33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Evaluating descriptors for the lateral translocation of membrane proteins

Olga Domanova[1,2,3], Stefan Borbe[1], Stefanie Mühlfeld[4], Martin Becker[4], Ralf Kubitz[4], Dieter Häussinger[4], Thomas Berlage[1,2,3]

Abstract— **Microscopic images of tissue sections are used for diagnosis and monitoring of therapy, by analysis of protein patterns correlating to disease states. Spatial protein distribution is influenced by protein translocation between different membrane compartments and quantified by comparison of microscopic images of biological samples. Cholestatic liver diseases are characterized by translocation of transport proteins, and quantification of their dislocation offers new diagnostic options. However, reliable and unbiased tools are lacking. The nowadays used manual method is slow, subjective and error-prone. We have developed a new workflow based on automated image analysis and improved it by the introduction of scale-free descriptors for the translocation quantification. This fast and unbiased method can substitute the manual analysis, and the suggested descriptors perform better than the earlier used statistical variance.**

I. INTRODUCTION

Not only quantitative, but also spatial regulation is of an utmost importance for the function of individual cells and organisms. The term toponome describes the temporal and spatial organization of biological molecules, in particular proteins, within the structures of the organism, mainly on the level of cellular, subcellular and supercellular structures [1]. The toponomics analysis (topological proteomics or location proteomics) focuses on the underlying laws of this spatial arrangement [2], [3].

Spatial relations can be captured in different dimensionalities. Translocation studies are based on the quantitative assessment of marker concentrations in bounded regions (e.g., nucleus-to-cytoplasm, see [4]). The considered molecular markers can be classified as functional or structural, with the latter defining reference structures and the former being the object of interest in terms of spatial relations. Protein co-localization studies can be based on measuring and evaluating isotropic distributions of distances between pixels [1], [5]. If the former is described as a 2-dimensional analysis, the latter should be regarded as 0-dimensional. The method described here is based on 1-dimensional modeling (orthogonal section to a membrane segment).

This work was supported in part by B-IT Research School. The work was also supported by KFO 217, DFG

[1] Fraunhofer Institute for Applied Information Technology FIT,Schloss Birlinghoven, 53754 Sankt Augustin, Germany

[2] RWTH Aachen, Department of Computer Science, Ahornstraße 55, 52074 Aachen, Germany

[3] Bonn-Aachen International Center for Information Technology B-IT, Dahlmannstraße 2, 53113 Bonn, Germany

[4] Heinrich-Heine University of Düsseldorf, Medical Faculty, Department of Gastroenterology, Hepatology and Infectiology, Moorenstr. 5, 40225 Düsseldorf, Germany

Here we suggest an improvement of the automated method for the toponomics analysis [6]. Translocation of membrane proteins is quantified by a fast and objective method which can substitute the manual evaluation.

A. Membrane Protein Translocation Analysis

Relative quantification of subcellular protein translocation by fluorescence intensity analysis in microscopy images, either manually or automatically, is a widely-used approach. Samples for image acquisition are prepared according to standard protocols. Tissue samples are cut into thin sections (ca. 5 μm) and proteins are fixed. Structural and functional marker proteins are labeled with fluorochromes by means of immunohistochemistry. By confocal fluorescence microscopy, an advanced optical sectioning technique, images of thin slices (ca. 500 nm) of a thick specimen can be acquired [7] to reveal spatial information about proteins in a defined region. Fluorescence intensities are directly proportional to the protein concentrations providing thus valuable basis for the analysis.

B. Biological Model

We used the liver as a biological model for the improvement of the translocation analysis. It is a multifunctional organ with a prominent role in metabolism [8] and has numerous essential functions. Among them, detoxification, protein synthesis and formation of bile required for the digestion are of a great importance.

Hepatocytes, the most abundant cell population of the liver [9], generate bile flow within bile canaliculi by continuous secretion of bile salts and other solutes across the canalicular membrane. The production and export of bile influences digestion, and thus the normal functioning of an organism. Displacement of transport proteins from canalicular membranes [10] leads to the decreased export of bile and may result in cholestatic liver diseases [9]. The short-term regulation of canalicular transport proteins is determined by rapid endo- and exocytosis of bile salt export pump (Bsep)-bearing vesicles from and into the cell membrane [11]–[13]. Liver perfusion in rats demonstrated that hyperosmolarity leads to rapid retrieval of Bsep from the canalicular membrane, reduces bile acid secretion and results in cholestasis [13].

The canalicular membrane has a structure of microvilli (thickness ca. 100-200 nm) [8], which significantly increases its surface providing space for a large number of transport proteins. As it is highly folded and not homogeneous, the protein intensities are captured from those layers which

978-1-4244-4122-8/11/$26.00 ©2011 IEEE

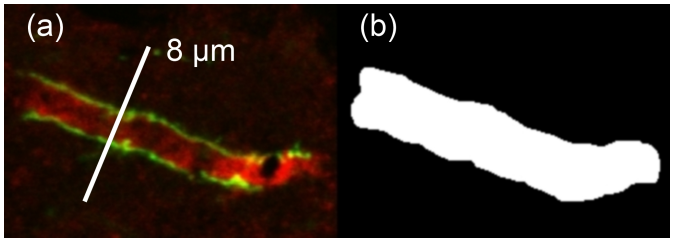

Fig. 1. (a) Representative canaliculus suitable for the analysis (green - Zo-1, red -Bsep). Profiles are extracted orthogonally to the canaliculus (white line). (b) Corresponding result of the foreground segmentation (white region).

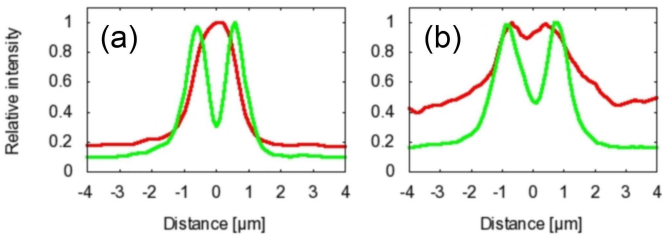

Fig. 2. Average intensity profiles of the (a) control and (b) test datasets. Green - Zo-1, red - Bsep.

are cut by an optical image plane of approximately 500 nm. Consequently, an obtained intensity distribution profile reveals a composition of two functions: a protein distribution across the individual microvillus and spatial folding of the microvilli. Fluorescent microscopy records the product of these functions without being able to decompose it. Therefore we do not aim at obtaining the function of the transporter distribution across the individual microvillus and work with the extracted *integral* intensity profile.

Several toponomic localization studies of Bsep were published [11], [12] comparing transport protein distribution by manual processing of microscopic images. Zonula occludens 1 (Zo-1), a tight junction associated protein, was selected as a structural marker, as it delimitates the canaliculus and allows its localization. An example of a canaliculus suitable for the analysis is shown in Fig. 1. Tight junctions (green lines) run along the two neighbouring hepatocytes and can be clearly distinguished. They have to be straight and run parallel to the image plane of a confocal microscope, while the Bsep transport proteins are mostly localized in the canalicular membrane between them. Nowadays, such images are processed and the protein translocation is evaluated manually.

II. METHODS

Rat liver tissue is a suitable model, as it is easier to obtain than human samples. Two datasets were prepared according to the known protocol: rat livers perfused under normo- and hyperosmolar conditions [13]. Images obtained by confocal laser scanning microscopy were studied for the toponomics of Bsep.

A. Manual Analysis

For manual analysis, measurements are taken where the Zo-1 immunostaining is undisrupted and run in parallel, the membrane segments are sufficiently long, straight and uniform [13]. As shown in Fig. 1, the profile is drawn orthogonally to a segment, as the protein distribution across the membrane is of interest. Canalicular diameter in the rat livers normally varies in the range of [0.8, 2.5 μm] (empirically determined), so that a profile of 8 μm covers the canaliculus and a significant part of the adjacent cells. The fluorescence intensity values from the structural and the functional markers are extracted along the profiles. Generally, about 100 profiles per sample are extracted on different positions to have a statistical basis. This evaluation method is time consuming, error-prone and subjective, as an expert decides where to measure profiles based on experience.

B. Intensity Profiles Extraction

Our workflow for the automated image processing and extraction of the intensity profiles was published earlier [6]. Briefly, canalicular membrane objects are identified in the confocal microscopy images using the foreground-background segmentation function of the Zeta software (Fig. 1) [14]. Following skeletonization of the foreground segments and pruning of the obtained skeleton yield centric lines, indicating directions to which canaliculi spread. Further, small and bent fragments unsuitable for the analysis are eliminated.

Intensity profiles are extracted orthogonally to the skeleton for both structural and functional markers at every selected position. Firstly, a tangent is fitted to identify the direction of this particular membrane segment. Then, an orthogonal line is drawn through this position spreading equally to both sides of the crossing point. Along this line the pixel intensities are quantified. Similar to [13], the width of the profile was increased by calculating the average of several lines. Later, intensity profiles with suboptimal Zo-1 distributions are eliminated (for details see [6]). The respective accepted Bsep profiles are passed to further analysis.

C. Improvement of the Intensity Profile Evaluation

Bsep profiles can be evaluated visually by comparing average profiles for each dataset. Fig. 2 shows average intensity profiles for the control dataset (livers perfused in a normoosmolar buffer) and a test dataset (perfusion in a hyperosmolar buffer). Compared to the control, average intensity profiles from the test dataset show translocation of Bsep from the canalicular membrane into the cytoplasm by increased intracellular fluorescence intensity (increased intensities on the tails of the profiles).

Nevertheless, for a statistical analysis, numerical descriptors have to be established to quantify the translocation. We have evaluated a number of descriptors to identify those achieving the highest discrimination rate between different conditions. Previously, the statistical variance of Bsep profiles was used to detect the internalization [11]–[13], and no subregions were distinguished in the profiles. However, different zones are clearly delimited by the Zo-1 maxima (tight junctions). The part of the profile between these

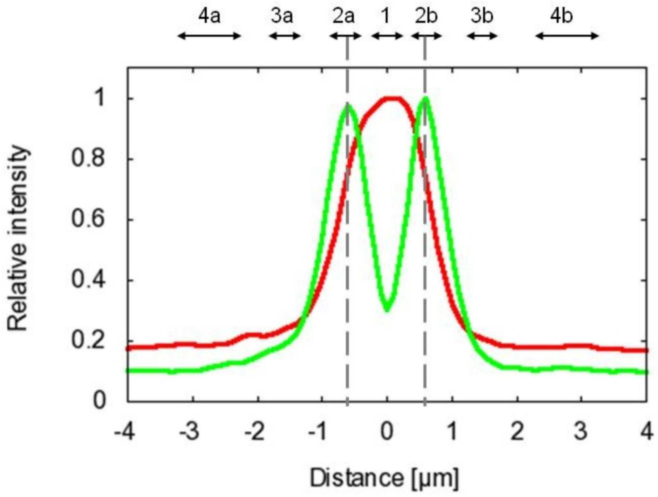

Fig. 3. Suggested zones for intensity profiles

maxima is considered to be the interior of the canaliculus, while the parts outside the tight junctions are considered to be cytoplasm. Restricting the analysis to particular zones might improve the statistical significance of a descriptor.

Zones with fixed length are calculated for the accepted Zo-1 profiles and applied to the respective Bsep profiles (Fig. 3). Firstly, two peaks are identified, and the center of the profile is found. Zone 1 has a length of 0.5 μm, and is centered between the peaks. Its length was chosen according to the empirical determination of canalicular width [0.8, 2.5 μm]. The integral intensity of zone 1 reveals the amount of Bsep in the center of the canaliculus. Zones 2a and 2b (0.5 μm each) are centered on the peaks. Their intensities indicate the protein concentration in the canalicular membrane. To analyze the internalization of transport proteins, intracellular zones were defined. Zones 3a and 3b (0.5 μm each) describe intracellular fluorescence intensities close to the canalicular membrane. Zones 4a and 4b are situated closer to the centers of the hepatocytes and measure 1.0 μm each.

In the considered case, we assume a translocation along the profile axis (from center to periphery). We are therefore looking for a numerical descriptor that varies with the distribution and maximally discriminates between the two physiological states and respective transporter distribution patterns. Integrals of zonal intensities can be compared as absolute numbers within one image because the zone length was kept constant in all profiles. However, due to limitations in the standardization of experimental conditions, immunohistochemical experiments do not allow absolute quantity measurements. Therefore, relative descriptors are preferred, such as ratios of intensity integrals. In the following, a number of such scale-free descriptors will be evaluated (Table I). One of them is presented as an example:

$$D = \frac{sum(Zone2a + Zone2b)}{sum(Zone4a + Zone4b)} \quad (1)$$

Descriptor D is calculated as a ratio of Bsep fluorescence intensities at the peaks of the Zo-1 profile relative to the intensities of Bsep in the cytosol far from the canalicular membranes. Bsep internalization affects *sum(Zone2a + Zone2b)* by broadening of the Bsep intensity profile and leads to an elevated fluorescence intensity due to Bsep vesicles in the cytoplasm which increases *sum(Zone4a + Zone4b)*. As the lengths of zones are kept constant for all profiles, the values of the descriptor can be easily compared and are expected to decrease under cholestatic conditions.

TABLE I
SUGGESTED DESCRIPTOR FORMULAS FOR THE EVALUATION OF THE PROTEIN TRANSLOCATION

Descriptor	Formula
X	sum1 / sum2
Y	sum1 / sum3
Z	sum1 / sum4
A	sum1 / (sum2 + sum3)
B	sum1 / (sum3 + sum4)
C	sum2 / sum3
D	sum2 / sum4
E	sum2 / (sum3 + sum4)

D. Statistical tests

The here suggested toponomics analysis aims at the detection of the membrane protein translocation. Therefore, the null hypothesis states that the Bsep distributions in both datasets are equal. Applied statistical tests are expected to reveal low p-values, indicating significant protein translocation in the test dataset compared to the control dataset.

The datasets are compared based on variables' values. We have performed both evaluations on the statistical variances and on the newly developed descriptors. An unpaired Wilcoxon rank sum test was chosen, as it does not make any assumptions on variables' distributions [13], can compare datasets of different sizes, and is widely used in medicine.

The number of the automatically extracted intensity profiles is much greater than the number of profiles obtained manually [6]. As the Wilcoxon rank sum test is sensitive to a number of data points, it has to be performed on equal sample sizes (e.g. 100 vs. 100) for objective comparison of the methods. Therefore, a sampling of 100 profiles from the automatically extracted ones was repeated 100 times, and the median p-values were reported. Two datasets are considered to be significantly distinct if the p-value < 0.05.

III. RESULTS

Two representative datasets of rat liver specimens were prepared by the University Clinic Düsseldorf and will be referred to as K (control samples, perfused in a normoosmolar buffer) and T (livers, perfused in a hyperosmolar buffer), respectively. K comprises 10 images of different regions of a sample, while T contains 11 images. Automated image processing and profile extraction led to obtaining ca. 4000 profiles per image. After the selection process, ca. 150 profiles per image were left. A manual profile extraction (100 profiles per dataset) was performed. Descriptors were

TABLE II

COMPARISON OF THE CONTROL AND TEST DATASETS. P-VALUES OBTAINED IN THE WILCOXON RANK SUM TESTS.

Descriptor	P-value	
	Manual	Automated
X	1.3 e-3	5.5 e-3
Y	1.6 e-21	4.7 e-14
Z	3.3 e-24	2.0 e-18
A	2.0 e-10	1.4 e-6
B	5.7 e-24	6.7 e-17
C	3.7 e-26	6.1 e-20
D	4.3 e-27	4.4 e-23
E	1.1 e-27	1.1 e-22
Variance	3.4 e-17	4.3 e-13

calculated and used for the comparison of the datasets by the Wilcoxon rank sum test (Table II). The expected result of the automated translocation analysis is to detect an increased amount of Bsep in intracellular vesicles.

P-values are very low and reveal significant differences in the descriptor variables' distributions in the dataset T relative to K. As the suggested descriptors are derived from the Bsep intensity profiles, we conclude, that the lateral translocation of the transporter took place. An automated profile extraction and the introduction of the new descriptors led to the improvement of p-values from e-17 (manual method, statistical variance) to e-23 (automated method, descriptor D).

IV. DISCUSSION

Manual evaluation of any data is subjective and error-prone simply due to a human factor. Consequently, an automatization of the data pre-processing, information extraction and evaluation is advantageous, as it can be objective and based on reproducible criteria. The here suggested automated evaluation of the lateral protein translocation extracts much more data points (4000 vs. 10), is faster (minutes vs. hours) and reproducible in comparison to the manual method.

Statistical tests have shown that p-values for the Bsep translocation detected by the automated method are only slightly worse than those obtained by a human expert, while the effect remains significant (p-value < 0.05). The results of both methods correlated well, which was also shown on 6 further experiments with varying biological conditions. Additionally, most of the newly suggested descriptors performed better than earlier used statistical variance, both in the automated and manual evaluations (e.g. D vs. statistical variance: e-27 vs. e-17, e-23 vs. e-13).

As it was mentioned before, the introduction of the zone model for intensity profiles allowed incorporation of the knowledge on the canalicular structure, such as tight junction positions. Oppositely, the statistical variance is calculated for the whole profile without distinguishing any subregions. Therefore, lower p-values achieved in Wilcoxon rank sum tests by descriptors rather than by the statistical variance met our expectations.

The most significant results were achieved by descriptors D and E, which are calculated from the intensities in the zones 2, 4 and 2, 3 and 4, respectively. Descriptors which incorporated integral intensities from zone 1 (X,Y,A,B) performed several orders worse. Therefore zone 1 is considered to be unreliable for the analysis.

V. CONCLUSIONS AND FUTURE WORKS

The suggested automated evaluation for the lateral protein translocation has proven to be comparable to the results of the manual processing. New descriptors for the intensity profiles led to more significant results than the statistical variance used earlier. Consequently, slow and subjective manual data processing can be substituted by the faster and completely reproducible method based on a greater number of extracted data points. Time saved for human experts is extremely valuable as it can be used for the operations which cannot (yet) be automated. Furthermore, the suggested translocation quantification can have a clinical application in diagnosis, therapy and drug discovery.

In the future, we plan to apply this method on different biological models and extend it to handle samples of human organism. Moreover, we aim at discovering the mathematical model for the lateral protein translocation in individual microvilli by decomposing the integral intensity function used here.

REFERENCES

[1] W. Schubert, "Topological proteomics, toponomics, melk technology," *Adv Biochem Eng Biotech*, vol. 83, pp. 189–209, 2003.

[2] R. F. Murphy, "Putting proteins on the map," *Nat Biotechnol*, vol. 24, pp. 1223–1224, 2006.

[3] M. Zhou and T. D. Veenstra, "Multiplexed immunofluorescence for the interrogation of cellular protein complexes," *Expert Rev Proteomics*, vol. 3, pp. 581–583, 2006.

[4] S. A. Haney, *High Content Screening: Science, Techniques and Applications*. Hoboken, New Jersey: John Wiley & Sons, 2008.

[5] J. Helmuth, G. Paul, and I. Sbalzarini, "Beyond co-localization: inferring spatial interactions between sub-cellular structures from microscopy images," *BMC Bioinformatics*, vol. 11, no. 1, p. 372, 2010.

[6] O. Domanova, S. Borbe, S. Mühlfeld, R. Kubitz, D. Häussinger, and T. Berlage, "Protein translocation analysis by automated image processing," in *Proc. Eighth IASTED International Conference on Biomedical Engineering (Biomed 2011)*, Innsbruck, Austria, Feb. 2011.

[7] J. A. Conchello and J. W. Lichtman, "Optical sectioning microscopy," *Nat Methods*, vol. 2, pp. 920–931, 2005.

[8] D. Häussinger, R. Kubitz, R. Reinehr, J. Bode, and F. Schliess, "Molecular aspects of medicine: from experimental to clinical hepatology." *Mol Aspects Med*, vol. 25, no. 3, pp. 221–360, 2004.

[9] A. Esteller, "Physiology of bile secretion." *World J Gastroenterol*, vol. 14, no. 37, pp. 5641–5649, 2008.

[10] D. Keppler and J. König, "Hepatic secretion of conjugated drugs and endogenous substances." *Semin Liver Dis*, vol. 20, pp. 265–272, 2000.

[11] R. Kubitz, M. Wettstein, U. Warskulat, and D. Häussinger, "Regulation of the multidrug resistance protein 2 in the rat liver by lipopolysaccharide and dexamethasone." *Gastroenterology*, vol. 116, pp. 401–410, 1999.

[12] M. Schmitt, R. Kubitz, M. Wettstein, S. vom Dahl, and D. Häussinger, "Retrieval of the mrp2 gene encoded conjugate export pump from the canalicular membrane contributes to cholestasis induced by tert-butyl hydroperoxide and chloro-dinitrobenzene." *Biol Chem*, vol. 381, pp. 487–495, 2000.

[13] M. Schmitt, R. Kubitz, S. Lizun, M. Wettstein, and D. Häussinger, "Regulation of the dynamic localization of the rat bsep gene-encoded bile salt export pump by anisoosmolarity." *Hepatology*, vol. 33, pp. 509–518, March 2001.

[14] Zeta, trainable image analysis for high content analysis. [Online]. Available: http://www.fit.fraunhofer.de/projects/life-science/zeta.html