A120 · Mining protein–RNA complexes for recurring RNA-binding motifs

Presenter: Sharra MN Lewis — Worcester Polytechnic Institute

Authors: Sharra Lewis, Aya Narunsky

Session: Day 2 · Fri Oct 2, 2026 · 2:15–4:15 PM

RNA-binding proteins (RBPs) are a large class of proteins that bind RNA and play central roles in essential cellular processes. Disruption of interactions between RBPs and their RNA targets has been associated with a variety of pathologies, including neurological disorders such as Alzheimer's disease, ALS, and fragile X syndrome, making these interfaces compelling drug targets and highlighting the importance of understanding these interfaces at the molecular level. Despite their importance, the sequence and structural determinants of RNA binding remain incompletely understood, in part due to the relatively small number of available protein-RNA complex structures. Identifying the molecular motifs that mediate these interactions could help bridge this gap. Here, we investigate protein-RNA complexes from the Protein Data Bank (PDB) to identify RNA-binding motifs in proteins. We collected a nonredundant dataset of 400 protein-RNA complexes and searched these proteins for recurring sequence fragments, referred to as "themes." These themes were analyzed using InterProScan to determine whether they correspond to previously characterized protein domains or functional motifs involved in RNA binding. A total of 120 themes were identified, 12 of which did not correspond to any annotated RNA-binding domain or motif. The results suggest that recurring short sequences may represent conserved regions associated with specific molecular functions, even among proteins lacking obvious homology, and may provide insight into the emergence and evolution of these functions. These recurring sequences may enable prediction of candidate RNA-binding regions in proteins even in the absence of structural data, extending insights beyond structurally characterized protein-RNA complexes.

Mining protein–RNA complexes for recurring RNA-binding motifs

Sharra Lewis[1] and Aya Narunsky[1]

[1]Department of Chemistry and Biochemistry, Worcester Polytechnic Institute, Worcester, MA

RNA-binding proteins (RBPs) are a large class of proteins that bind RNA and play central roles in essential cellular processes. Disruption of interactions between RBPs and their RNA targets has been associated with a variety of pathologies, including neurological disorders such as Alzheimer's disease, ALS, and fragile X syndrome, making these interfaces compelling drug targets and highlighting the importance of understanding these interfaces at the molecular level. Despite their importance, the sequence and structural determinants of RNA binding remain incompletely understood, in part due to the relatively small number of available protein-RNA complex structures. Identifying the molecular motifs that mediate these interactions could help bridge this gap. Here, we investigate protein-RNA complexes from the Protein Data Bank (PDB) to identify RNA-binding motifs in proteins. We collected a nonredundant dataset of 400 protein-RNA complexes and searched these proteins for recurring sequence fragments, referred to as "themes." These themes were analyzed using InterProScan to determine whether they correspond to previously characterized protein domains or functional motifs involved in RNA binding. A total of 120 themes were identified, 12 of which did not correspond to any annotated RNA-binding domain or motif. The results suggest that recurring short sequences may represent conserved regions associated with specific molecular functions, even among proteins lacking obvious homology, and may provide insight into the emergence and evolution of these functions. These recurring sequences may enable prediction of candidate RNA-binding regions in proteins even in the absence of structural data, extending insights beyond structurally characterized protein-RNA complexes.