

Representation of RNA-Protein Complexes in Topology Graphs

Mariella Zunker¹, Jan Niclas Wolf¹, Jörg Ackermann¹, and Ina Koch¹

Molecular Bioinformatics, Institute of Computer Science, Johann Wolfgang Goethe University, Robert-Mayer-Str. 11-15, 60325 Frankfurt am Main, Germany

The Protein Topology Graph Library (PTGL) is a web server and database of topology graphs for their visualization and analysis [1,2]. We added the parsing of RNA to the software and integrated vertices for RNA chains into the complex graph (CG) definition. The integration of protein as well as RNA chains in CGs enables the analysis of topological interactions of RNA-protein complexes. The representation of RNA-protein complexes in topology graphs allows the application of graph algorithms aiding in classification and structural comparison. The abstraction level of protein chains is an ideal level of resolution for studying spatial relationships in an RNA-protein complex.

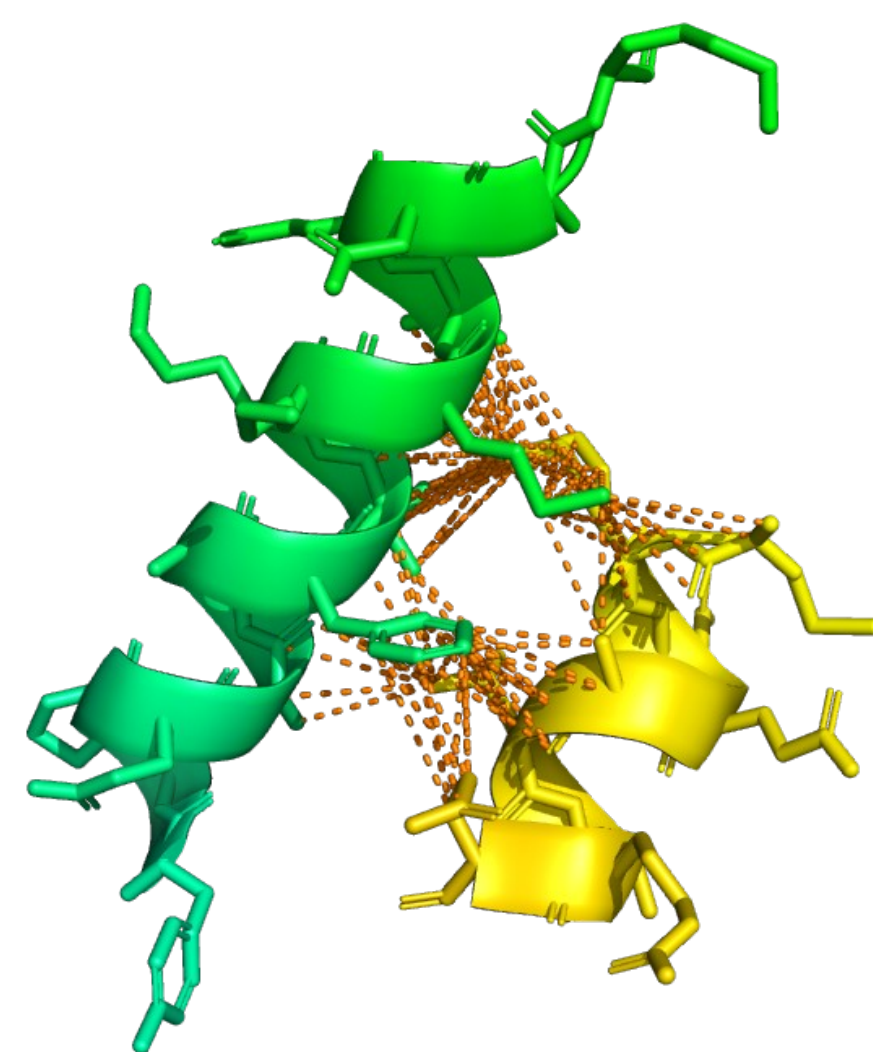
Methods – The Protein Topology Graph Library

Information on a Molecule's Atoms, Structure, and Coordinates

AA	Atom	X-	Y-	Z-Coord
ALA	N	265.643	286.241	166.773
ALA	C	264.215	286.041	166.970
ALA	C	263.671	287.020	167.997
ALA	O	264.410	287.841	168.541
ALA	C	263.931	284.614	167.395
GLN	N	262.370	286.925	168.260
GLN	C	261.679	287.870	169.121
GLN	C	260.525	287.159	169.812
GLN	O	259.895	286.265	169.241
GLN	C	261.178	289.071	168.306
LYS	C	260.534	286.779	173.968
LYS	C	260.422	285.268	174.038
LYS	C	259.317	284.852	174.994
LYS	N	259.609	285.247	176.403
GLU	N	257.288	286.713	170.342
GLU	C	256.002	287.036	169.741
GLU	C	254.879	286.993	170.769

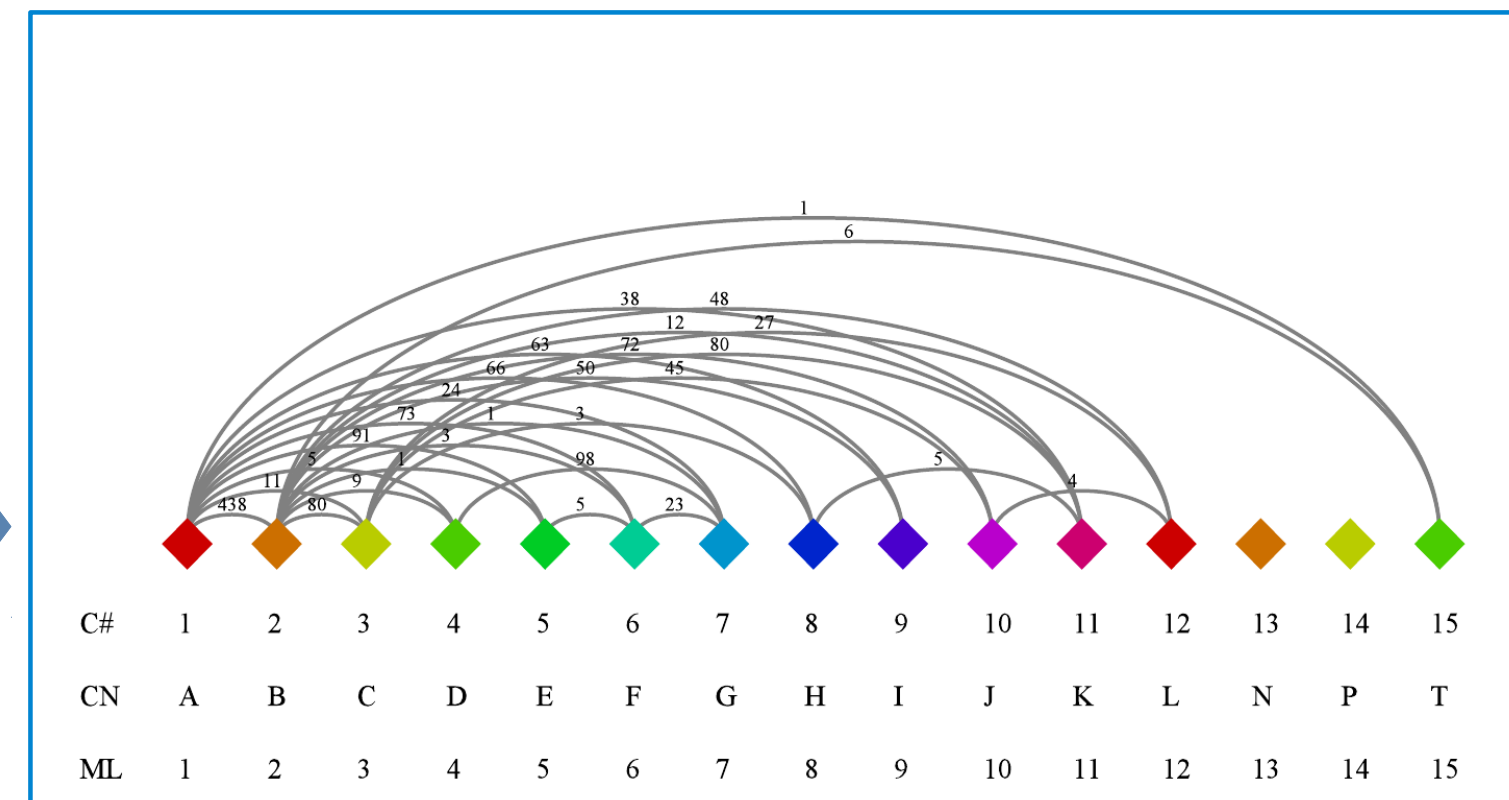
PDB-File [3]:
Atom
Coordinates.

Contact Computation



Computation: Atom contacts.

Depiction of Contacts of Structural Components in Complex Graphs

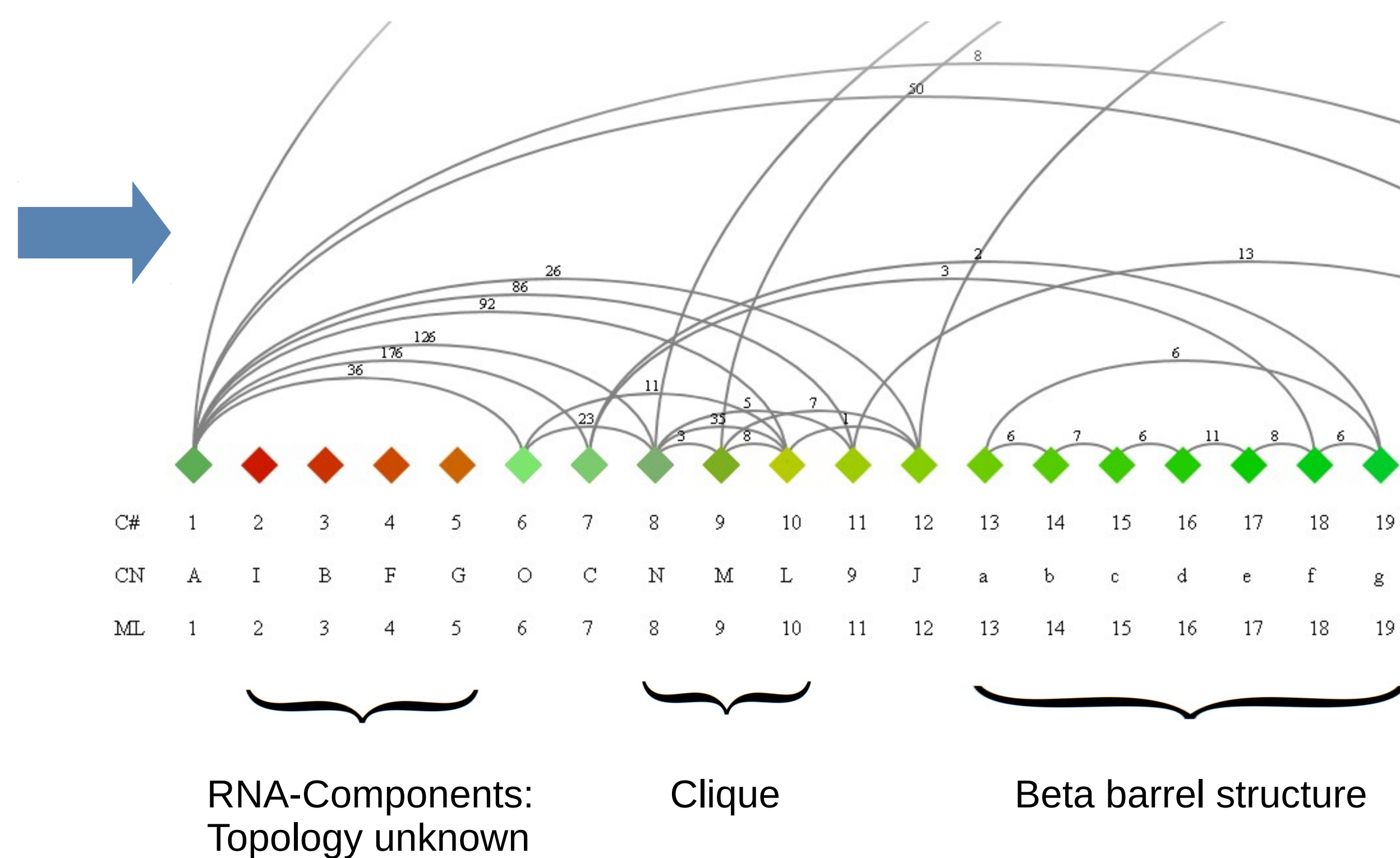
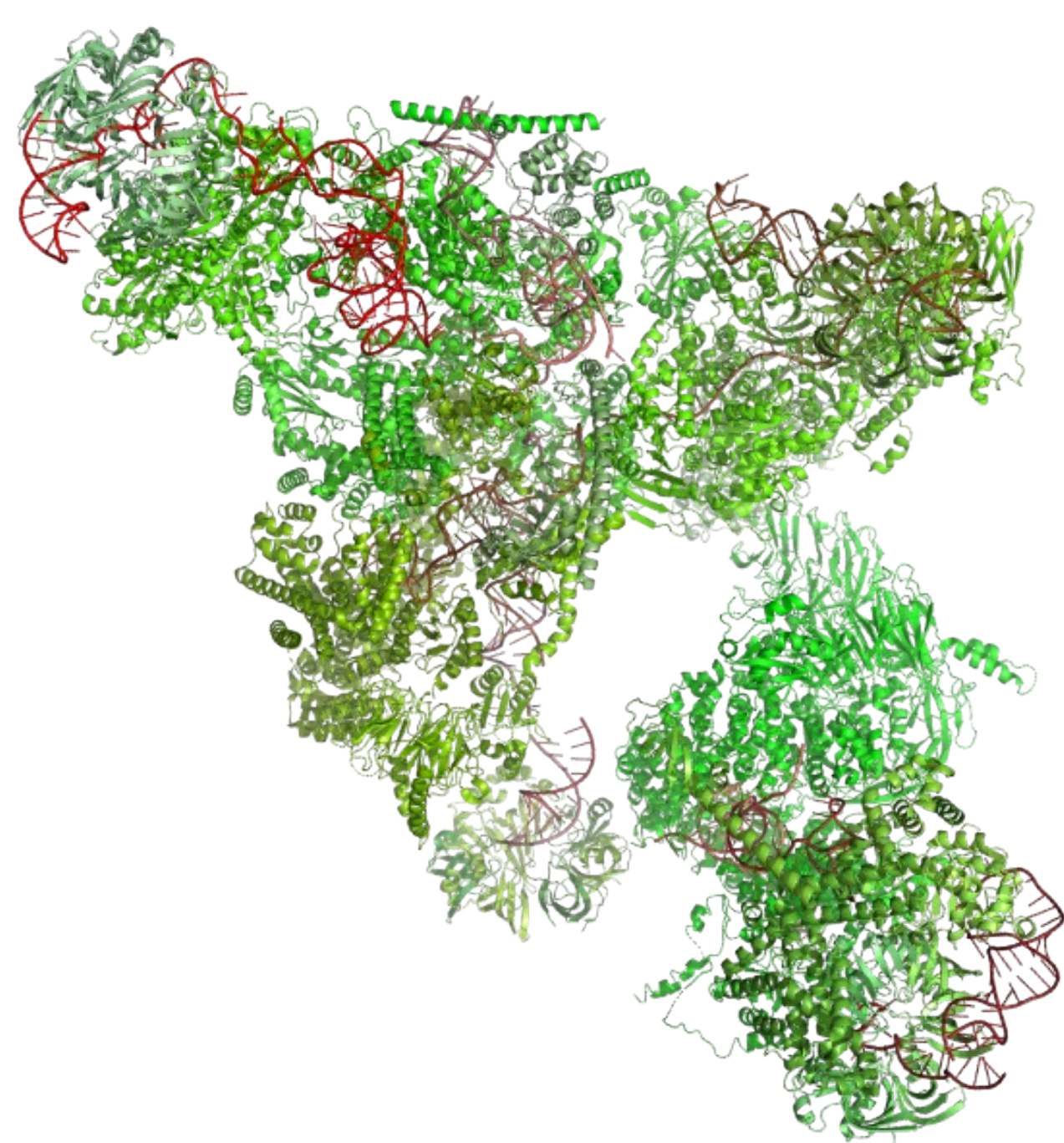


Output: Complex graph.

Results

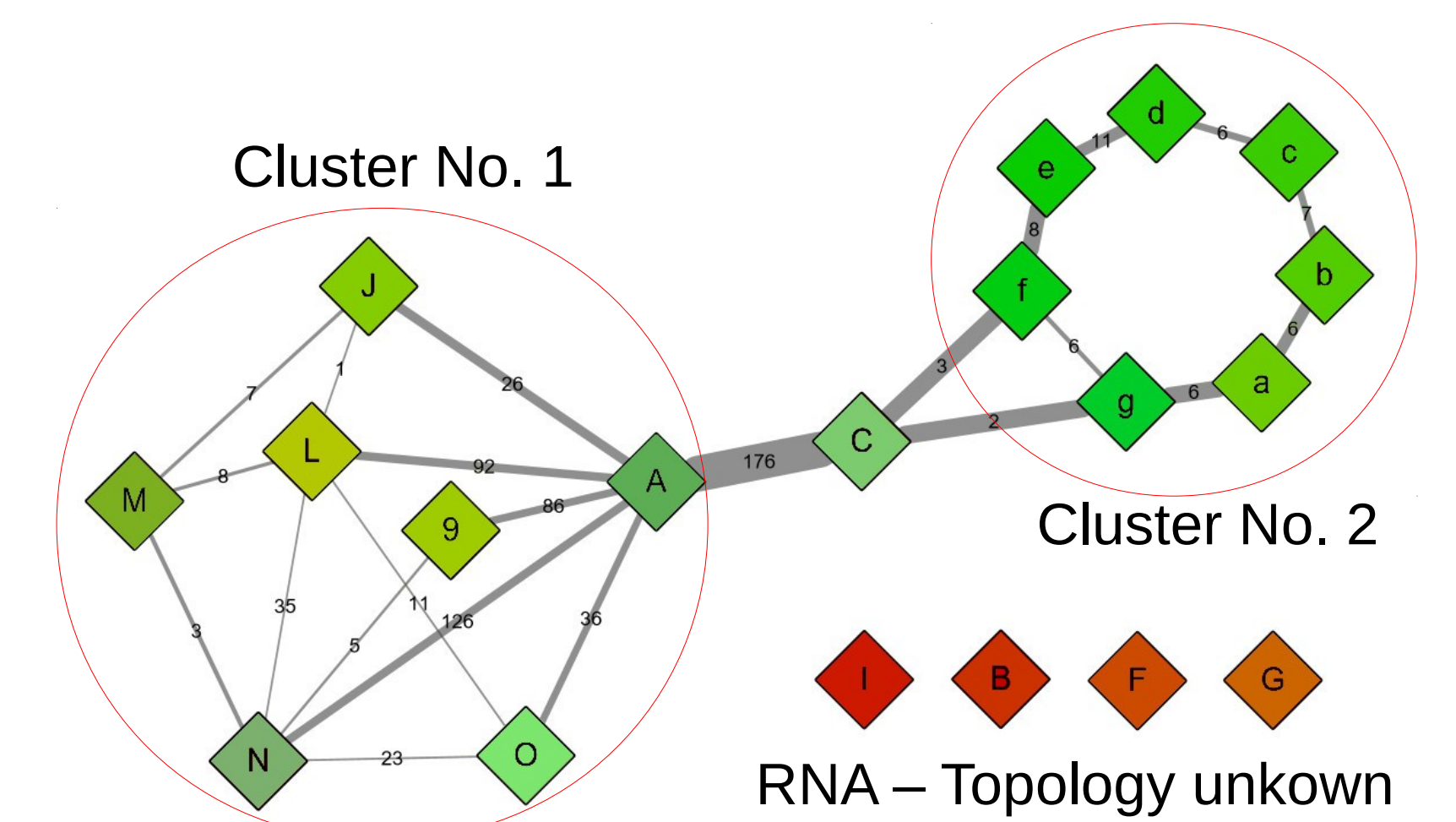
Extension of PTGL: Computation of RNA data [4].

New Complex Graph: Includes RNA components.

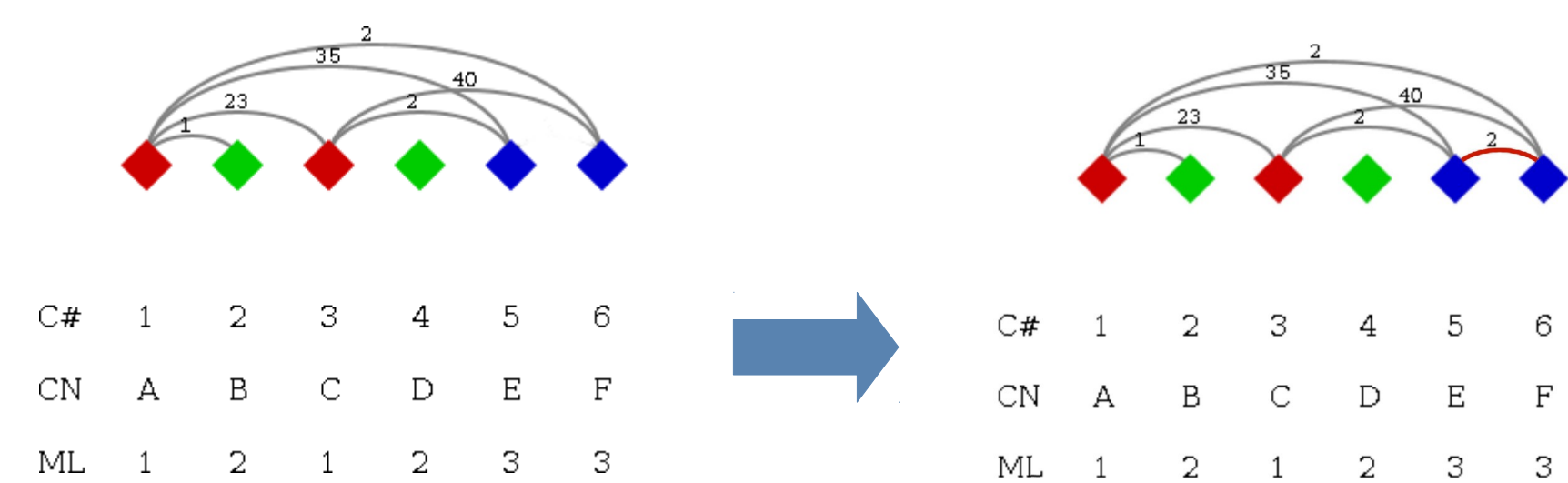


Analytical possibilities

- Betweenness centrality and other graph analytics



- Binding dynamics

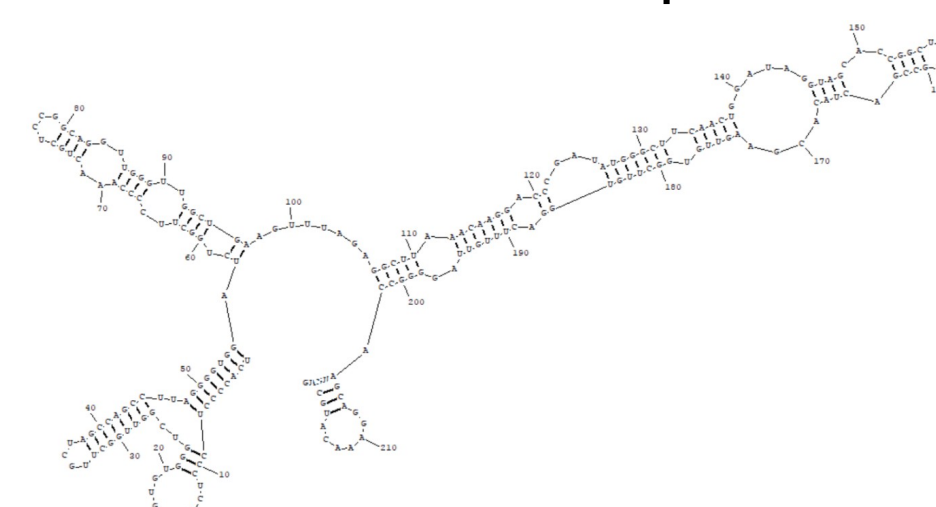


- Comparison of structurally conserved regions

Conclusion and Outlook

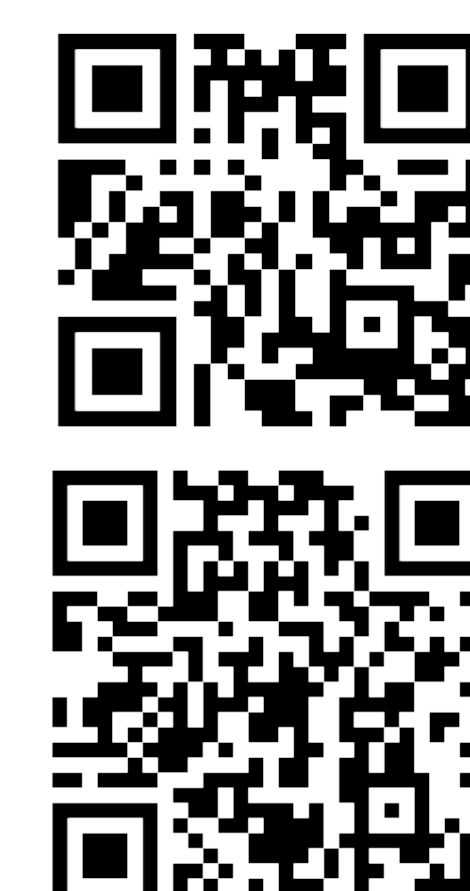
- Visualization of complex topologies
- Topology graphs can be retrieved and analyzed using the PTGL database
- RNA is now included in the topology graphs
- Application of graph algorithms

- Possibly inclusion of RNA secondary structure for contact computation



- Assembly pathways of RNA-protein complexes

Access



PTGL Website

<http://ptgl.uni-frankfurt.de>

Source Code

<https://github.com/MolBIF-FM/vplg>

References

- [1] Schäfer T., Scheck A., Bruneß D., May P., Koch I. (2015). The new protein topology graph library web server. *Bioinformatics*.
- [2] Wolf, J.N., Keßler, M., Ackermann, J., Koch I. (2020). PTGL: extension to graph-based topologies of cryo-EM data for large protein structures. *Bioinformatics*.
- [3] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. (2000) The Protein Data Bank *Nucleic Acids Research*, 28: 235-242.
- [4] Zhan, X., Yan, C., Zhang, X., Lei, J., Shi, Y. (2018). Structures of the human pre-catalytic spliceosome and its precursor spliceosome. *Cell Res* 28, 1129-1140.