Evaluating Computational Methods for ssDNA Aptamer 3D Structure Prediction

Selma Bengaouer [(1)], Bérangère Avalle [(1)], Irene Maffucci [(1)]

(1) Université de technologie de Compiègne, UPJV, CNRS, Enzyme and Cell Engineering, Centre de recherche Royallieu, CS 60 319 - 60 203, Compiègne Cedex, France.

Aptamers' inherent flexibility enables them to adopt specific 3D conformations, allowing high-affinity and selective binding to diverse target molecules, often with dissociation constants in the nM to pM range. Thus, high-resolution structural determination is crucial to understand their function, though the existing experimental methods are often costly and long. In this context, *in silico* 3D structure prediction is an efficient alternative to study aptamers function and design. Many algorithms have been developed for 3D structure prediction, with a focus on RNA. However, the growing interest in single stranded DNA (ssDNA), due to its greater stability and ease of synthesis, has highlighted the need to adapt these methods for DNA aptamers. To explore this potential, this study assessed three RNA 3D structure prediction methods (RNAComposer, SimRNA, and Vfold3D) based on their performance in the CASP15 challenge to evaluate their applicability to ssDNA. At this scope, a dataset of 93 experimentally determined ssDNA structures, including challenging motifs such as G-quadruplexes and pseudoknots, was built. Various metrics, such as RMSD, GDT_TS, INF, and TM-scores, were employed to benchmark the accuracy of the predictions. Vfold3D demonstrated the most proficiency in capturing global folds and interaction networks; however, it encountered challenges with smaller structures, faltering 15 out of the 93 cases. In contrast, SimRNA and RNAComposer successfully predicted all structures, with SimRNA slightly outperforming RNAComposer. However, all methods encountered difficulties with G-quadruplexes, and with structures containing multiple loops or long-distance interactions, which strongly increase the intrinsic flexibility of ssDNA. Therefore, these results indicate that much effort has still to be done for the modeling of ssDNA and that their structures need to be investigated considering their dynamics, e.g. by means of enhanced sampling molecular dynamics techniques.