Title: **Selected aspects of the interactions between the human Dicer ribonuclease and nucleic acids**

Author: Agnieszka Szczepanska

Abstract

Dicer ribonucleases belong to the Ribonuclease III (RNase III) family, which is a group of endoribonucleases that are specific for double-stranded RNAs (dsRNAs). Dicers are mainly known for their important role in the biogenesis of small regulatory RNAs (srRNAs), such as microRNAs (miRNAs) and small interfering RNAs (siRNAs). Recently, there has been an increasing number of reports describing Dicer activities unrelated to its RNase activity and srRNA production. This dissertation is devoted to the human Dicer ribonuclease (hDicer). hDicer, like other vertebrate Dicer proteins, is a multi-domain protein composed of an N-terminal helicase domain, a domain of unknown function 283 (DUF283), Platform, PAZ domain, a connector helix domain, two RNase III domains (RIIIa and RIIIb), and a C-terminal dsRNA-binding domain (dsRBD). The region spanning the Platform–PAZ–Connector helix domains is called the PPC cassette, and it plays a key role in the recognizing and anchoring canonical substrates of Dicer, such as miRNA precursors (pre-miRNAs) and siRNA precursors (pre-siRNAs).

Previous studies conducted at the Institute of Bioorganic Chemistry Polish Academy of Sciences have revealed that the DUF283 domain of hDicer binds single-stranded RNAs and DNAs but not the double-stranded RNA and DNA substrates. Interestingly, further studies showed that the DUF283 domain can promote base pairing of complementary nucleic acid molecules. Subsequent studies revealed that the full-length hDicer also exhibits such activity, thus acting as a nucleic acid annealer. Based on the results of these studies, it was hypothesized that the DUF283 domain is crucial for the annealing activity of hDicer. To verify this hypothesis, two variants of hDicer, lacking the amino acid sequence encoding the DUF283 domain, were created: (i) the ΔDUF(630-709) variant and (ii) the ΔDUF(625-752) variant. The ΔDUF(625-752) variant, in addition to the deletion of the DUF283 domain, also lacked amino acids flanking this domain. The analyses showed that in vitro, the ΔDUF(630-709) variant presented similar RNase activity towards the pre-miRNAs and pre-siRNAs used in the study, as the wild-type hDicer. In contrast, the ΔDUF(625-752) variant showed significantly weaker RNase activity towards the same substrates. Likewise, in cellulo studies of the ΔDUF(625-752) variant demonstrated that this variant produced the selected miRNAs with lower efficiency, as compared to the ΔDUF(630-709) variant, or hDicer. However, both variants did not facilitate base pairing of complementary nucleic acids. The obtained data strongly support the hypothesis that the DUF283 domain is crucial for the nucleic acid annealing activity of hDicer.

In the cell, Dicer ribonucleases can bind not only pre-miRNA and pre-siRNA substrates, but also other RNAs, e.g., mRNAs, or long non-coding RNAs. Given this fact, another research was conducted to determine the role of the DUF283 hDicer domain in the binding of cellular RNAs. In this study, the irCLIP-seq (infrared crosslinking immunoprecipitation followed by NGS sequencing) approach and human embryonic kidney cells (HEK 293T) were used, as well as DICER1 knock-out cell lines producing: (i) the ΔDUF(630-709) variant, (ii) the ΔDUF(625-752) variant, or (iii) the wild-type hDicer (rescue control). The collected results revealed that

the RNA pools bound by the ΔDUF hDicer variants and the wild-type full-length hDicer were different. These results suggest that the DUF283 hDicer domain is involved in the recognition and binding of a specific pool of cellular RNAs. Interestingly, the collected data revealed that the pool of RNAs bound by hDicer included RNAs rich in guanosine (G) tracks. Bioinformatics analyses showed that most of these RNAs have a potential to adopt G-quadruplex structures. Consequently, another research goal was to investigate the interactions between hDicer and nucleic acid molecules adopting G-quadruplex structures. The collected results showed that hDicer binds both RNA G-quadruplexes (G4-RNA) and DNA G-quadruplexes (G4-DNA), and that this binding presumably occurs within the PPC cassette of hDicer. Based on the results of the molecular modeling, it was hypothesized that pre-miRNAs and molecules adopting the G-quadruplex structures anchor within the same region of the PPC hDicer cassette. Indeed, when G4-RNA or G4-DNA was added to the reaction mixture, inhibition of pre-miRNA processing by hDicer was observed. The knowledge gained during these studies expands our understanding of the molecular basis of interactions between Dicer ribonucleases and RNA or DNA molecules adopting G-quadruplex structures.

In conclusion, the results of the experimental research conducted in the course of this dissertation support observations that the role of Dicer ribonucleases may extend far beyond the srRNA biogenesis pathways.