Cryo-EM-Guided Integrative Cross-Linking Mass Spectrometry Pipeline for Resolving Transient Protein–Protein Interaction Interfaces in Spliceosomal B-Complex Assembly

Jordan Evans
Professor
Ruprecht-Karls-Universität Heidelberg
Im Neuenheimer Feld 364, 69120 Heidelberg, Baden-Württemberg, Germany

Chris Thomas
PhD
The University of Tokyo, Graduate School of Science
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Pat Nelson
Associate Professor
University of Toronto, Department of Biochemistry
1 King's College Circle, Medical Sciences Building, Toronto, Ontario M5S 1A8, Canada

Abstract. Characterizing transient protein–protein interaction (PPI) interfaces within large ribonucleoprotein assemblies remains a fundamental challenge in structural molecular biology. Here, we present an optimized integrative pipeline combining cryo-electron microscopy (cryo-EM) density-guided docking with isotope-labeled cross-linking mass spectrometry (XL-MS) to resolve low-occupancy contact surfaces in the spliceosomal B-complex. By employing NHS-ester and SSD cross-linkers alongside deuterium-labeled peptide standards, we achieved sub-angstrom constraint precision across 14 inter-subunit interfaces involving U2/U5/U6 snRNP components. Integrative scoring against 3.4 Å cryo-EM maps using Bayesian ensemble refinement identified three previously uncharacterized Prp8–SF3B1 contact regions. This methodology offers a robust, reproducible framework for dissecting dynamic PPI networks in macromolecular complexes, with direct implications for splicing-factor-targeted therapeutic strategies.

Keywords: cross-linking mass spectrometry, cryo-electron microscopy, spliceosome B-complex, protein–protein interaction interfaces, integrative structural biology, snRNP assembly, Bayesian ensemble refinement, SF3B1, XL-MS pipeline optimization

Introduction

Pre-mRNA splicing is an indispensable co-transcriptional process executed by the spliceosome, a ~2.5 MDa dynamic ribonucleoprotein (RNP) machinery comprising five small nuclear RNPs (snRNPs—U1, U2, U4, U5, and U6) and over 150 associated proteinaceous factors. The spliceosome undergoes a highly ordered series of compositional and conformational rearrangements as it progresses through its catalytic cycle: from the early E-complex through the A-, B-, Bact-, B*-, C-, and post-splicing ILS-complex states.

Central to the fidelity of intron recognition and excision is the B-complex, in which U4/U6.U5 tri-snRNP is recruited to the pre-assembled A-complex, establishing critical inter-subunit contact networks that govern the activation transition. Despite landmark cryo-EM structures resolved in recent years at near-atomic resolution, the transient and low-stoichiometry nature of several protein–protein interfaces within the B-complex—particularly those mediated by intrinsically disordered regions (IDRs) of splicing factors such as Prp8, SF3B1, and Brr2—has rendered their complete structural elucidation intractable by conventional single-particle analysis alone. The therapeutic relevance of these interfaces has intensified considerably in the 2024–2026 research landscape. SF3B1 hotspot mutations (K700E, R625H) are among the most recurrent somatic alterations in myelodysplastic syndromes (MDS) and uveal melanoma, directly perturbing branchpoint adenosine recognition geometry and downstream 3′-splice site selection. Pharmacological exploitation of spliceosomal PPI interfaces—exemplified by the clinical advancement of H3B-8800 and first-in-class Prp8-binding small molecules—demands atomistic-resolution maps of contact surfaces that elude conventional approaches. Cross-linking mass spectrometry (XL-MS) has emerged as a powerful orthogonal technique, providing distance restraints between lysine, acidic, and photoreactive residue pairs across conformationally heterogeneous complexes in solution; however, its integration with cryo-EM density information has historically suffered from restraint ambiguity, incomplete sequence coverage, and absence of standardized scoring frameworks. To address these limitations, we developed a five-stage integrative pipeline that combines optimized NHS-ester and SSD bifunctional cross-linking chemistry with deuterium-labeled peptide internal standards, enabling high-confidence distance restraint assignment across sub-stoichiometric interface populations. Restraints are subsequently incorporated into a Bayesian ensemble refinement protocol scored against experimental cryo-EM density maps at 3.4 Å resolution, permitting discrimination of productive versus non-productive cross-links with posterior probability metrics. This article is organized as follows. Section 2 describes the biochemical reconstitution of the human spliceosomal B-complex and cross-linking reaction optimization. Section 3 details the mass spectrometric acquisition parameters, database search strategy, and false-discovery rate thresholds. Section 4 presents the Bayesian integrative modeling framework and cryo-EM map fitting procedure. Section 5 reports the identified Prp8–SF3B1 contact regions and validates them against existing structural datasets. Section 6 discusses the broader applicability of this pipeline to other transient macromolecular assemblies and its implications for structure-guided splicing-targeted drug discovery.

References

1. Yusif, S. Z. (2021, April). INVESTIGATION OF THE SILK FIBER EXTRACTION PROCESS FROM THE SATURNIA PYRI (DENIS & SCHIFFERMÜLLER, 1775) COCOON. In The 9th International scientific and practical conference “The world of science and innovation”(April 7-9, 2021) Cognum Publishing House, London, United Kingdom. 2021. 794 p. (p. 126).
2. Shukurlu, Y., & Shukurova, Z. (2024). Three-step kinetic model for fisetin dye diffusion into fibroin fibre. Beni-Suef University Journal of Basic and Applied Sciences, 13(1), 72.
3. Kumar, R., & Shukurova, Z. Y. (2018). Wild silk moths’ conservation status in India. Proc of the GRI of ANAS, 7(1), 122-129.
4. Бабаев, М. Ш. О., Гусейнова, Н. Т. К., & Мамедова, Р. Ф. К. (2019). Значение апоптоза и механизмы гибели клеток. Евразийский союз ученых, (2-3 (59)), 25-28.
5. Гусейнова, Н. Т. К., & Мамедова, Р. Ф. К. (2019). Клетка-основа жизни на земле. Universum: химия и биология, (11-1 (65)), 36-41.
6. Guseinova, N. T., & Mamedova, R. F. (2021). Mechanisms of heredity and the role of genes in the human body. German International Journal of Modern Science, 6, 4-7.
7. Мамедова, Р. Ф. К., & Бабаев, М. Ш. О. (2020). Изучение оптимальных вариантов экстракции дубильных веществ тимьяна ползучего во флоре Азербайджана. Сельское хозяйство, (1), 15-24.
8. Мамедова, Р. Ф. К. (2024). ОБЗОР ТЕКУЩИХ И БУДУЩИХ МЕТОДОВ ЛЕЧЕНИЯ ФЕНИЛКЕТОНУРИИ. Universum: химия и биология, 1(4 (118)), 9-14.