Andrey Cherstvy
ICS-2, FZ Juelich, 52425 Juelich, Germany
e-mail: a.cherstvy@gmail.com
A. Kolomeisky
Rice University, Houston, Texas 77005, USA
A. Kornyshev
Imperial College London, SW7 2AZ, London, UK

Protein-DNA interactions: reaching and recognizing the targets

Search and recognition of targets on DNA by DNA-binding proteins is a vital biological process. Some proteins find their target sequences on DNA with the rates that are 100-1000 times faster than predicted by Smoluchowski diffusion in 3D space. It is often claimed that the reduction of dimensionality from 3D in solution to 1D on DNA is the basic key to understand this facilitated diffusion of DNA-sliding proteins. Recent experiments have shown however that protein diffusion along DNA is often much slower than in solution (see data of Ref. [1] for the lac repressor). Thus, the 3D1D space reduction by itself does not ensure a faster target search. That controversy pushed us to revisit the problem [2].

We present two theoretical models that describe some physical and chemical aspects of protein target search and mechanism of DNA-protein electrostatic recognition. First, we consider the protein target search as a sequence of cycles of 3D diffusion in solution and 1D sliding along DNA. Our non-equilibrium model accounts for protein binding/unbinding to DNA [2]. The model contains a new correlation term, missing in previous theories, that comes from the accurate description of protein diffusion process in stochastic DNA-protein potential. We show that the search time is optimal for an intermediate strength of protein-DNA interactions and intermediate protein concentrations. The fast search is achieved by a parallel scanning of DNA by many proteins. Both conclusions are consistent with the outcomes of recent large-scale Monte Carlo simulations of protein diffusion [3].

Then, we focus on DNA-protein electrostatic interactions, known to give a large contribution to protein-DNA binding affinity. Contrary to hydrogen bonding, electrostatic protein-DNA forces are believed to be largely insensitive to DNA sequence. We show however how the complementarity of charge patterns on target DNA sequence and on a model protein can result in electrostatic recognition of a specific track on DNA. This recognition provokes protein pinning near this homologous region on DNA. We obtain analytical expressions for the shape of the capturing well and typical times proteins spend in it before thermal escape. These times are often long enough to allow a reorganization of the protein structure, so-called interaction-induced protein folding, and formation of stronger (hydrogen) bonds with DNA. One can thus suggest a two-step mechanism for DNA-protein recognition [2]: electrostatically mediated protein sliding and pinning followed by chemical recognition interactions.

This mechanism of protein-DNA recognition is reminiscent of charge adjustment predicted by us for sequence-specific DNA-DNA electrostatic interaction [4]. The

charge complementarity is also known to dominate the formation of many protein-protein complexes in solution [5], rendering such charge zipper complexation pretty general.

Theoretical model of protein-DNA charge recognition has been validated by our recent analysis of real DNA-protein complexes [6]. Structure visualization for many DNA-binding proteins indeed reveals a close proximity of positively charged protein residues (Arg, Lys, and Hist) to negative DNA phosphate groups [6]. A detailed computational analysis of Protein Data Bank files of crystallized DNA-protein complexes performed has indicated several important features. We have observed for instance that in particularly for large structural proteins such as nucleosome core particles, the sequence-specific DNA-protein charge zipper effects are strongly pronounced. Namely, the distribution of Lys and Arg on the protein surface in the vicinity of bound DNA fragment is adjusted to provide a better fit to sequence-specific pattern of DNA phosphates. This indicates sequence-specificity of electrostatic interactions for these complexes, the fact largely overlooked in literature before. Analysis of relatively small DNA-protein complexes, that implement standard motifs of DNA recognition, on the contrary, did not reveal any statistical preference in distribution of positively charged protein amino acids with respect to the contacting DNA phosphates [6,7].

References

[1] R. Austin et al., Phys. Rev. Lett., 97 048302 (2006).
[2] A. G. Cherstvy et al., J. Phys. Chem. B, 112 4741 (2008).
[3] R. K. Das and A. B. Kolomeisky, PCCP, 12 2999 (2010).
[4] A. G. Cherstvy et al., J. Phys. Chem. B, 108 6508 (2004).
[5] A. J. McCoy et al., J. Mol. Biol., 268 570 (1997).
[6] A. G. Cherstvy, J. Phys. Chem. B, 113 4242 (2009).
[7] A. G. Cherstvy, Phys. Chem. Chem. Phys, accepted (2011).