MSc Barbara Szutkowska

Doctoral thesis abstract *„Influenza A virus vRNA secondary structure in biological environment”*

Abstract

As of today, four types of Influenza viruses were characterized – A, B, C and D. Due to epidemic and pandemic potential, zoonotic Influenza A virus (IAV) strains are considered the most human-threatening. The phenomena of the antigenic shift and drift underlie the virulence of the IAV and contribute to rapid genetic evolution. Fast evolution is supported by a small (13 kbp) genome, which consists of single-stranded RNAs (vRNA) called segments. In the case of the Influenza A virus, the genome consists of the 8 single-stranded vRNAs, which along with the viral polymerase complex and multiple copies of stabilizing NP protein (nucleoprotein) form ribonucleoprotein complex (RNP). The genomic vRNAs are of negative-sense. This means that during the viral replication cycle, the vRNAs are replicated into complementary RNAs (cRNA), which are the template for the synthesis of new vRNAs. The vRNAs are also templates for the transcription of mRNAs which are translated into viral proteins.

RNA function depends on its secondary structure and the possibility of interactions with biomolecules such as RNA-RNA or RNA-protein interactions. In the case of influenza virus vRNA, a number of structural motifs have been described that perform function at different stages of the IAV lifecycle. The disturbance of these motifs contributed to the reduction of viral proliferation. Although we have knowledge of the *in vitro* folding of selected IAV vRNA segments and *in virio mapping* data for some strains have been published, we still lack the full knowledge of the vRNAs secondary structure folding in natural environment.

The main objectives of the research described in this doctoral thesis was to determine the vRNAs secondary structures in the biological environment: in cellular lysates (*in vivo-like*), in IAV-infected cells (*in cellulo*), and in virion (*in virio*). The pandemic strain A/California/04/2009 (H1N1) responsible for pandemics in the years 2009-2010 was used for the research.

In the first place, the vRNA8 structure in the lysate of IAV-infected MDCK (Madin Darby Canine Kidney) cells was determined. The RNAStructure program and experimental data from chemical mapping (DMS, 1M7) were used for the secondary structure prediction. The global structure, as well as local structure with maximum pairing distance, were generated during the prediction. This enabled determination of the impact of limited base-paring distance on the

secondary structure prediction and identification of structural motifs forming regardless of the parameters used. Next, the base-pairing conservation (based on structural-sequence analysis of all IAV strains) of vRNA8 structure was calculated. The analysis revealed structural motifs with high structural conservation among different IAV strains.

The research focused on the secondary structure of vRNAs *in virio* and *in cellulo* was described in the second part of this doctoral thesis. In the first place, the methods of chemical probing of vRNAs in virion as well as IAV-infected human lung cells (A549) were established. Another essential step was the development of NGS library preparation protocols. The sequencing results enabled the calculation of single-nucleotide reactivities to chemical reagents (NAI, DMS) for the full-length of all vRNA segments *in virio* as well as the full-length vRNA5, 7, and 8 *in cellulo*. The results obtained for three independent biological replicates showed a very high correlation, which proved the reproducibility of the results. Based on the sequencing results, common RNA structural motifs were revealed for *in virio* and *in cellulo* structures. The obtained results were also compared to the known *in virio* and *in vitro* vRNA secondary structures of other IAV strains, which allowed to indicate local structural motifs common to many IAV strains. The research revealed new structural motifs in certain vRNA segments as well.

Next, the vRNA8 secondary structures determined *in virio* and *in cellulo* were compared with the secondary structure obtained in cellular lysate. The comparison of the secondary structures showed that some of the motifs were common to all three analyzed vRNA8 structures. Importantly, some of the vRNA8 motifs predicted in cell lysates were unique to either the *in virio* secondary structure or *in cellulo* structure. This indicated that the structure in lysate was an intermediate structure between *both structures*.

The MFE (Minimum Free Energy) secondary structures of vRNA5, 7 and 8 in *in virio* and *in cellulo* conditions differ in global folding. Most differences were observed in the long-range base-pairing between distant nucleotides. These changes suggest high structural dynamics of RNA in the biological environment. Still, some structural motifs, especially local, were preserved probably due to their functional significance.

The research described in this doctoral thesis enables a very wide analysis of IAV vRNA secondary structures in the different biological environments. For the first time, vRNA structures were determined under *in cellulo* conditions and also compared with vRNA structures under other biological conditions. This research contributed to a better understanding of viral biology from the IAV genomic RNA point of view. Described conserved vRNA structural

motifs may be an ideal target for universal, strain-independent inhibition of Influenza A virus at the RNA level.