Nonspecific DNA-Protein Interaction: Why Proteins Can Diffuse along DNA
Authors: V. Dahirel, F. Pailluson, M. Jardat, M. Barbi and J.-M. Victor, Phys. Rev. Lett. **102**, 228101(2009)

and

Searching Fast for a Target on DNA without Falling to Traps
Authors: O. Bénichou, Y. Kafri, M. Sheinman and R. Voituriez, Phys. Rev. Lett. **103**, 138102(2009)

Recommended with a commentary by Daan Frenkel, Dept. of Chemistry, U. Cambridge.

One of the key discoveries in molecular biology is the observation that the expression of proteins by DNA is tightly regulated by repressor and promotor proteins. So-called DNA binding proteins (DBPs) can attach to specific target sites close to the gene that codes for a particular protein. When bound, such DNA-binding proteins can either enhance or suppress (one or the other) the translation of the specific gene into RNA. As there are many ($\mathcal{O}(2\ 10^4)$ in the human genome) genes, there are many regulatory motifs and it seems amazing that DBPs can find their target among the millions or even billions of DNA base pairs in the genome of a particular organism.

In the 1980's, Berg and von Hippel showed that the optimal search strategy would be one where the DBPs spend part of the time diffusing along the DNA and part diffusing through space until they rebind on some very different part of the DNA [1, 2]. The optimal ratio of sliding diffusion and through space diffusion is determined by the ratio of the diffusion constants of DBPs along the DNA and through space.

The Berg and von Hippel model has been hugely influential but it does not explain how DNA can perform sliding diffusion: clearly the DBP must be bound to the DNA, but sufficiently weakly that it can easily slide yet sufficiently strongly that it detaches only rarely. Moreover, it must bind strongly to its target.

In a recent article Dahirel et al. present model calculations that may offer an explanation. The work of Dahirel et al. starts with the observation that DBPs are typically concave on the side that faces the DNA. In other words: the DNA double helix fits more or less snugly inside the concave groove of the DBP. In their calculations, Dahirel et al. consider a simplified model

where the DNA is described as a charged cylinder and the proteins as oriented compact objects with various geometrical shapes. The simulations show that the DNA-protein interaction free energy has a minimum at a finite surface-to-surface separation, provided that the part of the protein that faces the DNA is concave with a radius of curvature that is equal (in absolute value) to that of the DNA. The depth of this free-energy minimum and its location depend on the charge of the protein: the lower the surface-charge density of the protein, the weaker the attraction and the larger the equilibrium surface-to-surface distance. The short-range repulsion is due to the accumulation of unbalanced counter-ions in the gap between protein and DNA. The same study makes it also possible to understand specific binding that is due to localised hydrogen bonds. If there is a matching pattern of potential hydrogen-bonding partners, the protein experiences a strong attraction to the DNA (as is to be expected), even though a small barrier must still be overcome to reach the binding site.

Although the model put forward by Dahirel et al. is elegant and appealing, it does not address another mystery associated with the search of DBPs for their target site: the problem is that there are many other sites on the DNA that may bind the DBP fairly strongly, thus slowing down the sliding diffusion. In fact, a naive analysis would suggest that the time it would take a protein to reach its target site is much, much longer than the timescales deduced from experiments on gene regulation.

A possible resolution of this second puzzle was recently put forward by Bénichou et al. This work points out that it is actually misleading to characterise the speed of the search by the mean time it takes a DBP to reach is target, as, in the presence of strong traps, this time may be orders of magnitude larger than the typical search time. This is particularly true in the case where several DBPs are searching for the same target a situation that is to be expected as cells tend to synthesize regulatory proteins in bursts. A human analogy can help to understand this effect. Suppose that a single explorer sets out to search for a hidden treasure. The search is dangerous and the probability that the explorer will be captured or even killed during the trip is appreciable. The average search time is dominated by these game-stopping events and can even be infinite. The same holds if N explorers are sent out on the same quest: as the average search time for every individual is infinite, so is the overall average of the mean first passage time. However and this is the key point of the paper of Bénichou et al. the typical search time (say, the time when there is a 50% probability to have found the target) may be many orders of magnitude shorter than the average. Using realistic values of

the rate of trapping and unbinding of DBPs on non-target sites, Bénichou et al. show that the typical search time to find the correct regulatory sequence is compatible with available experimental data, thus providing a possible resolution of what is known as the speed-stability paradox of target searching by DBPs.

References

[1] O.G. Berg and P.H. von Hippel, J. Biol. Chem **264**, 675(1989)

[2] R.B. Winter, O.G. Berg and P.H. von Hippel, Biochemistry **20**, 6961 (1981)