A new consensus GC-DNA motif for the ancient Smad4 family

SY07-08

M.J. Macias[I,II]

[I]IRB Barcelona, Barcelona, Spain, [II]Icrea, Barcelona, Spain

All activated Smad transcription factors (R-Smads) require Smad4 as a partner for gene regulation, development, and regeneration. The GTCT motif has been considered the main Smad binding site for all R-Smads and Smad4 proteins. Remarkably, Smad proteins also occupy the human genome at GC-rich *cis*-regulatory elements (CREs), lacking GTCT sites. Indeed, several GC motifs have been identified as Smad1/5 binders in the BMP pathway; however, specific GC-sites have not been characterized for Smad4 and for TGF-β activated Smad2/3 proteins. Here, we elucidate the structural basis for the specific and high affinity binding of human Smad4 to new GC-motifs from the *Goosecoid* (*Gsc*) promoter, a mesoendoderm differentiation gene. These sites were narrowed down using EMSA and CRISPR/Cas9 deletion experiments and high-resolution NMR and X-ray crystallography. Binding to the GTCT sites was shown in the past, to interact specifically with 3-bp. Remarkably, using the same DNA-binding region, Smad4 binds to the new sites reading 4-bp. The complex of *Trichoplax adhaerens* Smad4 MH1 domain with the same DNAs suggests that binding to these motifs is conserved along Smad4 evolution. We also discovered that the plasticity of the MH1 domain β-hairpin to recognize several DNA sequences is favored by its conformational flexibility, which provides a basis for understanding the functional adaptability of Smad proteins, and of Smad4 in particular, as a shared mediator in TGF-β and BMP signaling pathways.