Short thesis for the degree of Doctor of Philosophy (PhD)

Intact protein analysis using capillary zone electrophoresis-mass spectrometry in different capillaries

Narmin Hamidli

Supervisor: Dr. Attila Gáspár

UNIVERSITY OF DEBRECEN

Doctoral School of Chemistry

Debrecen, 2023

I. Introduction and objectives

Proteins possess inherent complexity. This is linked to their high molecular weight, post-translational modifications (PTMs), mutations, isoforms and microheterogeneities. Subtle modifications in the protein structure may induce the formation of functionally varying proteoforms, changing the efficacy and safety of proteins and protein complexes. The latter, in turn, may adversely impact medical and biological activity of molecules. Biopharmaceutical industry has long been willing to address this challenge, which necessitates the effective proteoform identification and characterization, that are usually deduced from the determination of the structure and the precise intact mass of the protein. For decades, mass spectrometry has been primarily responsible for this task.

The complete sequence of intact proteins and protein complexes may be well characterized using intact and native top-down mass spectrometric (TDMS) approaches, accordingly. Upon the intact analysis of therapeutic proteins TDMS often incorporates the chromatographic or electrophoretic separation step prior to MS analysis in order to simplify the distinct proteoform spectra. For the study of intact proteins, capillary electrophoresis has evolved into a mature and appealing technique. Capillary zone electrophoresis (CZE), capillary gel electrophoresis (CGE) and capillary isoelectric focusing (CIEF) have gained prominence in this field, while only CZE enables a straightforward and unchallenging connection, mass determination and protein characterization with an MS detector.

Although the coated capillaries have gained popularity in the last few decades for use in protein analysis, numerous established methods continue to rely on bare fused silica (BFS) capillaries. A number of studies in the scientific literature compare the performance of BFS capillaries to that of their coated counterparts. However, these comparisons are often presented in an unfair manner. The overlooked fact in these research is that the methods designed for protein analysis in coated capillaries may not necessarily be applicable to BFS

capillaries, which need further adjustment of conditions (particularly pH) for optimum performance and a fair comparison.

Based on the survey of the literature about intact protein analysis using CZE-MS the goals of my PhD research are (1) study of the separation parameters for BFS, polybrene (PB) and linear polyacrylamide (LPA) coated capillaries, enabling proper MS coupling; (2) the comparison of the analytical performance obtained with different capillaries; (3) investigation of CE-UV and CE-MS sensitivities attainable for simple and complex protein mixtures; (4) application of the proposed CZE-MS methods for pharmaceutical and biological samples.

II. Methods

A CE 7100 instrument (Agilent, Waldbronn, Germany) was used to perform CE separations. Unless otherwise specified, fused silica capillaries (50 μm I.D. and 370 μm O.D., Polymicro, Phoenix, AZ, USA) with the length of 65 cm (effective length (L_{eff}): 57 cm for UV) were chosen for the analyses with UV as well as MS detection. A 200 nm detection wavelength was selected for UV on-capillary detection. A sample introduction has been always performed hydrodynamically (50 mbar, 2 s) at the cathodic end of the PB, SMIL coated capillaries and at the anodic end of the BFS and LPA-coated capillaries. The separation voltage was found to be optimal as +20 kV for BFS, +30 kV for LPA and -30 kV for PB and SMIL capillaries. A nitrogen cylinder was utilized for the high pressure (max. 8 bar) rinsing. The operation of the CE instrument and the interpretation of the collected data have been accomplished by OpenLAB CDS Chemstation software (version B.04.02, Agilent).

The MS analysis was carried out via the MaXis II Ultra High-Resolution ESI-QTOF MS (Bruker, Karlsruhe, Germany) instrument. By the use of the co-axial ESI sprayer interface (G1607B, Agilent), CE and MS coupling has been made possible. The sheath liquid, which contained isopropyl alcohol and water (1:1) with 0.1% v/v FA (at 0.7 μL/min flow rate), was

delivered using a 1260 Infinity II isocratic pump (Agilent). External mass calibration was conducted on MS analyses using ESI-MS Tuning mix (part number: G2431A, Agilent) and on MS/MS analyses using Na-formate calibrant solutions. Background correction has been applied to electropherograms. The MS electrospray ion source settings for the five protein mixture analyses were as follows: capillary voltage: 4.5 kV; end plate offset: 500 V; nebulization pressure: 0.5 bar; dry gas temperature: 220 °C; dry gas flow rate: 8.0 L/min; MS scan range: 500-2400 m/z; spectra rate: 1.5 Hz. The following parameters have been applied throughout the analysis of human insulin and its analogues: capillary voltage: 4.5 kV; end plate offset: 500 V; nebulization pressure: 0.3 bar (following 500 s no pressure application after sample introduction); dry gas temperature: 200 °C; dry gas flow rate: 4.0 L/min; MS scan range: 600-2500 m/z; spectra rate: 3 Hz. The mass resolution for the insulin formulations were in the range of 65000-94000. MS/MS settings were tuned in a following way: mass scan range: 20-1800 m/z; spectra rate: 1 Hz; collision energy (CID): 45 eV. As a precursor ion, the most abundant charge state (5+) was selected. The measurements were monitored through otofControl software version 4.1 (build: 3.5, Bruker) and evaluated using Compass DataAnalysis software version 4.4 (build: 200.55.2969).

III. New scientific results

1. **We demonstrated that when the separation of proteins is performed in BFS capillary at low pH (pH = 1.8), minimal adsorption of components, good precision and separation efficiency similar to or even better than those obtained with the coated capillaries can be achieved.**

In case of the BFS capillary, it has been demonstrated that the pH of the BGE and capillary conditioning have a significant influence on protein adsorption and consequently, the reproducibility and sensitivity of the method. Strongly acidic pH (pH = 1.8) and thorough capillary preconditioning enabled adequate precision with RSD% values of 0.56-0.78% and 1.7-6.5% for

migration times and peak areas, respectively (n=10) (Fig 1). Separation efficiency data with minimum adsorption on the capillary surface comprised 27000/m-322000/m which were similar and sometimes higher than the values acquired in the PB-coated (5400/m-44000/m) and LPA-coated (7000/m-118000/m) capillaries.

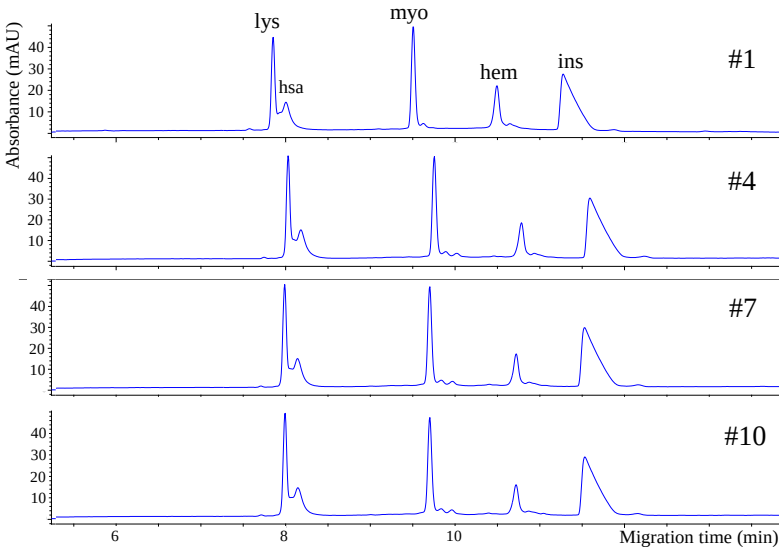

Figure 1. BFS capillary separation of 5 protein mixture. The repetitions are reflected by the 1st, 4th, 7th and 10th electropherograms. Sample: lysozyme (lys), human serum albumin (hsa), human insulin (ins), hemoglobin (hem) and myoglobin (myo).

2. The PB and the LPA-coated capillaries demonstrated their excellent separation power using BGE with pH higher than 2.5 for the CZE separation of simple and complex protein mixtures.

2.1 Using PB-coated capillary, for the first time, different proteoforms of human hemoglobin besides its globin chains have been separated and identified.

PB-coated capillary has been shown to be effective at resolving several components, namely, the various proteoforms of hemoglobin. In particular, the separation of hemoglobin globin chains (αa and βa with M of 15125 Da and 15867 Da, accordingly) as well as a few unrecognized proteoforms have been achieved in this capillary (Fig. 2).

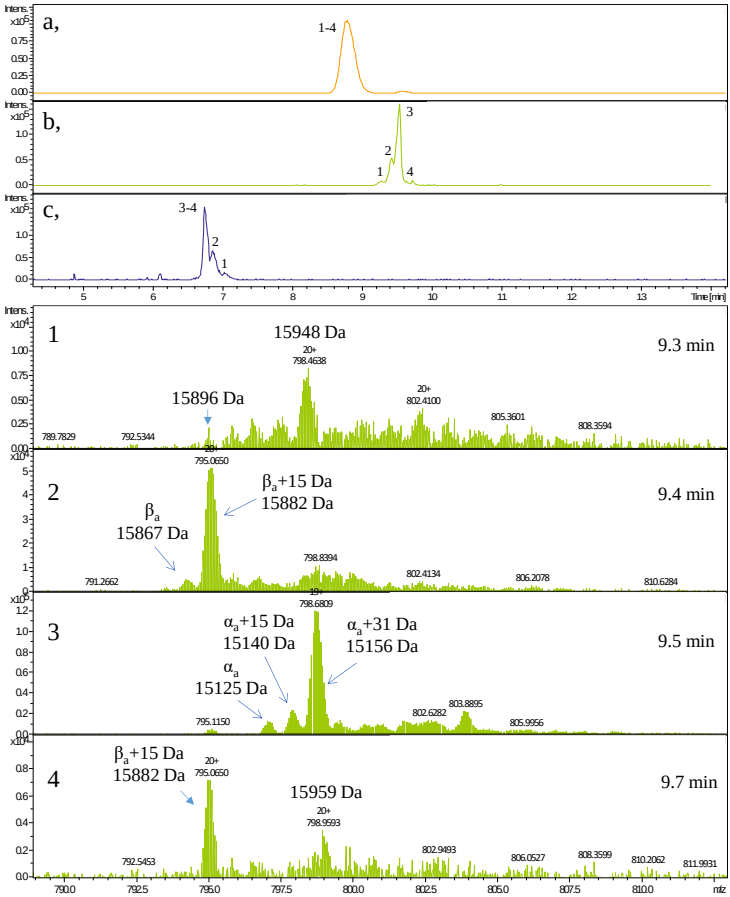

Figure 2. Hemoglobin analysis in (a) BFS, (b) PB-coated and (c) LPA-coated capillaries using CZE-MS. 1-4 indicate the spectra from (b). Sample: 5 mg/mL hemoglobin.

2.2 During the study of complex, protein-rich cobra venom sample PB or LPA coated capillaries have resolved a higher number of proteins and peptides compared to BFS capillaries.

Each capillary applied in optimized separation conditions proved to be relevant for the analysis of the protein-rich sample. As a result, BFS capillary allowed the identification of 29 venom proteins, while LPA-coated capillary has aided the separation of more components and identification of 31 proteins (Fig.3).

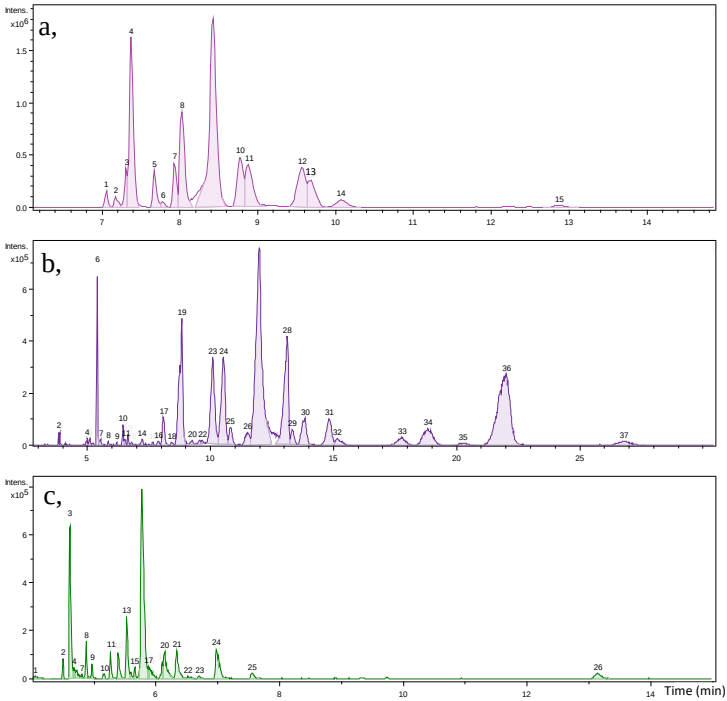

Figure 3. CE-MS analysis of cobra venom in uncoated and coated capillaries. (a) BFS capillary, (b) PB-coated capillary, (c) LPA-coated capillary.

PB-coated capillary, however, could achieve to separate 52 components, 38 of them being proteins. In PB-coated capillary the increased

throughput may be attributed to an extended analysis time and/or its superior resolving ability for smaller species.

3. **CZE has been proved to be suitable for the separation of proteins, while on-line UV photometric detection, particularly for large proteins, has often exhibited higher sensitivity than ESI-MS.**

As the detection technique changes from UV to MS, the following system parameters was found to be affected: (1) detection sensitivity degrade and (2) off-capillary MS detection slightly reduced the separation efficiency. Large proteins form a higher number of charge states and several adducts in ESI-MS spectra which reduces the signal intensity for a given analyte peak in the electropherogram. HSA, for instance, produced more than 40 charge states (too ambiguous to identify) and did not appear as a distinct peak (Fig. 4).

Regarding the venom sample, ESI-MS possesses a fairly good detection sensitivity for components with molecular masses lower than 25 kDa (Fig. 3), while it appeared to be less sensitive for larger proteins with a broad distribution of charged forms (and with various adducts). As a result, ESI-MS was unable to confirm the presence of the numerous large proteins present in venom (with masses varying between 30 and 100 kDa). Even though those proteins can be determined by UV detection, the identification is not yet achievable.

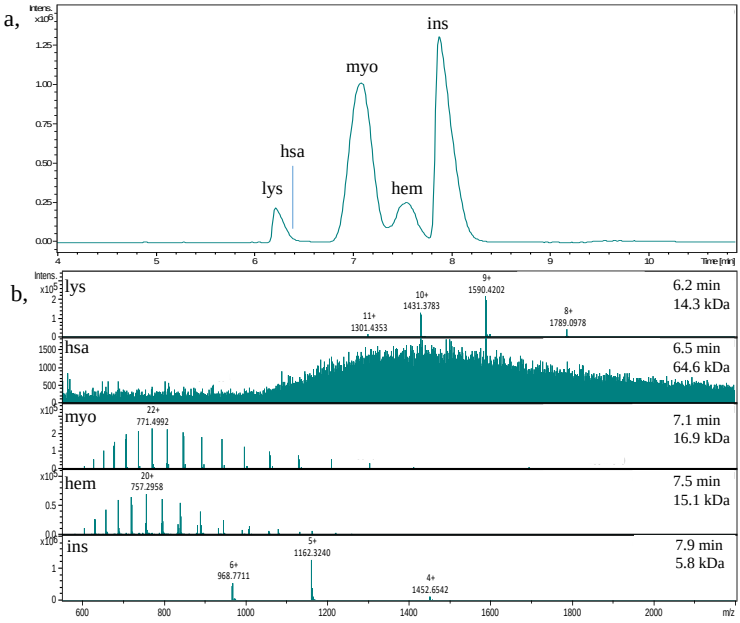

Figure 4. CZE-MS analysis of a protein mixture. (a) A base peak electropherogram and (b) mass spectra of separated proteins.

4. For the first time, we have developed a CZE-MS method to separate human insulin and its 6 analogues having only minimal differences between them.

As these insulins are small proteins with very similar masses, it is beneficial to determine whether established CZE method enables the differentiation and quantification of all components. The separation of seven recombinant insulins have been performed using PB-DS (dextran sulfate)-PB successive multi-ionic layer (SMIL)-coated capillary (Fig. 5). Based on 10 consecutive repetitions, the migration time reproducibility was in the range of 0.08-0.5 RSD%, while the peak areas were higher (5.8-8.8 RSD%) due to the wider peak shapes and overlapping of insulins. Compared to separation performances achievable with

the different capillaries SMIL-coated capillaries produced the greatest separation efficiency values (17000/m-192000/m).

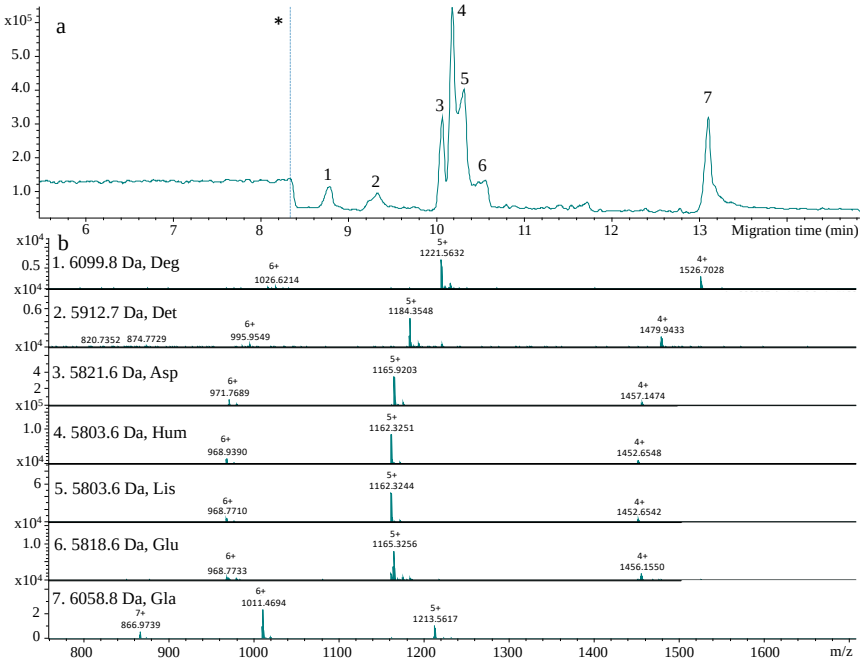

Figure 5. The CZE–MS separation of human insulin and its 6 analogues in SMIL-coated capillary. (a) base peak electropherogram; (b) mass spectra of separated insulins; ∗ the point that ESI nebulization pressure was turned on.

5. It was proved that the proteins with very similar structural conformations such as human insulin and insulin aspart can be characterized by MS/MS analysis using CID technique.

While examining the proteins, structural data may be required in addition to the separation that could be obtained through the fragmentation of the molecular ions. When the proteins feature a few disulfide bridges between cysteine residues, they could be efficiently disrupted using some fragmentation strategies (such as ECD), however, they commonly appear to be resistant to the

CID technique unless a protein reduction is conducted prior to analysis. As insulins keep three S-S linkages on the peptide backbone, the aforementioned challenge holds true for them as well. Yet, it was possible to distinguish insulins since the varying amino acid residues are positioned outside the disulfide loop. Due to the presence of diagnostic fragments, the distinction between these insulins was possible without uncertainty.

For the intact MS/MS analysis, aspart insulin was compared to human insulin, which presents a variation in a single amino acid at position β28 (Pro is substituted with Asp). Due to this small alteration in the structure, a 17.9742 Da mass difference is obtained. As a result, the MS/MS spectra of the product ions with the replaced amino acid residue revealed the existence of β28-containing fragments (Fig. 6).

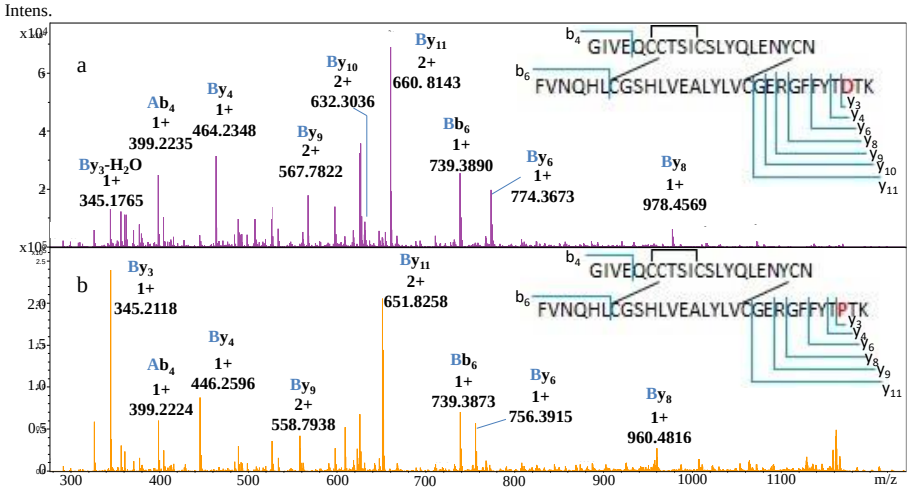

Figure 6. The spectra of (a) aspart and (b) human insulin from MS/MS analysis. The legends for the annotated peaks contain the chain (in blue) and the fragment type. Structure of the analogues are indicated, highlighting the difference in amino acid sequences in red as well as identified fragment types.

IV. Possible utilization of the results

Top-down proteomics is a rapidly evolving scientific field necessitating new insights about the CZE separation of intact protein mixtures and subsequent MS analyses. For the study of intact proteins, we compare the performances of three different capillaries in their optimal operating settings. This data can provide clear guidance to scientific and industrial community for optimizing experimental parameters and selecting the most suitable capillary (coated or uncoated) depending on the analytical goal. Additionally, the discussion about the detection sensitivities of UV photometry and ESI-MS for the analysis of small and large components has been provided, which may be beneficial information during the choice of detection technique.

The developed CZE methods have been qualitatively and quantitatively validated using seven recombinant insulins, demonstrating the applicability of the method for identifying each insulin in pharmaceutical formulations. The method is potentially relevant for the analysis of pre-mixed insulins or counterfeit insulin mixtures, where only small differences from the seven examined analogues may be anticipated.

The applicability of CZE-MS to other TDMS examinations of proteins have been presented. These areas cover a wider range from the analysis of deamidated forms (with 1 Da mass difference only) to native TDMS of proteins and very large components like mAbs (about 150 kDa molecular mass).

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/148/2023.PL
Subject: PhD Publication List

Candidate: Narmin Hamidli
Doctoral School: Doctoral School of Chemistry
MTMT ID: 10088368

List of publications related to the dissertation

Foreign language scientific articles in international journals (2)

1. **Hamidli, N.**, Pajaziti, B., Andrási, M., Nagy, C., Gáspár, A.: Determination of human insulin and its six therapeutic analogues by capillary electrophoresis - mass spectrometry.
   J. Chromatogr. A. 1678 (16), 1-8, 2022. ISSN: 0021-9673.
   DOI: https://doi.org/10.1016/j.chroma.2022.463351
   IF: 4.601 (2021)
2. **Hamidli, N.**, Andrási, M., Nagy, C., Gáspár, A.: Analysis of intact proteins with capillary zone electrophoresis coupled to mass spectromery using uncoated and coated capillaries.
   J. Chromatogr. A. 1654, 1-9, 2021. ISSN: 0021-9673.
   DOI: http://dx.doi.org/10.1016/j.chroma.2021.462448
   IF: 4.601

DEBRECENI EGYETEM Egyetemi és Nemzeti Könyvtár

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

List of other publications

Foreign language scientific articles in international journals (2)

3. Nagy, C., Andrási, M., **Hamidli, N.**, Gyémánt, G., Gáspár, A.: Top-down proteomic analysis of monoclonal antibodies by capillary zone electrophoresis-mass spectrometry. *Journal of Chromatography Open. 2*, 1-21, 2022. EISSN: 2772-3917. DOI: https://doi.org/10.1016/j.jcoa.2021.100024

4. Andrási, M., Pajaziti, B., Sipos, B., Nagy, C., **Hamidli, N.**, Gáspár, A.: Determination of deamidated isoforms of human insulin using capillary electrophoresis. *J. Chromatogr. A. 1626*, 1-8, 2020. ISSN: 0021-9673. DOI: http://dx.doi.org/10.1016/j.chroma.2020.461344 IF: 4.759

Total IF of journals (all publications): 13,961
Total IF of journals (publications related to the dissertation): 9,202

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

11 May, 2023

DEBRECENI EGYETEM * Egyetemi és Nemzeti Könyvtár *