Crystallographic Studies of Nucleic Acid Sequence Readout by Proteins

Udo Heinemann

Crystallography, Max Delbrück Center for Molecular Medicine, Berlin, Germany
heinemann@mdc-berlin.de

The faithful readout of DNA or RNA sequences by proteins is of paramount importance in central processes in biology, in particular in gene expression regulation. Transcription factors recognize sequence motifs at enhancer or promoter sites on double-stranded DNA and regulate transcription initiation by binding to these sites. RNA-binding proteins often associate with single-stranded or stem-loop structures in their target, thereby regulating various processes including mRNA translation and half-life or microRNA biogenesis. Here, I will provide an account of recent crystallographic and biochemical studies on these proteins.

Krueppel-like transcription factors (Klf proteins) bind target DNA through their zinc-finger domains. Klf4 aligns its three zinc fingers along the major groove of target DNA, each forming multiple hydrogen-bonded contacts with base pairs in the recognition site [1]. The first Klf4 zinc finger was shown to inhibit cellular self-renewal, whereas the other two zinc fingers are important for induction of macrophage differentiation. The Grainyhead/CP2 family comprises transcription factors, which regulate epithelial differentiation, organ development and skin barrier formation. Human Grainyhead-like (Grhl) 1 has a structure reminiscent of tumor suppressor p53 and binds target DNA with 2:1 stoichiometry, employing a parsimonious DNA sequence readout strategy with few direct protein-base contacts [2]. Our study reveals how tumor-associated mutations inactivate Grhl proteins and contribute to carcinogenesis.

Roquin proteins mediate mRNA degradation by binding to a constitutive decay element (CDE) in the 3'-untranslated region and recruiting an RNA deadenylation complex. They preferentially target mRNAs of proteins acting in the immune system. Crystal structure analysis shows the conserved ROQ domain of Roquin-1 to adopt a winged helix-turn-helix fold and suggest, in combination with mutagenesis and biochemical analyses, that Roquin-1 dimers are involved in binding RNA stem-loop structures [3]. Roquin RNA targets are also recognized and cleaved by ribonucleases of the Regnase (regulatory RNase) family. Recently, we determined the crystal structure of a Regnase-3:RNA complex, revealing for the first time the structure of the unique Regnase CCCH zinc finger and the geometry of its RNA binding [4]. This structure allows a novel classification of CCCH zinc fingers and their interaction with single-stranded RNA. Finally, we studied protein-RNA interactions involved in let7 microRNA biogenesis. Crystallographic and biochemical analyses led us to propose a model of Lin28:pre-let7 binding where initial contacts are mediated by the cold-shock domain of Lin28 by which the RNA structure is remodeled to facilitate subsequent tight binding to the Lin28 zinc-knuckle domain [5].

[1] A. Schuetz, D. Nana, C. Rose, G. Zocher, M. Milanovic, J. Koenigsmann, R. Blasig, U. Heinemann, D. Carstanjen, *Cell. Mol. Life Sci.*, **68**, (2011), 3121-3131.
[2] Q. Ming, Y. Roske, A. Schuetz, K. Walentin, I. Ibraimi, K.M. Schmidt-Ott, U. Heinemann, *Nucleic Acids Res.*, **46**, (2018), 2082-2095.
[3] A. Schuetz, Y. Murakawa, E. Rosenbaum, M. Landthaler, U. Heinemann, *Nat. Commun.*, (2014), **5**:5701.
[4] A. Garg *et al*., unpublished
[5] F. Mayr, A. Schütz, N. Doege, U. Heinemann, *Nucleic Acids Res.*, **40**, (2012), 7492-7506.