Ben-Gurion University of the Negev

Computational Modeling of c-Myc-Max Bivalent Heterotetramer Using Alphafold 3

Maor Yehezkehely[a], Michael Povolotskii[a], Oren Ram[b], and David Lukatsky[a]

[a]Department of Chemistry, BenGurion University of the Negev; and [b]Department of Biological Chemistry, The Institute of Life Sciences, The Hebrew University of Jerusalem

Abstract

The c-Myc oncogene is a crucial transcription factor involved in cancer. It functions by forming a heterodimer with other factors, such as Max. Crystallographic studies have revealed that this structure consists of two heterodimers forming a bivalent heterotetramer. Using Alphafold3, for a set of human genome sequences, we confirmed that the c-Myc-Max complex indeed forms heterotetramer structure by looping DNA. Our findings reveal that DNA sequence length constitutes a key feature responsible for bivalent heterotetramer formation.

In our analysis, we utilized DNA sequences derived from MNChIP-Seq data of human embryonic stem cells, specifically those with high binding affinity to the c-Myc transcription factor[3]. We selected 189 genomic sequences containing two E-box motifs (CACGTG), separated by 50-150 base pairs (bp), with a margin of 6-bp on both side.

Each sequence was predicted using AlphaFold3[1] to determine if a bivalent structure, characterized by DNA looping, was formed.

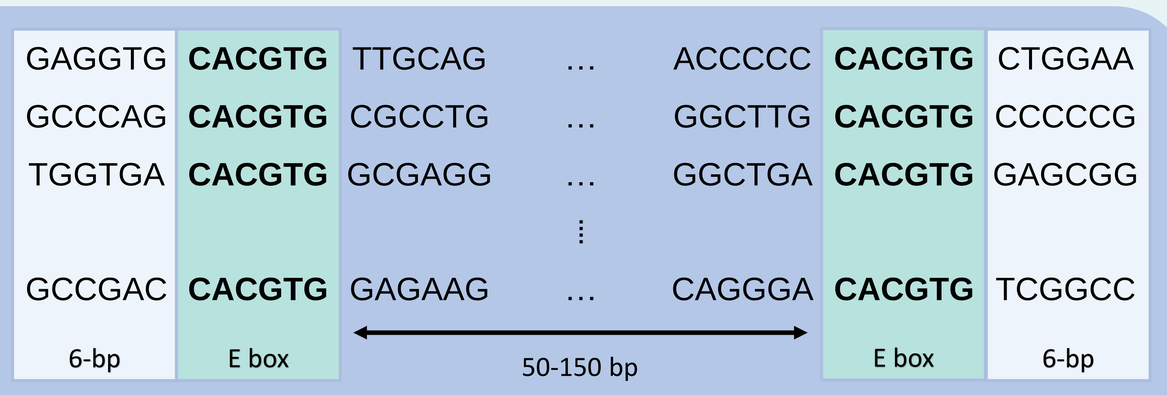


Fig 1 | Examples of the selected DNA sequence

Our analysis revealed that **83.07% (157 out of 189) of the structures form a configuration** resembling the bivalent heterotetramer by looping DNA. The control group was divided into four categories: two peaks (one short and one long) with two types of reshuffling (10 reshuffled sequences containing the binding site and 10 reshuffled sequences lacking it). **It was observed that only 5% (1 out of 20) of the control sequences without the binding motif form DNA loops.**

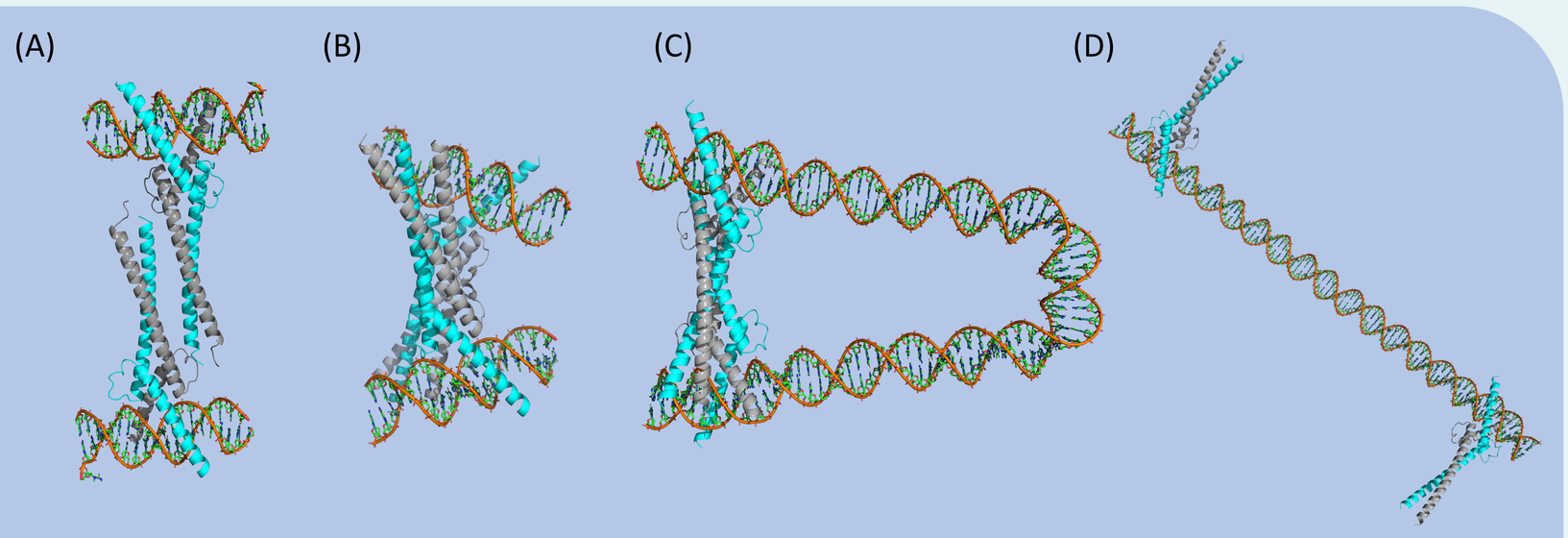
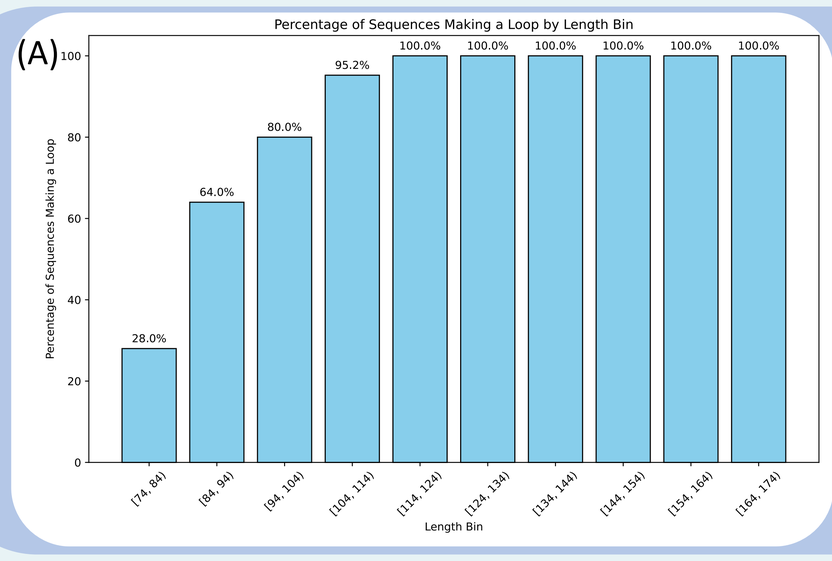


Fig 2| (A) Crystallographic structure ID 1NKP[2] (CGAGTAGCACGTGCTACTC) (B) Alphafold3 prediction of structure ID 1NKP (C) A typical predicted structure by alphafold3 (length 96-bp, 54.16% GC content) (D) A non-bivalent structure predicted by alphafold3 (length 95-bp, 76.84% GC content)



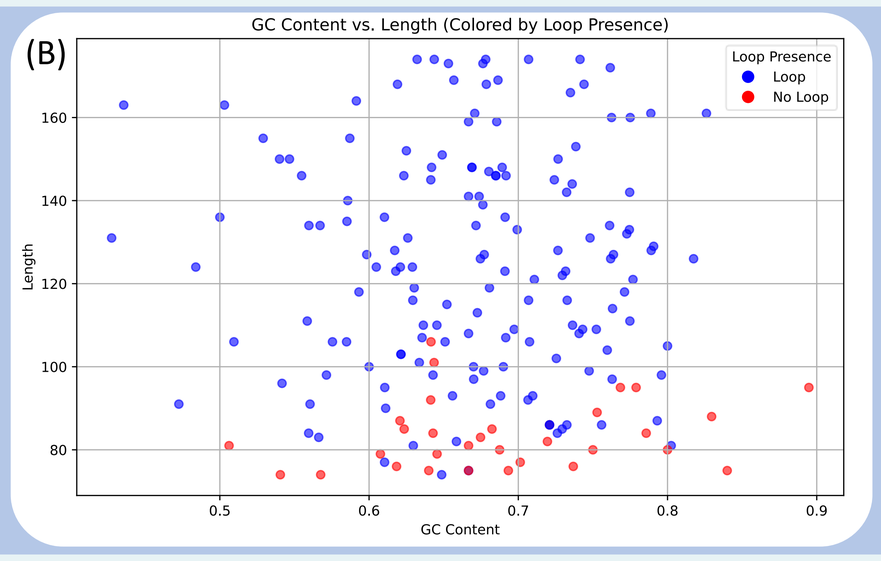


Fig 3 | (A) Percentage of DNA sequences forming loops within 10-bp length intervals. The percentage of looping DNA decreases for shorter sequences. (B) Scatter plot illustrating the relationship between DNA sequence length, GC content, and loop presence (blue points represent looping DNA, red points represent non-looping DNA).

In this study, using AlphaFold3 we reveal that c-Myc-Max forms a bivalent heterotetramer by looping genomic DNA. However, a small fraction of DNA sequences that we tested do not adopt this loop conformation upon c-Myc-Max-DNA binding. We observed that the DNA sequence length has a more significant influence on DNA looping than the GC-content. In particular, the degree of DNA looping decreases for shorter DNA sequences.

References:

1. Abramson, J., Adler, J., Dunger, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
2. Nair SK, Burley SK. X-ray structures of Myc-Max and Mad-Max recognizing DNA. Molecular bases of regulation by proto-oncogenic transcription factors. *Cell* **112**, 193-205 (2003).
3. Tsankov, A., Gu, H., Akopian, V. et al. Transcription factor binding dynamics during human ES cell differentiation. *Nature* **518**, 344–349 (2015).