Maciej Baranowski - Abstract

The subject of my thesis is the family of molecular chaperones called Hsp40 - a large, diverse family of proteins present in and essential for nearly any cell. These nearly ubiquitous, in most cases essential proteins are at the frontline of cellular proteostasis, recognizing unfolded polypeptides, protecting them and transferring to their Hsp70 partners, for either refolding or preparation for other chaperone or proteolytic pathways. Despite the importance of Hsp40 proteins, their mechanism of action remains poorly described, especially concerning substrate recognition and binding.

Substrate recognition and binding is known to occur on the C-Terminal Domain 1 (CTD1) of Hsp40. The CTD1 is a small (~80 residues) domain composed of two β-sheets and sometimes compared to the Immunoglobulin-Like Fold (ILF) domains. I noticed that arrangement of β strands in CTD1 is very unusual, looking like a fusion of two structural motifs: the Greek key and the psi loop. Since such arrangement is very characteristic and easy to recognize, I decided to check if there are other, non Hsp40, proteins, which would contain domains similar to CTD1 in terms of arrangement of secondary structure elements. I found that ILF is the only (speaking in CATH terminology) family of folds similar to CTD1, which led me to hypothesize that CTD1 and ILF families are directly evolutionarily related. Using only criterion of parsimony, I created a scenario describing how this transition might have had happened. This scenario in turn, when combined with analysis of multiple sequence alignments and careful comparison of CTD1 and ILF structures, suggested new details about the mechanism of substrate binding by CTD1: i) the substrates of Hsp40 are probably around 10 to 12 residues in length in contrast to 7-8 residue model peptides typically used for Hsp40 research; ii) the substrates are symmetric, composed of central part and two flanks; iii) the central part is characterised by at least two hydrophobic residues pointing in roughly the same direction; iv) the flanks are negatively charged.

For my next step, I performed extensive molecular dynamics simulations of an example of the CTD1 domain. Results of molecular dynamics simulations led to following observation: i) the two hydrophobic clefts on CTD1 which serve as a docking site for hydrophobic residues on the substrate strongly fluctuate, possibly allowing CTD1 to adopt to hydrophobic side chains of different sizes; ii) on one of the sides, the two β-sheets forming CTD1 can partially dissociate. This dissociation is driven by water separating conserved polar residues in CTD1.
Finally, I made an attempt to test my findings in *in vitro* experiments, but I failed to obtain a pure sample of the Hsp40 protein I selected. My findings bring a new light to the mechanism of substrate recognition and binding of the Hsp40 proteins, which are a central hub of the chaperone network. They strongly suggest that polypeptide strands recognized by Hsp40 are nearly identical

to strands recognized by their partners, the Hsp70 - with the exception that the ends of the strands are negatively charged for Hsp40 and positively charged for Hsp70. My results suggest new direction for experiments (both *in vitro* and crystallographic) on Hsp40 proteins, experiments which would deepen our understanding of the Hsp70/40 system.