DNA twisting in complexes with proteins: choosing metrics

[12]KA Ugryumov, [134]EO Vasilenko, [1]DN Kalyuzhny, [1]AA Anashkina, [1]YD Nechipurenko

[1]Institute of Molecular Biology RAS, [2]Moscow Institute of Physics and Technology, [3]Skolkovo Institute of Science and Technology, [4]Institute for Systems Biology and Medicine

We explore patterns of geometric change in DNA molecules within complexes involving proteins of various structures and functions. Structures from the PDB database [1], which contains atomic coordinates of DNA-protein complexes obtained through X-ray diffraction and electron microscopy, were analyzed. We selected metrics useful for describing local geometric changes in protein-binding DNA regions. As a measure of DNA twisting, we used the rotation angle between adjacent base pairs (Twist) [2]. The average deviation of this value in DNA-protein complexes from reference values, both unsigned (Twist Mean Absolute Difference, TMAD) and signed (TMD), was calculated. The reference Twist angle for free DNA was estimated in two ways: (1) assuming a canonical value of 36° for straight B-DNA [3], and (2) using the deep learning algorithm "Deep DNAshape" [4], which predicts DNA helical geometry based on nucleotide sequences. This comparison helps separate protein-induced deformation from the intrinsic "context effect" of the oligonucleotide itself [5]. To estimate shifts within base pairs, the Stretch and Opening parameters were used [6]. Comparisons between angles in DNA-protein complexes and the two references differ significantly. Relative to idealized straight DNA, twisting appears more pronounced; however, accounting for sequence context reveals that such deformations are partly inherent to the oligonucleotides. This suggests that a reduced requirement for additional twisting in a recognized DNA region may be an evolutionary factor for proteins. The extent of deformation correlates with protein function: on average, transcription factors induce weaker geometric alterations in DNA, while structural proteins cause stronger ones. Protein binding sites often correspond to larger Stretch and Opening values. The studied proteins generally promote DNA unwinding, as the mean TMD is negative. No significant differences were observed among transcription factors with different DNA-binding motifs. The developed algorithm enables the evaluation of deformation in any DNA-protein complex available in the PDB. This approach provides a pathway toward describing the mechanisms of DNA allostery [7].

[1] Norambuena, T., & Melo, F. (2010). The protein-DNA interface database. BMC bioinformatics, 11(1), 262.

[2] Calladine, C. R., Drew, H. R., & McCall, M. J. (1988). The intrinsic curvature of DNA in solution. Journal of molecular biology, 201(1), 127-137.

[3] Li, S., Olson, W. K., & Lu, X. J. (2019). Web 3DNA 2.0 for the analysis, visualization, and modeling of 3D nucleic acid structures. Nucleic acids research, 47(W1), W26-W34.

[4] Li, J., Chiu, T. P., & Rohs, R. (2024). Predicting DNA structure using a deep learning method. Nature communications, 15(1), 1243.

[5] Dlakić, M., & Harrington, R. E. (1996). The effects of sequence context on DNA curvature. Proceedings of the National Academy of Sciences, 93(9), 3847-3852.

[6] Kosikov, K. M., Gorin, A. A., Zhurkin, V. B., & Olson, W. K. (1999). DNA stretching and compression: large-scale simulations of double helical structures. Journal of molecular biology, 289(5), 1301-1326.

[7] YD Nechipurenko, EO Vasilenko, KA Ugryumov, AV Karnaukhov, VG Tumanyan. The allosteric effect of proteins and antibiotics binding on DNA: history and modernity. Molecular Biology, 2026, V. 60. In press.