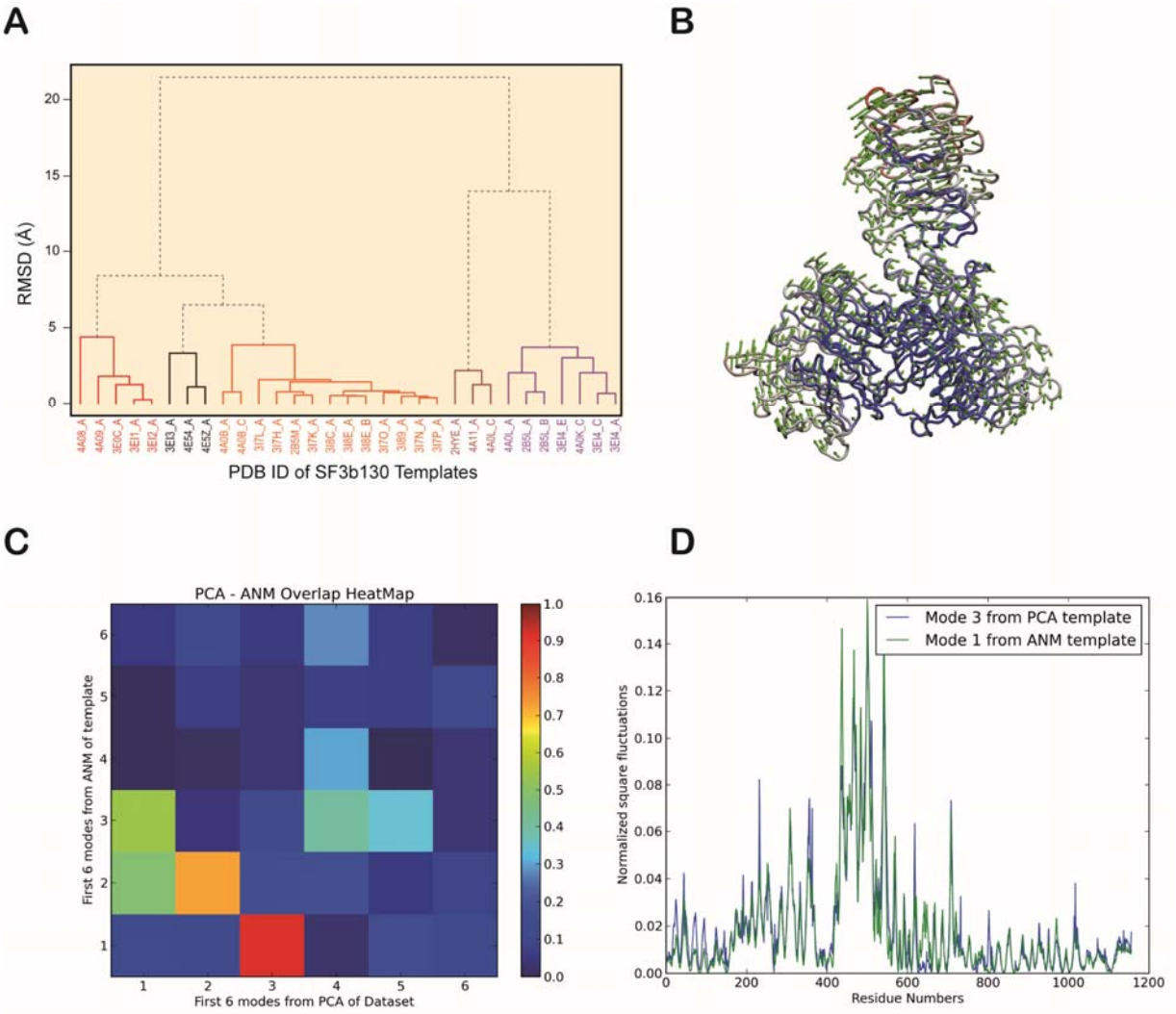


Figure S12 PCA and Normal mode analysis. (A) Structural variability of the homologue of SF3b130, DNA-damage binding protein. The dendrogram was generated based on RMSD based hierarchical clustering. **(B)** Vector field representation of the 1st mode in NMA analysis of the template showing the primary mode of motion. **(C)** The heat map showing the overlap between Principal components analysis (PCA) of SF3b130 models built on other templates and the SF3b130 based on 3EI3 template model normal mode **(D)** The normalized square fluctuations showing the high conformational flexibility of SF3b130 model especially around region 390 to 710 which correspond to the region showing high degree of rotational movement (Table S4).

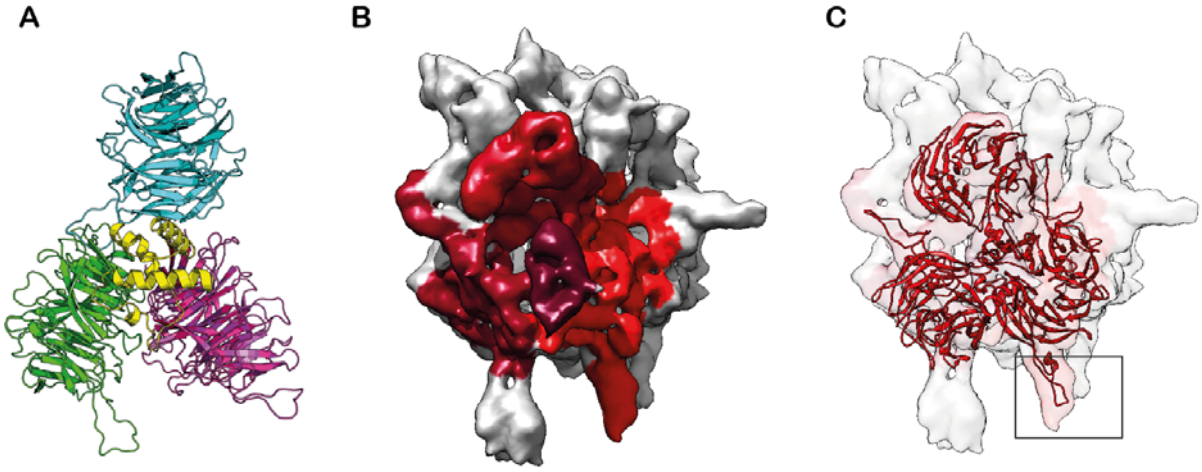


Figure S13. SF3b130 modeling and density fitting. (A) The comparative model of SF3b130. **(B)** The localized regions were segmented (four shades of red) and used for individual fitting of the domains. **(C)** The fitting of the individual domains of the comparative model of SF3b130 into the density segments. Moreover the features shown in SF3b130 structure compared to the template (box) and information from additional fitting methods (global density search) allowed unambiguous assignment of the densities for the three seven bladed WD40 β propeller domains of the protein and the helical domain.

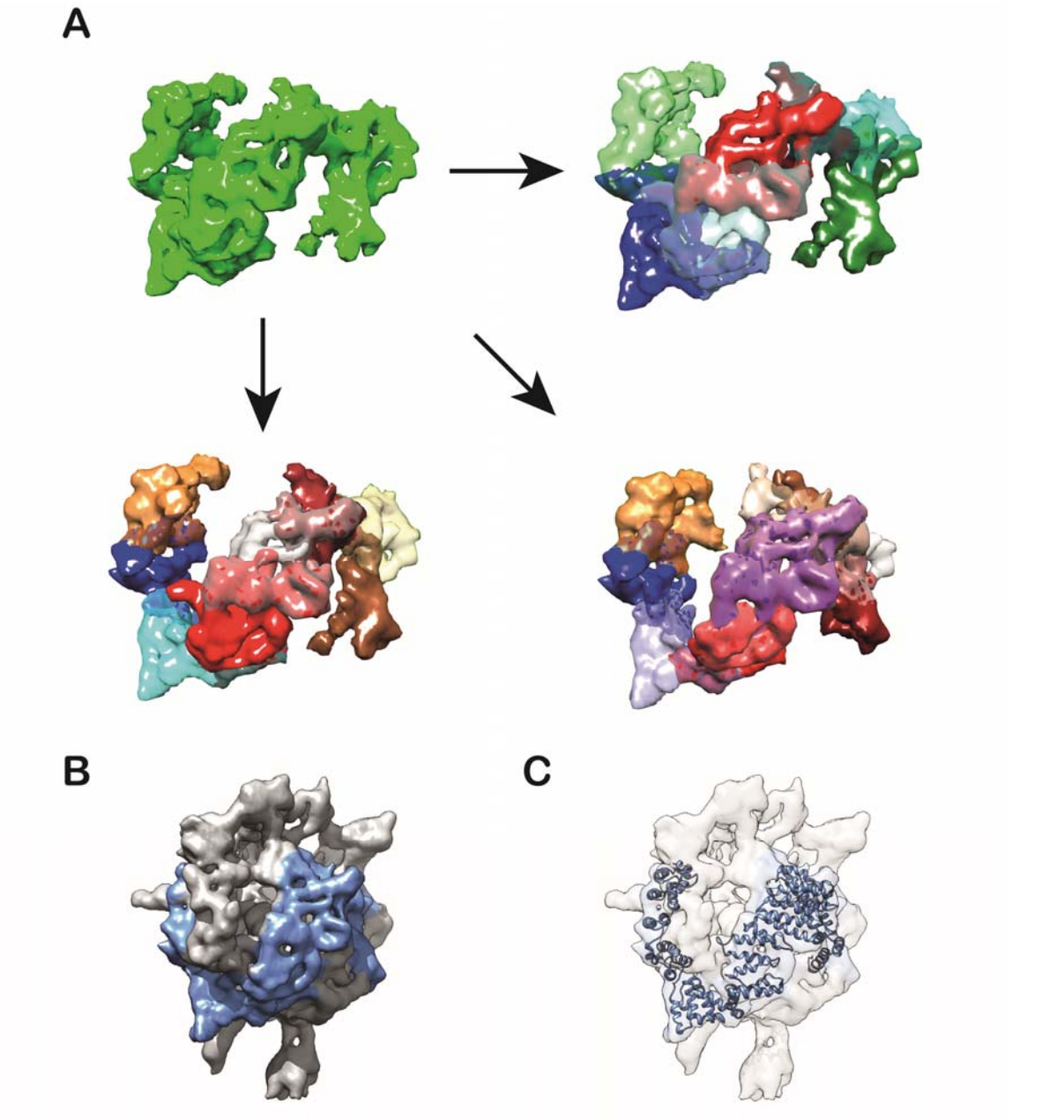


Figure S14. SF3b155 modeling and density fitting. (A) The density corresponding to SF3b155 was fragmented into overlapping sets of 6, 8 and 10 fragments. Fragmented SF3b155 structure was used for fitting into each of the overlapping density segments. The protein fragments were then used for building full length model of SF3b155. **(B)** Localized region of SF3b155 spanning substantial region of SF3b complex. **(C)** Fitted SF3b155 model.

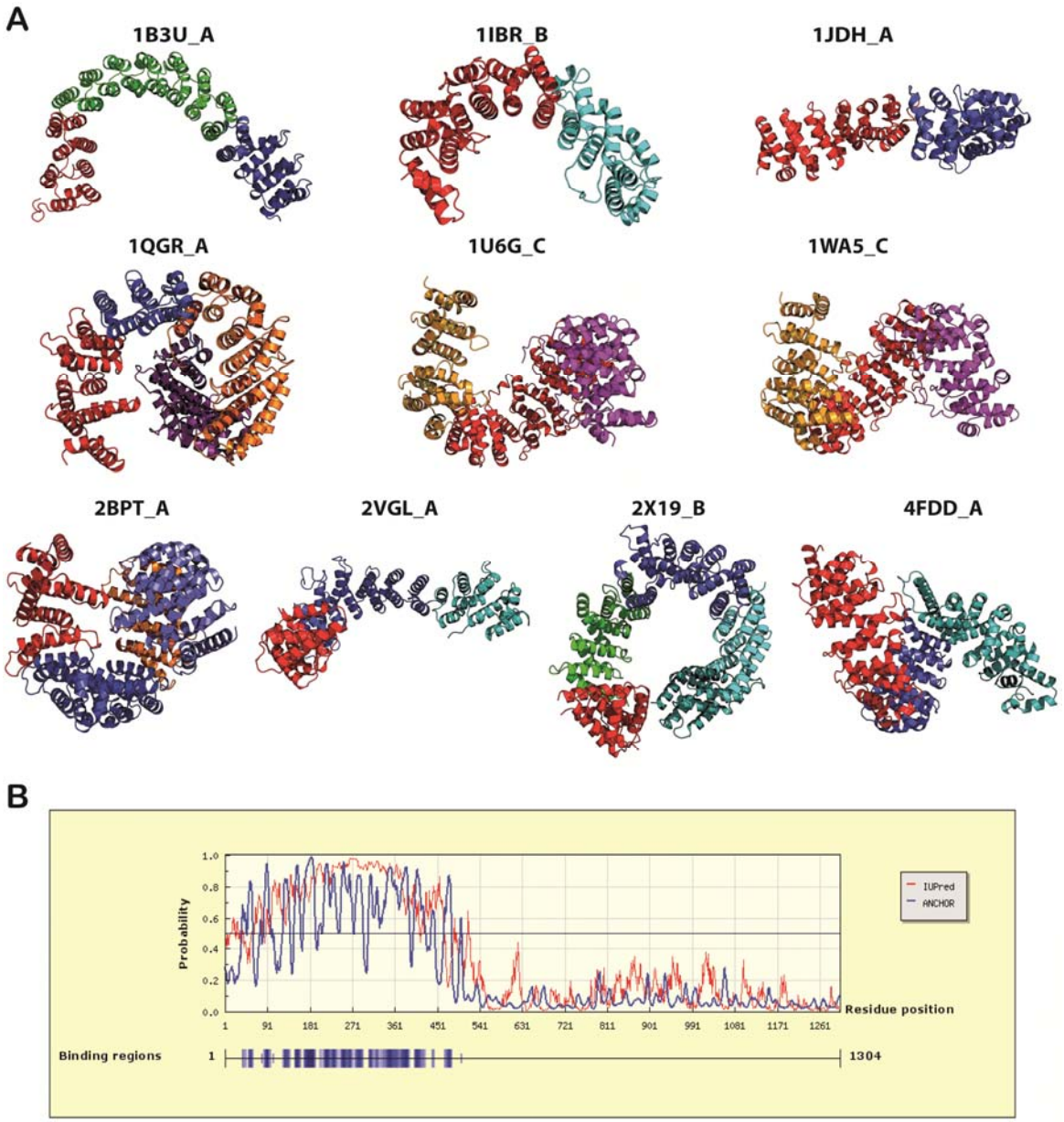


Figure S15. Templates hits with various curvatures and the disordered regions in SF3b155. **(A)** The template structures are colored according to the rigid units computed using NMA and HingeProt analysis. The rigid units are tethered by flexible inter-core regions which enable HEAT repeat proteins to adopt varying degree of curvatures. **(B)** The N-terminal intrinsically disordered region of SF3b155 predicted using "ANCHOR" server also showing the potential binding regions.

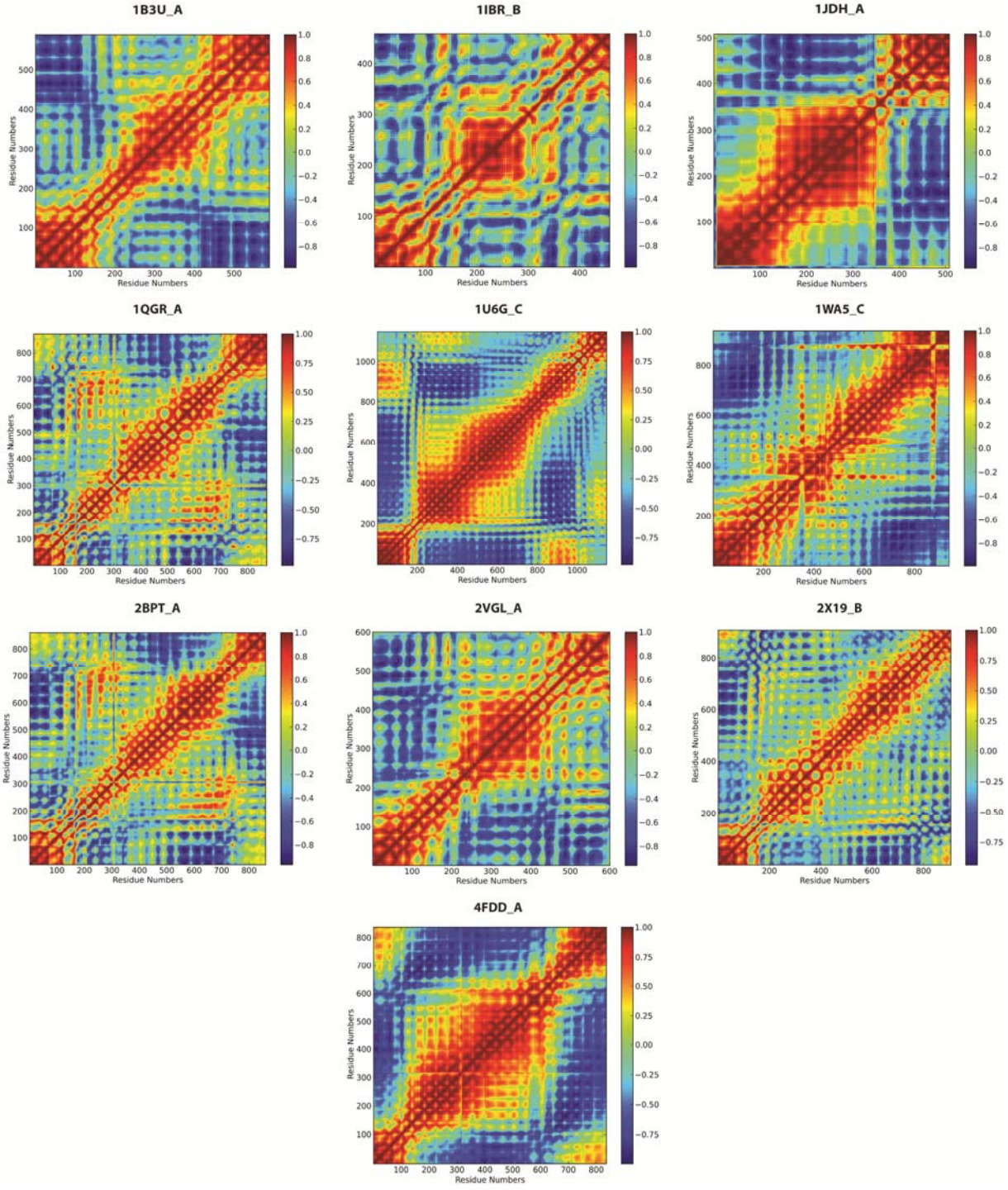


Figure S16. **Dynamic cross-correlation plots from Anisotropic Normal Mode (ANM) analysis.** The heat maps show the SF3b155 templates flexibility with the red regions showing the residues which have a correlated motion and the blue corresponding to the anti-correlated motions within the HEAT repeats of these proteins.

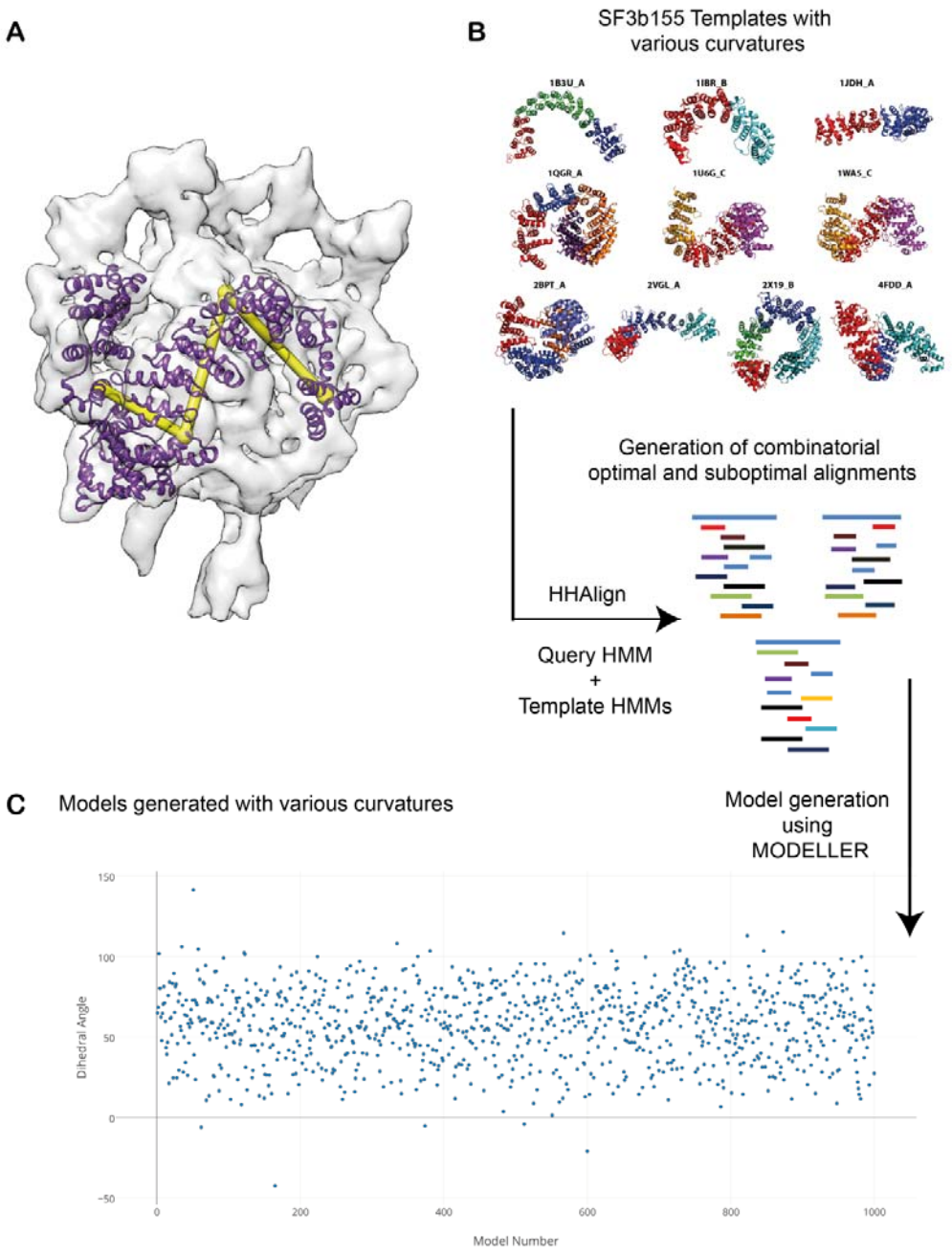


Figure S17. Hybrid fragment based modeling protocol. **(A)** SF3b155 has an S-shaped like curvature (yellow S-shaped marker). **(B)** Thus the hybrid fragment based modeling protocol was performed to see if we could capture this curvature without using the density information during the modeling process. **(C)** The models generated of various curvatures are shown in scatter plot based on the dihedral angles computed using the C-alpha positions of residues 630, 820, 1050 and 1155 (yellow markers). This shows our SF3b155 fragment based fitting protocol is unbiased.

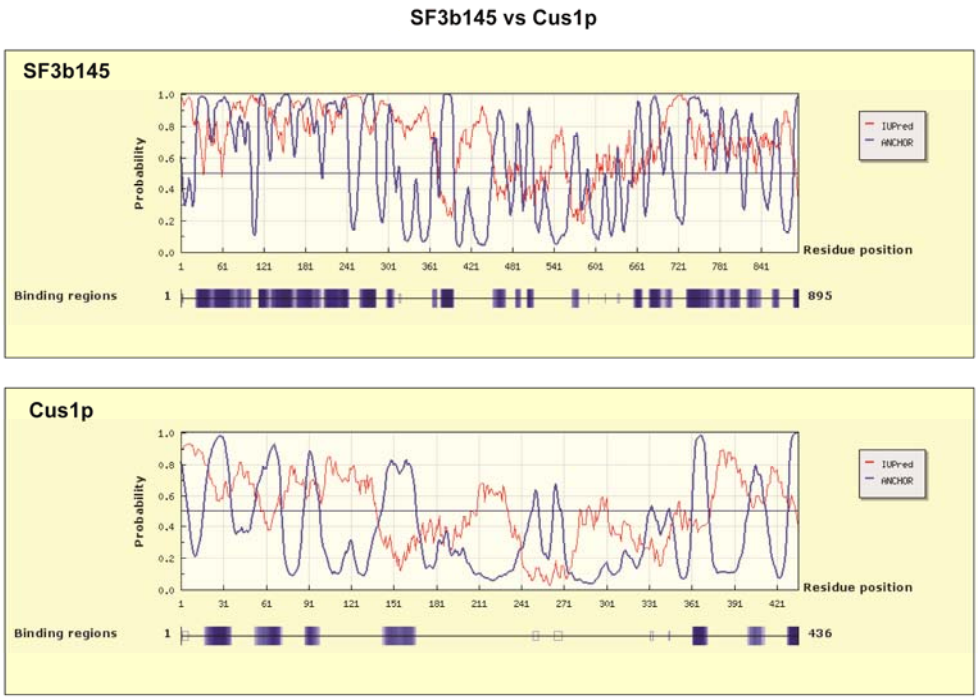


Figure S18. Intrinsically disordered regions of SF3b145 and Cus1p (Yeast homologue). The intrinsically disordered regions (IDR) for the proteins SF3b145 and Cus1p were generated using the "ANCHOR" server and the potential binding regions are also shown.

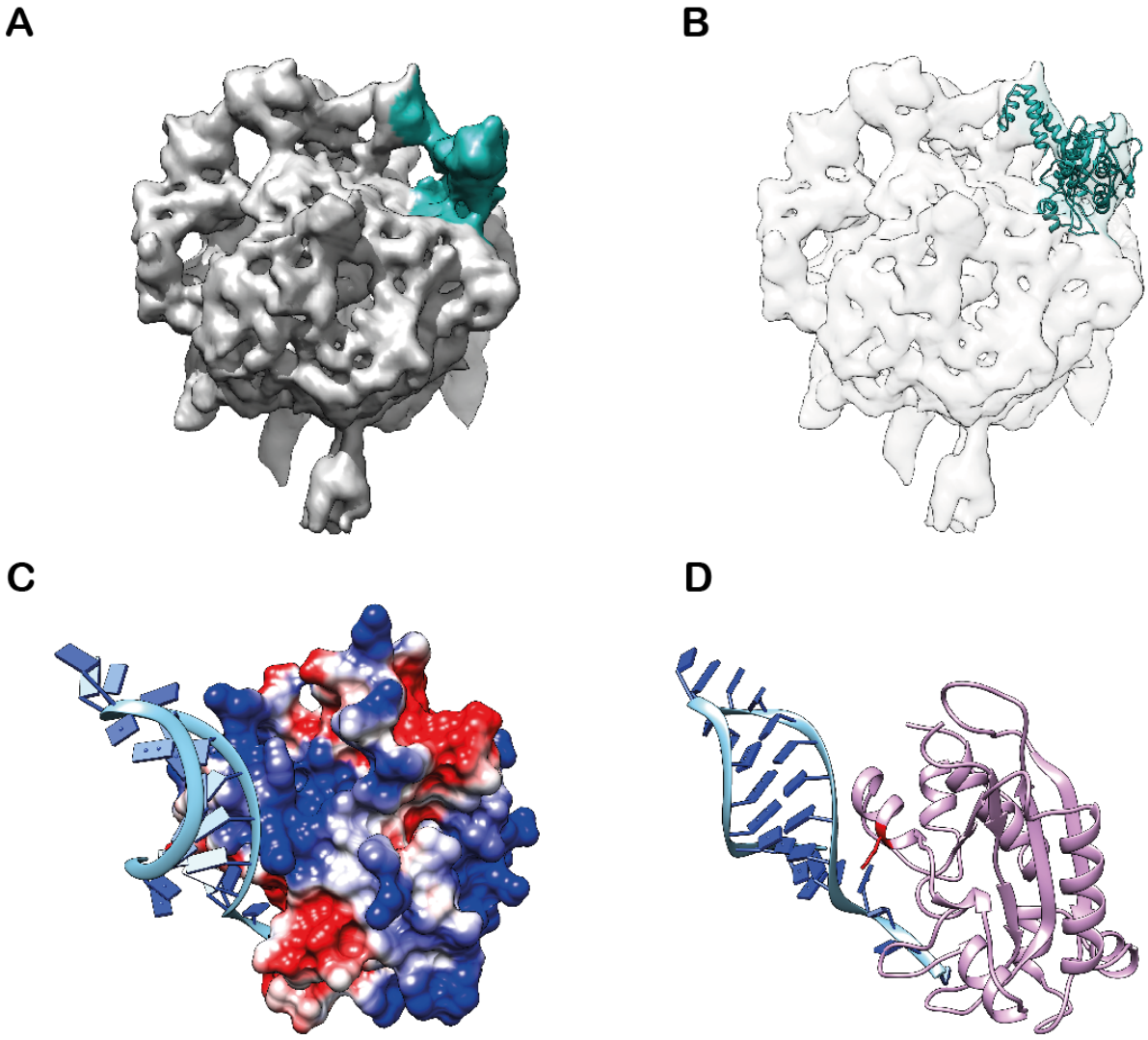


Figure S19. SF3b145 modeling and density fitting. (A) SF3b145 position in SF3b density obtained from whole map based fitting **(B)** Comparative model of SF3b145 RNaseH like fold along with SAP domain fitted into the localized density based on global search and along with interaction data (interacting with SF3b49 and SF3b130). **(C)** The electrostatic potential surface for SF3b145 shown along with stem-loop RNA interacting with highly positively charged surface (blue) **(D)** The Lysine 487 residue (marked red with side chain shown) interacts with the stem-loop RNA. This residue was proposed to interact with stem-loop RNA based on cold-suppressor mutation data of its yeast homolog Cus1p.

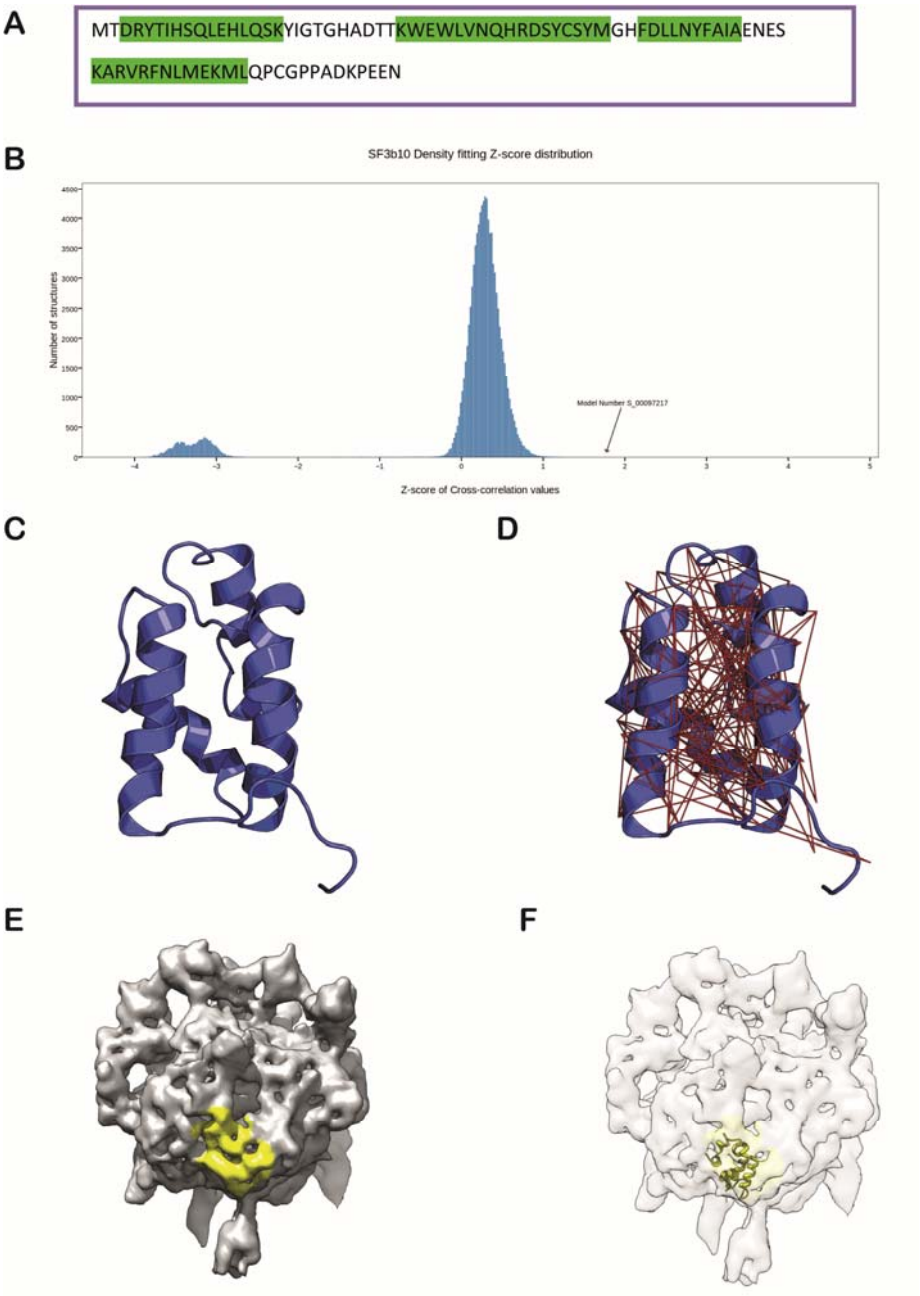


Figure S20. SF3b10 modeling and density fitting. (A) The alpha helical regions mapped onto SF3b10 sequence based on secondary structure prediction and used for *de novo* structure modeling. **(B)** The distribution of the density cross-correlation Z-scores for the 100,000 models with the Model #97217 (marked with arrow) at the tail end of the distribution. **(C)** The best *de novo* Rosetta model obtained from our robust scoring scheme. **(D)** Co-evolutionary residue-residue contacts used for re-ranking decoys mapped onto the model (97217) of SF3b10. **(E)** The localization information of SF3b10 derived from interaction data (with SF3b155) as well as density segmentation method. **(F)** SF3b10 fitted into the localized density.

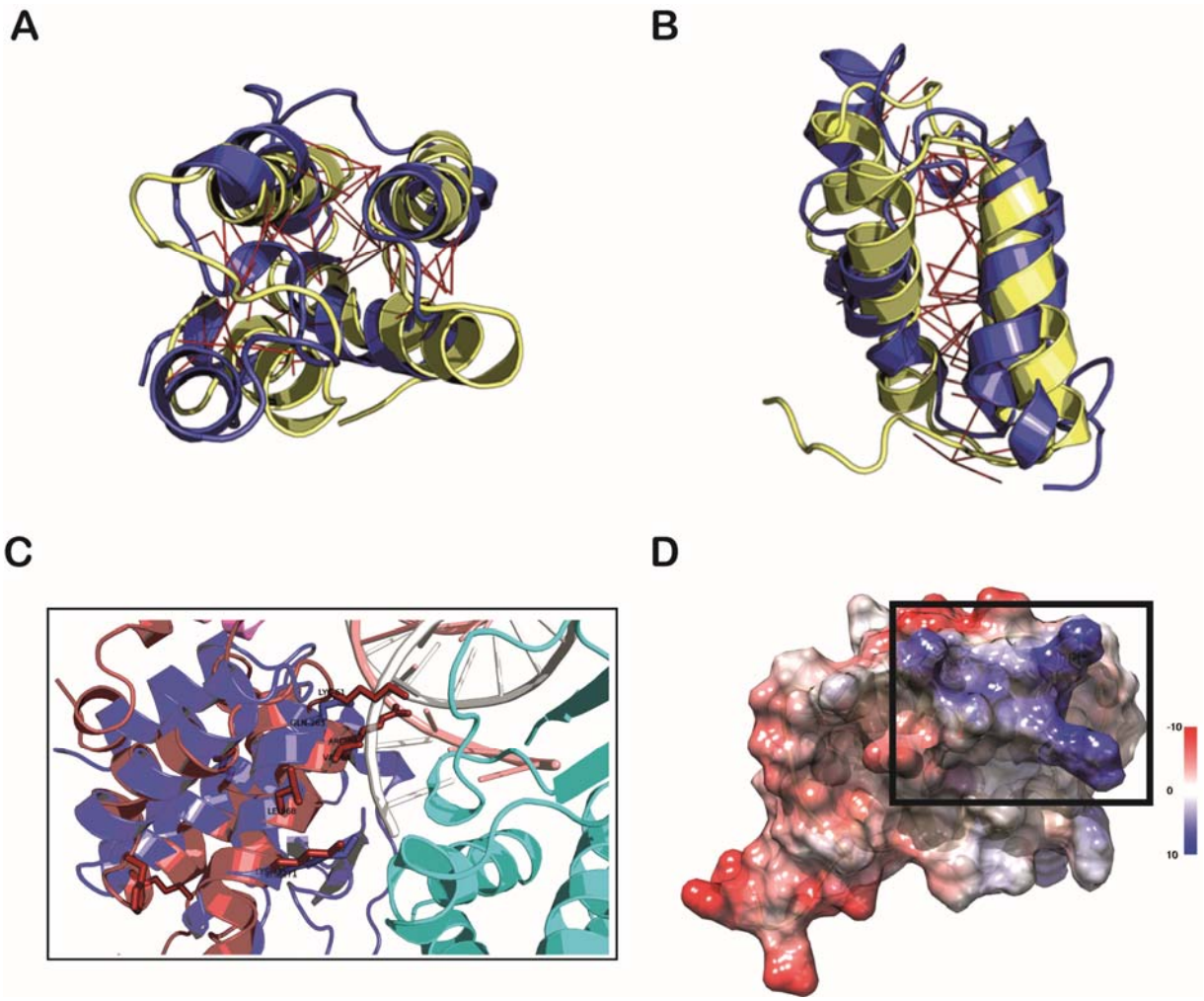


Figure S21. SF3b10 biological insights (A) Top view and **(B)** Side view of the structural alignment of the PDB ID: 2I5U with the Model #97217 shown along with common contacts mapped on to the structure which are mostly inter-helical. **(C)** Structural alignment of SF3b10 model with the protein VP35 shown along with the side chains of residues aligned at the equivalent positions. **(D)** Electrostatic surface potential highlighting the possible RNA binding patch (in box), positive surface is colored blue; negative surface is colored red.

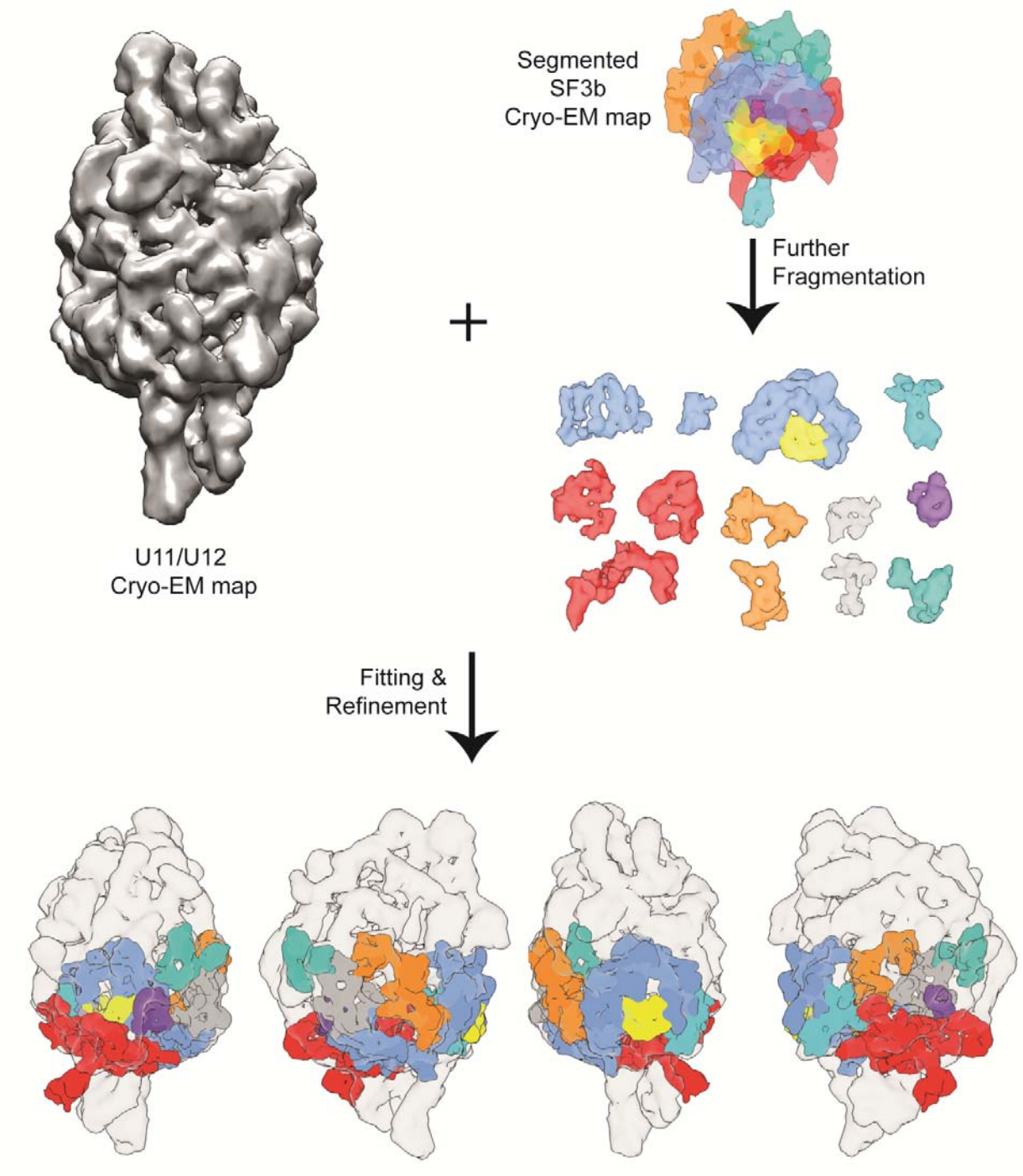


Figure S22. Modeling open form of SF3b complex. Segmented density of SF3b closed form coloured according to the individual subunit locations was iteratively fragmented resulting in 13 relatively large and rigid fragments. These fragments were used for fitting into the U11/U12 di-snRNP map (13.4Å). The fitting was performed using CHIMERA and the density-density cross correlation was maximized to obtain the open density map of SF3b complex.

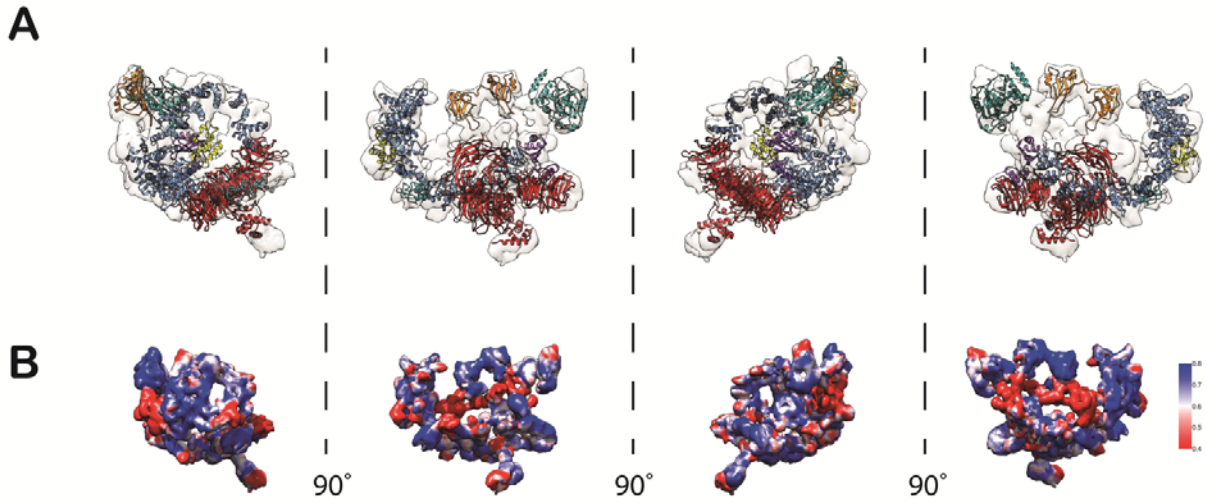


Figure S23. Modeling atomic level information for the open form of SF3b complex. (A) Fitting of atomic level models of SF3b components into the open density generated using U11/U12 as a guide map. **(B)** The local assessment of the atomic models into the density is computed using the local cross correlation.

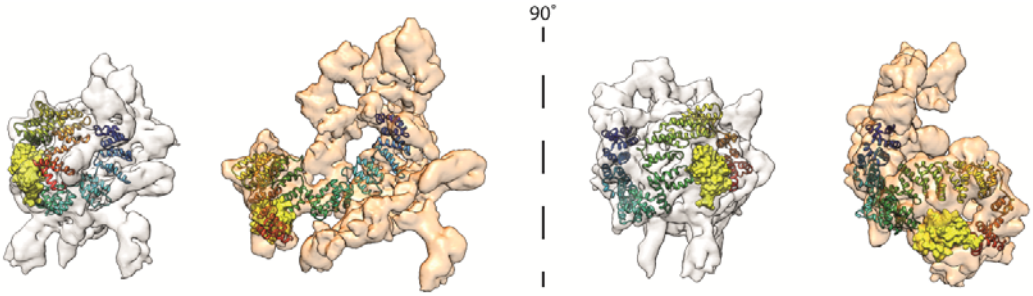


Figure S24. Comparison of open and closed forms of SF3b complex. The open and closed forms differ in the orientation of a flap, primarily composed of SF3b155. The SF3b10 (*yellow*) makes multiple contacts with the SF3b155 C-terminal region (flap), and maintains the integrity of the flap.

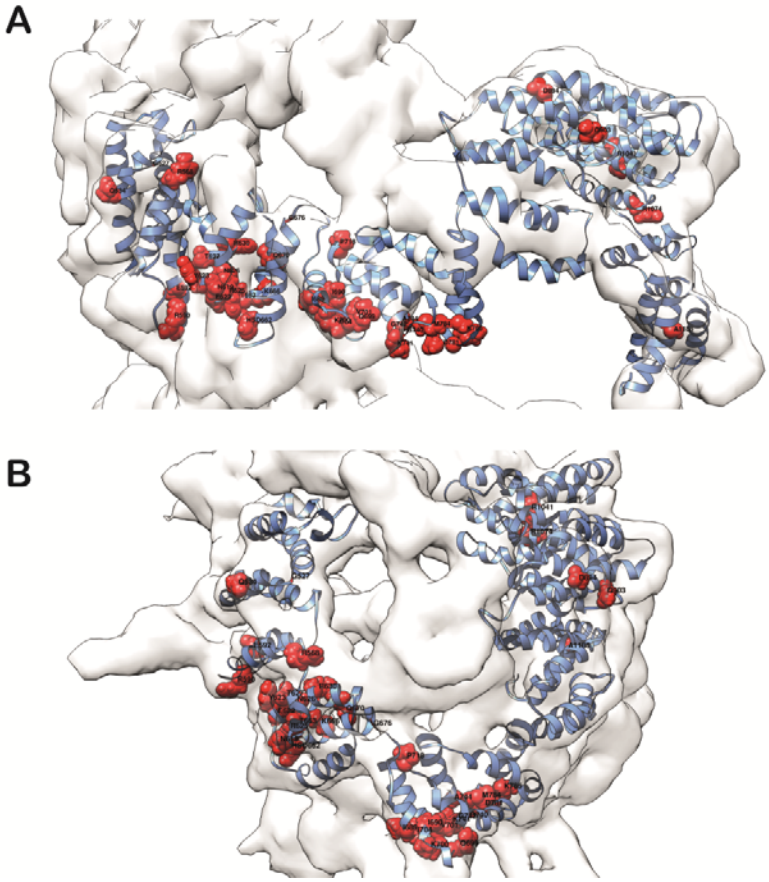


Figure S25. SF3b155 mutations. Cancer causing mutations of SF3b155 mapped onto the open **(A)** and closed **(B)** forms of SF3b complex. We note two clusters of sites on either sides of the hinge region which when substituted cause cancers. The cancer causing mutations were collated from (Kotake et al, 2007) and (Webb et al, 2013).

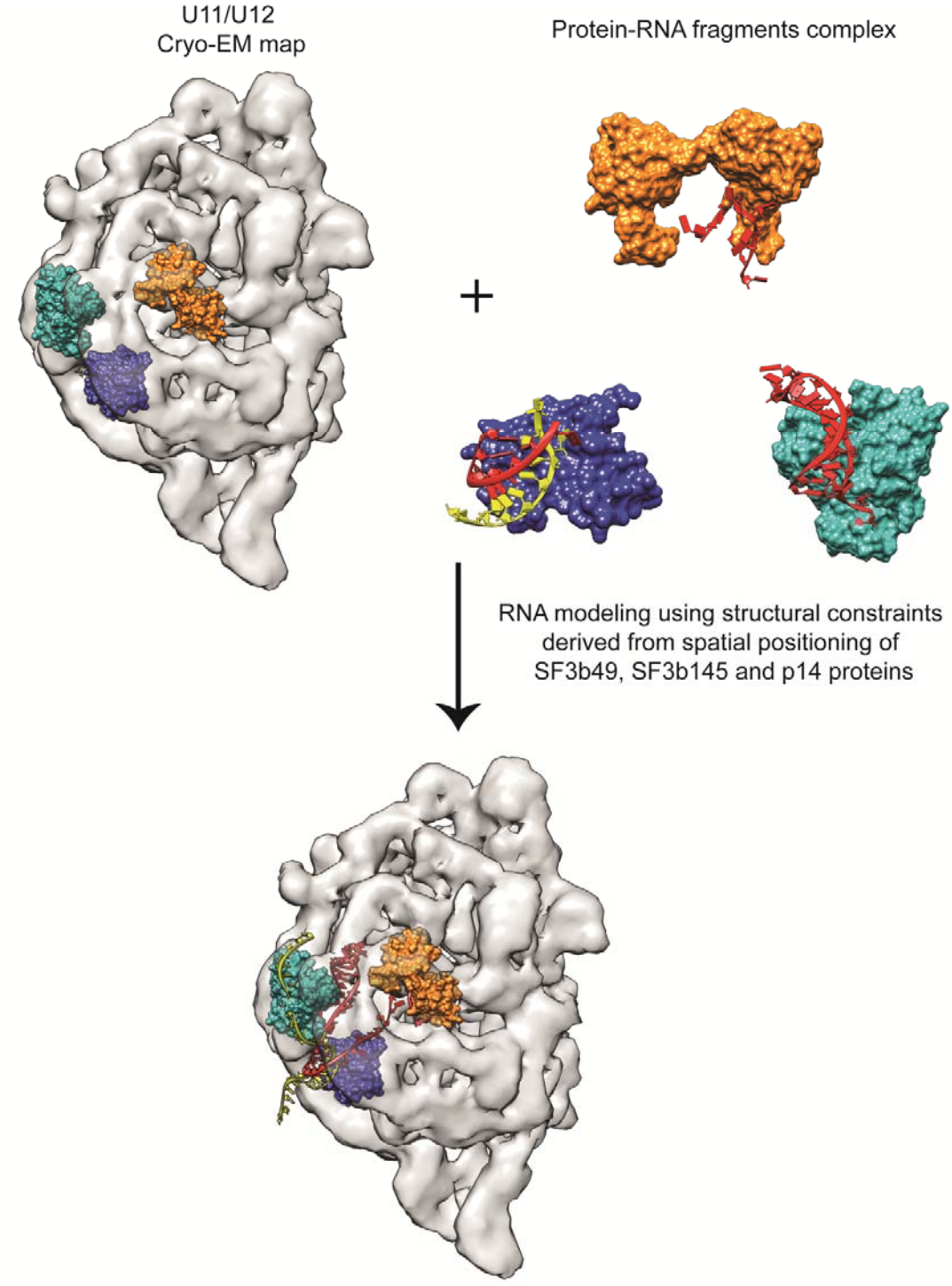


Figure S26. RNA modeling. The spatial positions of the three proteins SF3b49 (orange), SF3b145 (light sea green) and p14 (medium blue) along with the RNA fragments bound to them helped the structural modeling of U12 snRNA: pre-mRNA interaction.

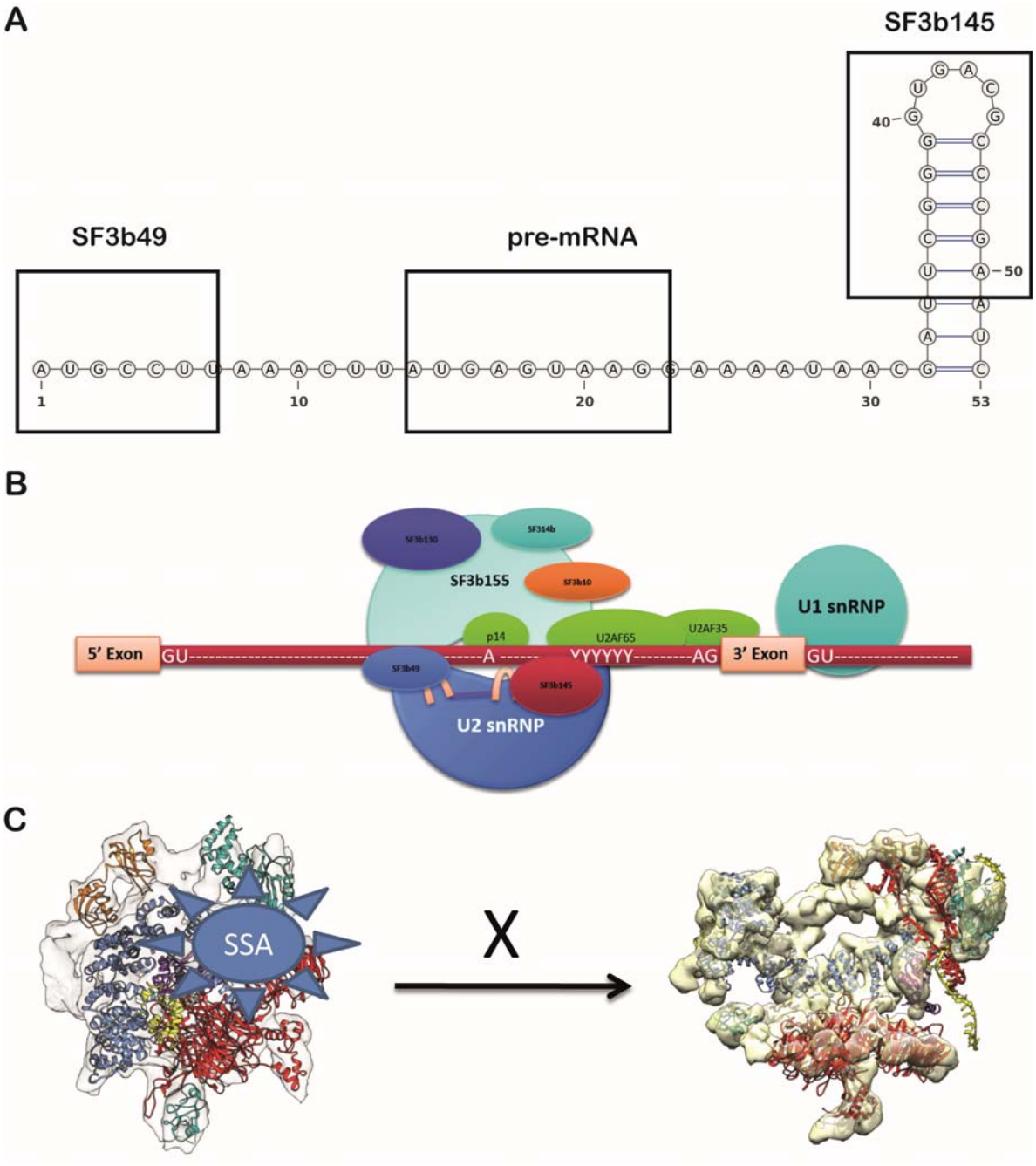


Figure S27. RNA protein interactions. (A) The region of U12 snRNA which was used for RNA structural modeling. The boxed regions show SF3b49, pre-mRNA and SF3b145 interacting regions with U12 snRNA. **(B)** The schematic shows the potential spatial arrangement of the proteins in U2 splicing pathway and its interaction with U2 snRNA based on previous literature and from our SF3b pseudo-atomic model. The spatial juxtaposition of proteins SF3b49, SF3b145 and p14 in U2 snRNP is also shown. **(C)** The binding of Spliceostatin A to SF3b130 and SF3b155 which might block the opening of SF3b.

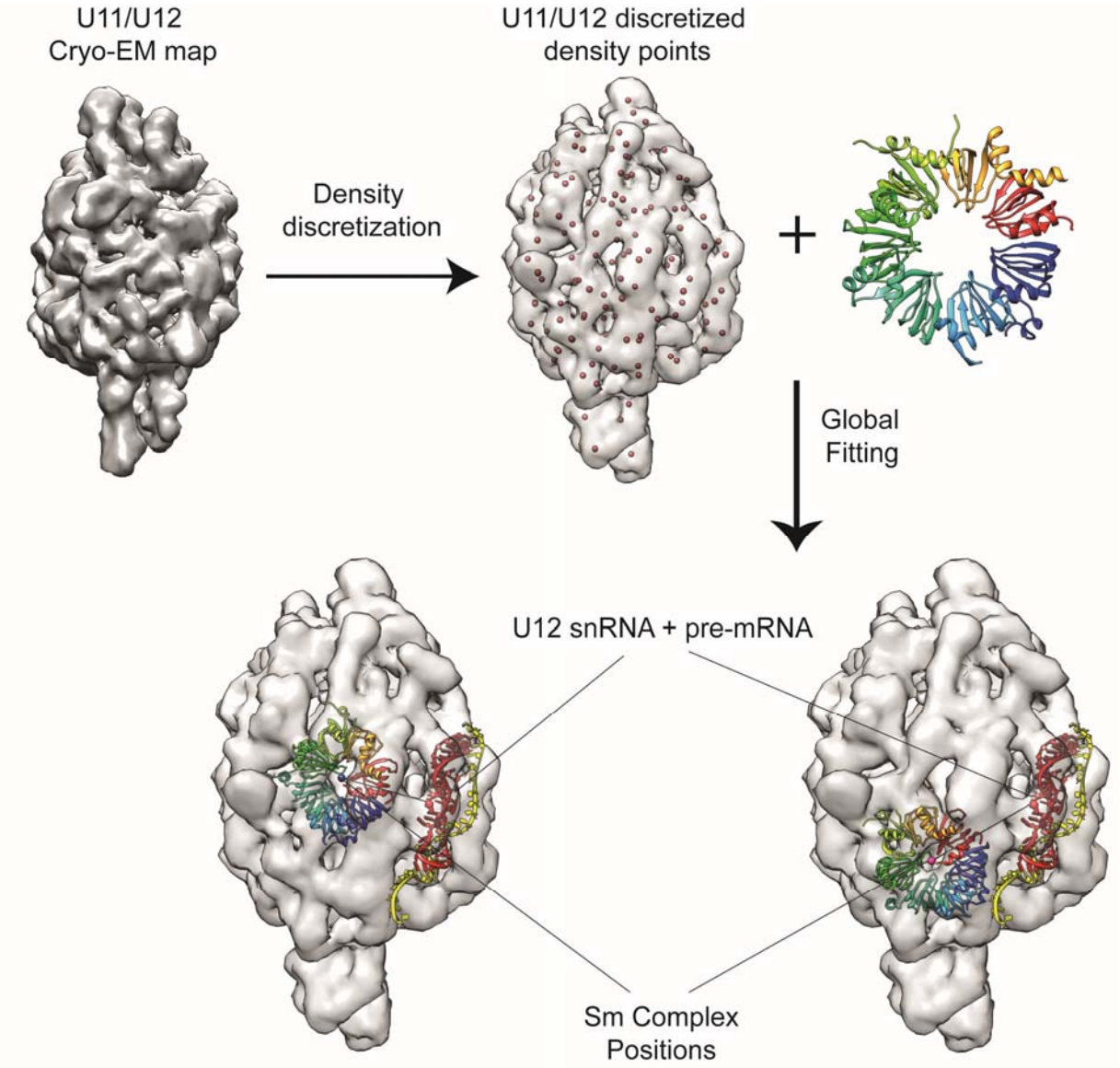


Figure S28. Probable U12 Sm complex positions in U11/U12 density map. A full atom Sm complex was generated using MaxSprout for the PDB ID 3PGW. A global fitting was performed using a discretized density cryo-EM map of U11/U12 di-snRNP with UCSF Chimera. This yielded two likely positions for this complex which were closer to U12 snRNA stem loop IIa among the top 20 best hits.

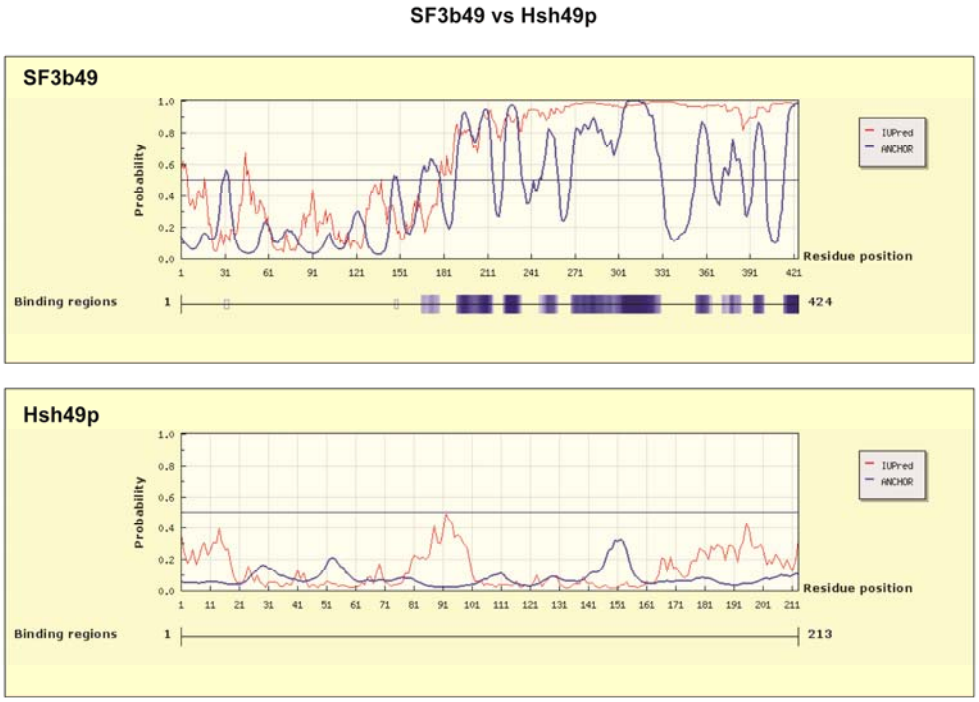


Figure S29. Intrinsically disordered regions of SF3b49 and Hsh49p (Yeast homologue). The intrinsically disordered regions (IDR) of the proteins SF3b49 and Hsh49p were generated using the "ANCHOR" server and the potential binding regions are also shown.

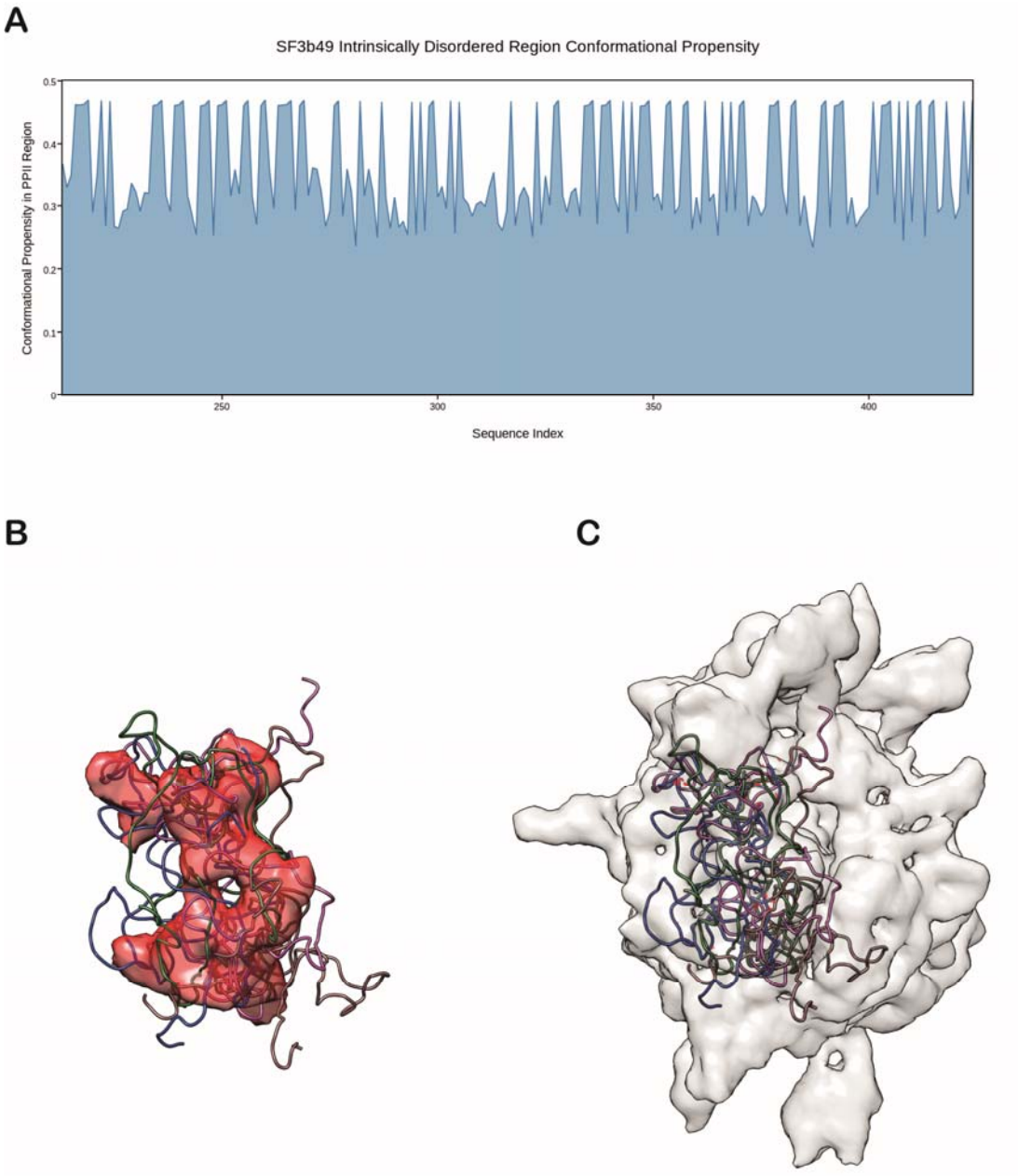
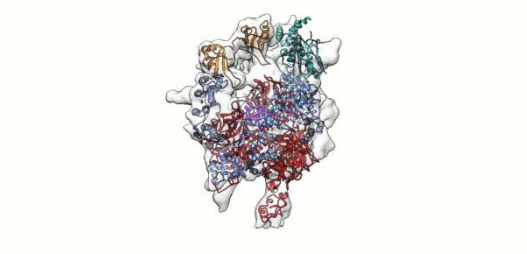


Figure S30. SF3b49 ensemble modeling. (A) The plot shows the conformational propensity of the polyproline II region (φ-75°±25°, ψ-145°±25°) of Ramachandran map for the population of 500,000 conformers generated for SF3b49 intrinsically disordered region. **(B)** The probable density region in SF3b cryo-EM map which is C-terminus to the SF3b49 RRM domain was used to filter the conformers. **(C)** The top SF3b49 IDR conformers which best match the shape fitted to the SF3b density are shown.

Supplementary Movie Legends

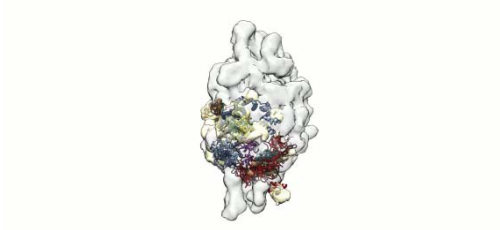


Supplementary Movie S1

This video shows the closed form of SF3b complex with atomic-level models of components fitted into the cryo-EM density map of SF3b. The individual components are shown in different colour (as in Figure 1).

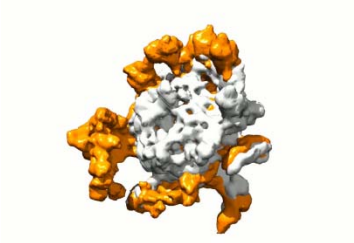
Supplementary Movie S2

This video shows the morphing between the different conformations of DNA Damage Binding Protein 1 (SF3b130 template).



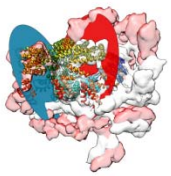
Supplementary Movie S3

This video shows the open form of SF3b complex with atomic-level models of components fitted into the cryo-EM density map of U11/U12 di-snRNP. The open form density is extracted from this map. The individual components are shown in different colour.



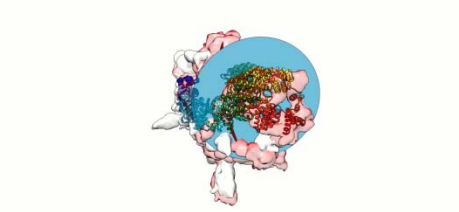
Supplementary Movie S4

This video shows the open and closed forms of SF3b density maps aligned with each other to show their differences.



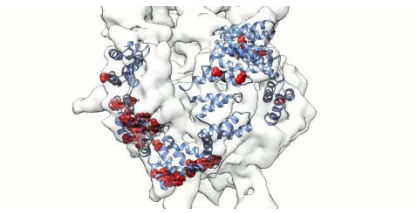
Supplementary Movie S5

This video shows the open and closed forms aligned with each other. The closed to open transition is governed by a flap (fitted discs) which moves about a hinge and is primarily formed by SF3b155 C-terminal region (as in Figure 4).



Supplementary Movie S6

This video shows the closed to open transition facilitated by SF3b155 C-terminal region in a 360 degree view (as in Figure 4).



Supplementary Movie S7

This video shows the mapping of the mutations causing cancer in SF3b155 concentrated around the hinge region and also the interacting surface with SF3b130 (as in Supplementary Figure S25).