

NMR studies on protein-nucleic acid interaction

Robert Kaptein

Published online: 9 May 2013
© Springer Science+Business Media Dordrecht 2013

A special section of this issue is devoted to the study of protein-nucleic acid interaction by NMR. This topic constitutes a fruitful area of research already for several decades. Interactions between proteins and nucleic acids are involved in many essential cellular processes such as replication, transcription, translation, DNA repair, and the regulation thereof. Structural and dynamics studies by NMR have provided crucial information on recognition processes and complex formation between proteins and nucleic acids. In the present collection of articles new data is presented and the authors also give a short review of their work in this area. One of the articles belonging to this collection (Barraud and Allain 2013) was already published in the January issue; it can be found at <http://dx.doi.org/10.1007/s10858-012-9696-4>.

NMR studies on protein-DNA interactions started in the early 80s when synthetic oligonucleotides became available. In 1987 the first structural model of a complex of lac repressor headpiece with a cognate DNA sequence was reported (Boelens et al. 1987). In 1993 atomic resolution NMR structures were obtained for DNA complexes of GATA-1 (Omichinski et al. 1993), Antennapedia homeo-domain (Billeter et al. 1993), and lac repressor headpiece (Chuprina et al. 1993), for reviews see (Jamin and Toma 2001; Campagne et al. 2011). During the mid-90s the first NMR derived protein-RNA complexes were published (Puglisi et al. 1995; Ye et al. 1995; Allain et al. 1996). This area is particularly successful as by now some 150 NMR structures of protein-RNA complexes have been deposited

in the PDB, which amounts to 40 % of these structures in the PDB (for a review see Dominguez et al. 2011).

In this issue, Gervais et al. (2013) review the structural biology of a recently discovered family of DNA-binding zinc finger domains: the Thanathos-Associated Protein (THAP) domains. A new structure is presented for THAP11, a human transcription factor involved in diseases such as colon cancer. The protein is homologous with THAP1 but differs in its DNA-binding properties.

Hennig et al. (2013) address the combined use of NMR with Small Angle X-ray and Neutron Scattering (SAXS and SANS) methods in structure determination of protein-RNA complexes. For large complexes where only sparse NMR data is available, the authors show that both SAXS and SANS, depending on the complex, can be beneficial in refining structures obtained by data-driven docking with HADDOCK (see also van Dijk et al. 2013). The method is illustrated with a ternary complex consisting of two RNA recognition motif (RRM) domains and an 18-mer mRNA control region. Given the complexity of the system the structure elucidation of this ternary complex is clearly a tour de force.

Another combination of methods is discussed by Koharudin et al. (2013). Here it is shown how NMR can help in finding suitable conditions for crystallization of a protein-DNA complex. An NMR analysis showed that for the protein CsrA the best DNA ligand is a stem-loop structure with the consensus GGA triplet in the loop. It was also found that the protein binds two DNA oligomers. Based on the NMR screens The CsrA-DNA complex could be crystallized.

Water molecules play an important role in interfaces of protein-nucleic acid complexes. Van Dijk et al. (2013) have extended their popular docking program HADDOCK to include interfacial water in docking nucleic acids to proteins. Benchmark studies on 30 high-resolution X-ray

R. Kaptein (✉)
Utrecht University, Padualaan 8, 3584 CH Utrecht,
The Netherlands
e-mail: R.Kaptein@uu.nl

structures show that the overall solvation level and water-mediated hydrogen bonds are well reproduced.

Proteins that bind to specific DNA sequences also bind to non-specific DNA. This property is assumed to facilitate the search for specific targets by a variety of mechanisms such as sliding along the DNA, hopping on and off DNA, and intersegment transfer (for proteins that have two DNA-binding sites). Loth et al. (2013) have investigated the sliding process for a lac repressor headpiece dimer. From line-broadening of backbone amides that form H-bonds in the non-specific complex they have determined the 1D diffusion constant for sliding. Curiously, the sliding rate they find is several orders of magnitude smaller than obtained from single molecule fluorescence studies and cannot account for enhanced target location by lac repressor.

Barraud and Allain (2013) address the problem of RNA binding by proteins that have two RRM. A nice overview is given of the various binding modes of these tandem RRM proteins revealed by both NMR and X-ray crystallography. An in-depth study of the prototypal two-RRM domain ribonucleoprotein, human hnRNP A1, shows that the relative orientation of both domains of the protein free in solution is quite different from that found in the crystal structure. In contrast, in complex with RNA a similar orientation is found. Apparently, the structure in the free state is sensitive to crystal packing forces. This study is a good example of the use of segmental isotope labeling in NMR structure determination of multi-domain RNA binding proteins.

It is clear that NMR of protein-nucleic interactions has come a long way since its beginning some 30 years ago. The papers in this collection present some exciting developments in this rapidly evolving field.

References

- Allain FH, Gubser CC, Howe PW, Nagai K, Neuhaus D, Varani G (1996) Solution structure of the N-terminal RNP domain of U1A protein: the role of C-terminal residues in structure stability and RNA binding. *Nature* 380:646–650

- Barraud P, Allain FH (2013) Solution structure of the two RNA recognition motifs of hnRNP A1 using segmental isotope labeling: how the relative orientation between RRM domains influences the nucleic acid binding topology. *J Biomol NMR* 55:119–138
- Billeter M, Qian YQ, Otting G, Müller M, Gehring W, Wüthrich K (1993) Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex. *J Mol Biol* 234:1084–1094
- Boelens R, Scheek RM, van Boom JH, Kaptein R (1987) Complex of lac repressor headpiece with a 14 base-pair lac operator fragment studied by two-dimensional nuclear magnetic resonance. *J Mol Biol* 193:213–216
- Campagne S, Gervais V, Milon A (2011) Nuclear magnetic resonance analysis of protein-DNA interactions. *J R Interface* 8:1065–1078
- Chuprina VP, Rullmann JAC, Lamerichs RMJN, van Boom JH, Boelens R, Kaptein R (1993) Structure of the complex of lac repressor headpiece and an 11 base pair half-operator determined by NMR Spectroscopy and Restrained Molecular Dynamics. *J Mol Biol* 234:446–462
- Dominguez C, Schubert M, Duss O, Ravindranathan S, Allain FH (2011) Structure determination and dynamics of protein-RNA complexes by NMR spectroscopy. *Prog NMR Spectrosc* 58:1–61
- Gervais V, Campagne S, Durand J, Muller I, Milon A (2013) NMR studies of a new family of DNA binding proteins: the THAP proteins. doi:10.1007/s10858-012-9699-1
- Hennig J, Wang I, Sonntag M, Gabel F, Sattler M (2013) Combining NMR and small angle X-ray and neutron scattering in the structural analysis of a ternary protein-RNA complex. doi:10.1007/s10858-013-9719-9
- Jamin N, Toma F (2001) NMR studies of protein-DNA interactions. *Prog NMR Spectrosc* 38:83–114
- Koharudin LMI, Boelens R, Kaptein R, Gronenborn AM (2013) A NMR guided approach for CsrA-RNA crystallization. doi:10.1007/s10858-013-9712-3
- Loth K, Gnida M, Romanuka J, Kaptein R, Boelens R (2013) Sliding and target location of DNA-binding proteins: an NMR view of the lac repressor system. doi:10.1007/s10858-013-9723-0
- Omichinski JG, Clore MG et al (1993) NMR structure of a specific DNA complex of Zn-containing DNA-binding domain of GATA-1. *Science* 261:438–446
- Puglisi JD, Chen L, Blanchard S, Frankel AD (1995) Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science* 270:1200–1203
- Van Dijk M, Visscher KM, Kastriitis PL, Bonvin AMJJ (2013) Solvated protein-DNA docking using HADDOCK. Accessed 20 April 2013
- Ye X, Kumar RA, Patel DJ (1995) Molecular recognition in the bovine immunodeficiency virus Tat peptide TAR RNA complex. *Chem Biol* 2:827