

Quantitative modeling of transcription initiation by bacterial RNA polymerase

Marko Djordjevic

Mathematical Biosciences Institute, The Ohio State University
email: mdjordjevic@math.ohio-state.edu

BACKGROUND

Bacterial RNA polymerase (RNAP) is a central enzyme in cell, which is responsible for gene transcription. The first stage of transcription is transcription initiation, whose crucial step is the open complex formation during which RNAP binds to promoter and separates the two strands of DNA. Over the last two decades a large amount of data on the open complex formation by bacterial RNAP has been obtained [1]. However, the mechanism by which RNAP forms the open complex is still unknown, and several qualitative hypotheses have been proposed [2]. In order to provide a theoretical framework needed to analyze the assembled experimental data, we developed the first quantitative model of the open complex formation by bacterial RNA polymerase [3]. The model is based on a theoretical biophysics approach, while statistical analysis and bioinformatic methods were used to test the model against available experimental data.

RESULTS

We first show that a simple model (which may follow from recent bioinformatic and experimental results), by which promoter DNA is melted in one step through thermal fluctuations, is inconsistent with experimental data. We next consider a more complex two step model. In this model the transcription bubble forms in the upstream edge of the melted region through thermal fluctuations facilitated by RNAP-DNA interactions, and consequently extends to transcription start site. We derive an explicit connection between the rate of transcription initiation and physical properties of promoter sequence and promoter-RNAP interactions.

We next compare our model with both biochemical measurements and genomics data. We first show that the biochemically measured transition rates from closed to open complex correlate very well with the values predicted from our model ($R^2=0.75$, with the P value of 10^{-3}). Furthermore,

we report even better agreement of the parameters predicted by our model and those inferred from the experimentally confirmed core promoter sequences (correlation constant of $R^2=0.93$, with the P value of 10^{-11}), as shown in Figure 1. No free parameters are used in model testing.

CONCLUSION

The good agreement of our model with the experimental data justifies the quantitative model that we developed, and it furthermore strongly supports the qualitative hypothesis on which the model is based. From a more practical point, our results allow estimating the rate of transition from closed to open complex for a given promoter sequence, which would otherwise require performing quite demanding experimental measurements. This, in turn, allows efficient 'engineering' of promoter sequences with desired kinetic properties. We expect that our model will provide a better practical and conceptual understanding of the transcription initiation process.

REFERENCES

1. Murakami KS, Darst SA (2003) Bacterial RNA polymerases: the whole story. *Curr Opin Struct Biol* 13: 31-39.
2. Young BA, Gruber TM, Gross CA (2004) Minimal machinery of RNA polymerase holoenzyme sufficient for promoter melting. *Science* 303: 1382-1384.
3. Djordjevic M, Bundschuh R. Open complex formation by bacterial RNA polymerase. submitted to *Biophys J* (2007).

Acknowledgements: This work was funded by NSF under Agreement No. 112050.

Figure 1: **Comparison of the model with genomics data.**

Blue dots give parameter values corresponding to contributions to transcription initiation rate due to the presence of a certain base at a certain position in the upstream part of the melted region. Values on the horizontal axis are the parameters inferred from experimentally detected core promoter sequences. Values on the vertical axis are the same parameters predicted from our model. Good correlation between the genomic inferred and the model predicted parameters can be observed. Red dashed curve is the linear fit to the data, and the obtained slope (1.2 ± 0.2) is in a very good agreement with the value of 1 predicted by our model.

