Conservation of DNA Curvature Signals in Gene Regulatory Regions

Ruy Jauregui[1] **Enrique Merino**[2]
ruy@kuicr.kyoto-u.ac.jp merino@ibt.unam.mx

[1] Bioinformatics Group, Institute of Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan
[2] Biotechnology Institute, National Autonomous University of Mexico, Cuernavaca Mor. 62210, Mexico

Keywords: DNA curvature, transcriptional regulation, orthologous groups.

1 Introduction

DNA curvature has been studied for more than 20 years and has been related to a broad spectrum of biological functions such as DNA replication, transcriptional regulation, and recombination. Studies on the relationship between DNA curvature and transcriptional regulation have been conducted for a relatively small number of genes and discrete *loci*, such as Sigma-54 dependent *glnAp2* and *glnHp2* genes [1], and IHF and CRP regulatory proteins [5, 8]. In this work we extend the previous notions about the biological relevance of DNA curvature as an element of transcriptional regulation by evaluating the conservation of DNA static curvature signals in the regulatory regions of orthologous groups of genes in 99 eubacterial and archaeal genomes.

2 Method and Results

DNA sequence was derived from the complete bacterial genomes available in the Entrez Genome Database. A 250 nt. window, containing 200 bases upstream and 50 bases downstream of the start codon of each coding sequence (CDS), was chosen as our analysis window, since more than 90% of the regulatory signals are found within this range in *E. coli* K12. Operons were predicted based on intergenic distances and genomic context as described by Moreno-Hagelsieb and Collado-Vides [6], and the regulatory region of each gene was considered as the upstream region of the first gene in the operon and was called the Minimal Upstream Regions set, (MURs).

DNA curvature was calculated using the algorithm BEND [2]. A curvature profile was obtained by assigning to each nucleotide of the genome a curvature value, expressed as a deviation angle from the helical axis per helical turn. Since each genome presents a distinctive curvature profile, a cutoff value of 3 SD from the genomic curvature mean of each organism was used to identify statistically significant signals in the set of MURs. Each gene selected this way was sorted into its corresponding orthologous group. Our orthologous groups were mainly those found in the COGs database [9].

Sixty-eight of the 4391 COGs studied presented a statistically significant number of curvature signals (at least 3 SD above the expected mean), including genes coding for DNA binding regulatory proteins such as HU, IHF and FIS, transposases, ribosomal proteins and aminoacil tRNA synthetases, translation factors, cell division, and flagellum biosynthesis among others. These COGs were classified according to their global functional characterization. Experimental data supporting a role of DNA curvature in transcriptional regulation for these genes was searched for in the literature, and in several cases our observations were supported by previous analysis of discrete *loci*.

3 Discussion

Among the relevant groups that were found in our analysis, COGs with proteins coding for the global regulators HU, IHF and FIS were found to present a significant number of genes with curvature signals in a broad spectrum of phylogenetically unrelated organisms, this results are coherent with previous experiments that demonstrate that these proteins bind to curved DNA and are autogenous transcriptional regulators of their own genes [3, 5, 7]. The two subunits of DNA topoisomerase IV (DNA Gyrase) were also found to present conserved curvature signals in their regulatory regions. DNA gyrase is known to be a homeostatic regulator of the nucleus supercoiling state, and to be regulated by FIS [4]. Our finding of several COGs related to cell division and flagellum biosynthesis was unexpected since there is no experimental evidence relating the regulation of these genes with DNA curvature; even tough their transcription is known to be dependent on global regulatory proteins such as HU and FIS. Our work demonstrates the prevalence of curvature signals as conserved elements in transcriptional regulation and extends the previous knowledge described for unique genes in single organisms into a genomic context.

References

[1] Carmona, M. and Magasanik, B., Activation of transcription at Sigma^{54}-dependent promoters on linear templates requires intrinsic or induced bending of the DNA, *J. Mol. Biol.*, 261(3):348–356, 1996.

[2] Goodsell, D.S. and Dickerson, R.E., Bending and curvature calculations in B-DNA, *Nucleic Acids Res.*, 22(24):5497–5503, 1994.

[3] Kohno, K., Wada, M., Kano, Y., and Imamoto, F., Promoters and autogenous control of the *Escherichia coli hupA* and *hupB* genes, *J. Mol. Biol.*, 213(1):27–36, 1990.

[4] Menzel, R. and Gellert, M., Regulation of the genes for *E. coli* DNA gyrase: homeostatic control of DNA supercoiling, *Cell*, 34(1):105–113, 1983.

[5] Miller, H.I., Kirk, M., and Echols, H., SOS induction and autoregulation of the *himA* gene for site-specific recombination in *Escherichia coli*, *Proc. Natl. Acad. Sci. USA*, 78(11):6754–6758, 1981.

[6] Moreno-Hagelsieb, G. and Collado-Vides, J., A powerful non-homology method for the prediction of operons in prokaryotes, *Bioinformatics*, 18(Suppl.1):329–336, 2002.

[7] Ninnemann, O., Koch, C., and Kahmann, R., The *E. coli fis* promoter is subject to stringent control and autorregulation, *EMBO J.*, 11(3):1075–1083, 1992.

[8] Perez-Martin, J. and De Lorenzo, V., Clues and consequences of DNA bending in transcription, *Annu. Rev. Microbiol.*, 51:593–628, 1997.

[9] Tatusov, R.L., Galperin, M.Y., Natale, D.A., and Koonin, E.V., The COG database: a tool for genome-scale analysis of protein functions and evolution, *Nucleic Acids Res.*, 28(1):33–36, 2000.