Anna Więch-Walów

Molecular analysis of the F domain of ecdysteroid receptor from Aedes aegypti.

Summary:

The functioning of multicellular organisms requires the precise control and coordination of many processes that take place in them. Nuclear receptors (NRs) are an example of proteins that regulate the transcription of genes at the molecular level and are important for the processes of development or the maintenance of homeostasis. NRs, a group of transcription factors (TFs), act dependent on the binding of signaling molecules and their disturbed function results in a number of undesirable effects.

The structure of NRs includes the N-terminal domain (NTD) - most often exhibiting features of intrinsically disordered regions (IDRs) and being responsible for transcription transactivation, the DNA binding domain (DBD), a flexible linker – hinge region and a ligand binding domain (LBD). Some NRs, at their C terminus, have F domain. The current knowledge about the structure and function of F domains is rather residual. Published results demonstrated, however, its involvement in LBD signaling molecules recruitment, an effect on the recruitment of transcriptional corepressors and coactivators, and facilitating the dimerization of NRs.

Originating from *Aedes aegypti*, the ecdysteroid receptor (EcR) has one of the longest amino acid sequences corresponding to the F domain (AaFEcR). Due to the significant role of AaEcR in the reproductive process of *A. aegypti*, the structure and function of domains that precede AaFEcR have been an area of concern before. By interacting with Ultraspiracle (Usp), another member of the NRs superfamily, AaEcR forms a functional ecdysteroid receptor (Usp/EcR). Usp/EcR is responsible for the control of the expression of the gene encoding vitellogenin (Vg) - early yolk protein, that is accumulated in mosquito oocytes. In the context of the *A. aegypti* involvement in the transmission of dengue or Zika fever viruses, detailed knowledge of the molecular mechanism of AaFEcR action may be the key to a controlled reduction of the mosquito population.

The aim of this study was the molecular analysis of the F domain from EcR (AaFEcR), that has not been characterized previously. To realize it, an efficient procedure of overexpression of the recombinant AaFEcR in fusion with the TF (trigger factor) chaperonin protein was developed in order to obtain its homogeneous preparation. For AaFEcR, *in silico* analyzes were performed, showing that the isolated AaFEcR exhibits features of intrinsically disordered proteins (IDPs). Lack of a stable tertiary structure under physiological conditions of AaFEcR has been confirmed by the analysis of the content of secondary structures using circular dichroism (CD) technique. The obtained results allowed to determine that in the AaFEcR structure there are both: disordered structures that can undergo structuring under the conditions of induced folding and globular structures - susceptible to chemical denaturation. The above features allowed to determine that AaFEcR has the conformation of a pre-molten globule (PMG). The analysis of the hydrodynamic properties of AaFEcR initially suggested that it undergoes oligomerization. However, the experiment of size-exclusion chromatography (SEC) and analytical ultracentrifugation (SV-AUC) showed that its high value of Stokes radius (R_S) is due to belonging to IDPs and the protein functions as an elongated monomer.

Subsequent *in silico* analyzes of the AaFEcR sequence revealed the presence of two distinctive motifs (HGPHPHPHG and QQLTPNQQQHQQQHSQLQQVHANG) that had previously been described as being able to coordinate metal ions. The potential ability of AaFEcR to bind

Zn^{2+}, Cu^{2+} and Ca^{2+} ions was investigated using CD, SEC, SV-AUC, mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy. AaFEcR has been shown to be capable of interacting with one or two Zn^{2+} ions, one, two or three Cu^{2+} ions and one Ca^{2+} ion interacting with the carboxyl terminus of the AaFEcR or carbonyl group in the molecule. Binding of the above ions does not result in change in the content of secondary structures in AaFEcR but a decrease in the volume of its molecule. Coordination of Zn^{2+} ions by full-length AaFEcR most likely takes place within the amino acid motif of HGPHPHPHG as evidenced by the observation of a slight chemical shift of the 74. glycine residue by NMR technique, but it does not result in changes to the tertiary structure of AaFEcR. Thus, the coordination of ions by AaFEcR should take place at the level of single amino acid residues, resulting in its molecule compensation.

Studies with the use of synthetic Ac-HGPHPHPHG-NH_2 and Ac-QQLTPNQQQHQQQHSQLQQVHANG-NH_2 peptides have shown that both motifs are able to coordinate Zn^{2+} and Cu^{2+} ions through nitrogen atoms from imidazole rings of the histidine residues. Due to the high content of glutamine residues, capable of forming hydrogen bonds in the Ac-QQLTPNQQQHQQQHSQLQQVHANG-NH_2 peptide, the interactions were considered more stable. The unique discovery, however, was the coordination of the Cu^{2+} ion by Ac-HGPHPHPHG-NH_2 peptide, resulting in the formation of a unique poly-l-proline type II (PPII) helix.