Characterization of transglutaminase 2 substrate specificity using phage display technology, logistic regression analysis and intrinsic disorder examination

Éva Csősz, Department of Biochemistry and Molecular Biology, University of Debrecen
Supervisor: Prof. Dr. László Fésüs
Molecular and Cellular Immunbiology Graduate School, Director: Prof. László Fésüs

Transglutaminase 2 (TG2) catalyzes the Ca^{2+}-dependent post-translational modification of proteins via formation of isopeptide bonds between their glutamine and lysine residues. The enzyme has more than 130 reported substrates but the exact mechanism by which its substrates are selected is still an enigma. As a first approach, we collected the known transglutaminase substrates into TRANSDAB Wiki (http://genomics.dote.hu/wiki), the transglutaminase substrate database and using the deposited information we attempted to find out the rules of TG2 substrate selection.

To study the preferred sequences around substrate glutamines we adapted the phage display technique selecting the glutamine donor substrates from a random heptapeptide library via their binding to recombinant TG2. The pQx(P,T,S)l consensus motif around glutamines was established, which is consistent with so far identified substrates. Database searches showed that several proteins contain peptides similar to the phage-selected sequences, and the N-terminal glutamine-rich domain of SWI1/SNF1-related chromatin remodeling protein p270 was chosen for detailed analysis. Mass spectrometry-based studies of a representative part of the SWI1/SNF1-related chromatin remodeling protein indicated that it was modified by TG2. Along with phage display technique *in silico* methods were used to compare the sequence context of substrate and non substrate residues to get a better understanding about principles of substrate selection of TG2. None of the results could give a full explanation how TG2 selects the different substrate glutamine and lysine residues.

Using the structural information on TG2 substrate proteins listed in TRANSDAB Wiki database a slight preference of TG2 for glutamine and lysine residues situated in turns could be observed. When the spatial environment of the favored glutamine and lysine residues were analyzed with logistic regression the presence of specific amino acid patterns were identified. Using the occurrence of the predictor amino acids as selection criteria several polypeptides were predicted and later identified as novel *in vitro* substrates for TG2. Studying the sequence of TG2 substrate proteins lacking available crystal structure the strong favorable influence on substrate selection of the presence of substrate glutamine and lysine residues in intrinsically disordered regions also could be revealed.

The collected sequence and structural data have provided novel understanding of how this versatile enzyme selects its substrates in various cell compartments and tissues and suggest that instead of the strict linear sequences spatial features must be considered as well to explain the complex physico-chemical interaction between TG2 and its substrates. It seems that in case of this enzyme a divergent substrate recognition system has evolved where beside the linear sequences, spatial structural features and the presence of intrinsic disorder can be significant in substrate selection. This may reflect the unique nature of how transglutaminase 2 works in almost all cellular compartments, including the cell surface and extracellular space. It is capable to perform diverse biochemical reactions, such as signal transduction through its GTPase activity, ATP hydrolysis, protein disulphide isomerase activity, integrin and fibronectin binding, while its major biochemical function is modifying protein bound glutamine residues whenever it becomes feasible. The need of substrate selection for this classical transglutaminase function may arise under very different circumstances making the flexible recognition mechanisms detailed in this work advantageous.

Keywords: TRANSDAB Wiki, phage display technology, logistic regression analysis, substrate preference, intrinsic disorder, three dimensional structure
Kulcsszavak: TRANSDAB Wiki, fág bemutató rendszer, logisztikus regressziós analízis, szubsztrátpreferencia, szerkezeti rendezetlenség, térszerkezet