PROTEIN CHEMISTRY AND MASS SPECTROMETRY IN STRUCTURAL ANALYSIS OF LARGE PROTEINS

Karel Bezouška[1,2], Petr Man[1,2], Petr Pompach[1,2], Petr Novák[1,3], Vladimír Havlíček[2], Rolf Hilgenfeld[4], Ondřej Plíhal[1,2], Jan Sklenář[1,2], Andrea Pišvejcová[2], Vladimír Křen[2]

[1]Department of Biochemistry, Faculty of Science, Charles University Prague, Hlavova 8, CZ-12840 Praha 2, Czech Republic;
[2]Institute of Microbiology, Academy of Sciences of Czech Republic, Vídenská 1083, CZ-14220 Praha 4, Czech Republic;
[3]Sandia National Laboratories, 7011 East Avenue, Livermore, CA, 94550, USA;
[4]Institute for MolecularBiotechmology,Beutenbergstrasse 11, D-07745 Jena, Germany.
Correspondence to bezouska@biomed.cas.cz

According to the textbook knowledge the use of automated Edman degradation and protein mass spectrometry for the determination of complete primary structure is limited to small proteins with molecular size less than 10 kDa, larger proteins being analyzed by DNA sequencing of the corresponding genes or cDNA clones. Here we present two examples of our recent work in which the sequence of rather large proteins has been determined completely or nearly completely by protein sequencing and mass spectrometry. The first example includes the pokeweed antiviral protein from *Phytolacca acinosa* (PAP-S) that belongs to the family of type-1 ribosome-inactivating proteins. The purified PAP-S proteins resolve on SDS electrophoresis into two closely related bands with Mr of about 29 kDa. Interestingly, the upper protein band, PAP-S_{up} has been shown to crystallize through carbohydrate-protein interactions based on a rare type of N-glycosylation, namely N-linked GlcNAc monosaccharide substitutions at the canonical Asn-Xxx-Ser/Thr [1]. The sequence of PAP-S_{up} is not known from the genetic data, but is essential for unambiguous solving of the crystal structure in positions that cannot be called directly from the electron density. We have thus determined the complete structure of PAP-S_{up} by Edman degradation of N-terminal and internal peptide sequences in combination with MALDI peptide mapping and tandem mass spectrometry using an ion trap. The complete sequence has 261 amino acids and includes three sites of the above N-glycosylation. The sequence coverage was 92 % by Edman degradation data, 93 % by peptide mapping and 90 % by tandem MS data. The second example is -N-acetylhexosaminidase from *Aspergillus oryzae* CCF1066, a robust extracellular secreted enzyme used in enzymatic syntheses of oligosaccharides and biotechnology [2]. This enzyme has 600 amino acids (including 6 cysteins and 6 sites of N-glycosylation), of which 466 has been verified by direct analysis of the protein (sequence coverage 77 %). Identification of large N-terminal segment in the protein proved difficult pointing to the fast cleavage of this protein segment. Enzyme is composed of cleaved signal peptide, the propeptide sequence involved in regulated secretion, the inactive zincin domain, and the catalytical domain belonging to family 20 of glycohydrolases.

Supported by grant from Ministry of Education of Czech Republic (MSM 113100001), by Institutional Research Concept No. A0Z5020903 for the Institute of Microbiology, and by grant from the Grant Agency of Czech Republic (203/01/1018).

1. T. Hogg, I. Kutá-Smatanová, K. Bezouška, N. Ulbrich, R. Hilgenfeld, *Acta Cryst.*, **D58** (2002) 1734-1739.
2. Z. Hunková, V. Křen, M. Ščigelová, L. Weignerová., O. Scheel, J. Thiem. *Biotechnol. Lett.* **18** (1996) 725-730.

DYNAMICS OF 1,2-DICHLOROETHANE IN THE ACTIVE SITE OF HALOALKANE DEHALOGENASE LINB: EFFECT OF SOLVENT AND HALIDE ION ON PRODUCTIVE BINDING

Michal Otyepka[1] and Jiří Damborský[2]

[1]Department of Physical Chemistry, Faculty of Science, Palacký University, tř. Svobody 26, 771 46 Olomouc, Czech Republic
[2]National Centre for Biomolecular Research, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic

1,2-dichloroethane (DCE) is a toxic and carcinogenic chlorinated compound that is not known to be formed naturally. As many other synthetic halogenated aliphatic compounds, DCE is rather resistant to biodegradation and persists in the environment. Nevertheless, several bacterial cultures that are able to use DCE as the only carbon and halogen source have been isolated. The most efficient catalysis of DCE has been observed with haloalkane dehalogenase DhlA from *Xanthobacter autotrophicus* GJ10. Even lower activity with DCE was observed for haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26. Crystallographic analysis of LinB-DCE complex showed non-productive binding of DCE to the enzyme active site, while molecular docking suggested that DCE molecule can possibly bind to the active site but is prevented by chloride ion and/or water molecules [1].

Two nanoseconds-long trajectories of LinB with different number of ligands bound to the active site were carried out and compared. The results show that productive binding of DCE (i.e., binding to the Michaelis-Menten complex) is blocked by the presence of chloride ion or water molecule in the halide-stabilization pocket of the active site. In case of empty halide-stabilization pocket, the productive binding of DCE occurs very rapidly (in less then 20 ps). On the other hand, DCE locks the chloride ion in the halide-stabilization pocket as was confirmed by steered molecular dynamics simulations and by fact that chloride ion can easy leave LinB active site in the system without DCE.