DOI 10.1002/pmic.200800598 *Proteomics* 2009, *9,* 1326–1343

RESEARCH ARTICLE

Biospecific irreversible fishing coupled with atomic force microscopy for detection of extremely low-abundant proteins

Alexander Archakov, Yurii Ivanov, Andrey Lisitsa and Victor Zgoda

Institute of Biomedical Chemistry, Moscow, Russia

Received: July 15, 2008
Revised: September 15, 2008
Accepted: October 3, 2008

In the absence of an analog of PCR for proteins, the concentration detection limit (DL) becomes a real challenge. The problem may be solved by means of a combination of biospecific irreversible fishing with atomic force microscopy (AFM). AFM offers the ability to register individual molecules and their complexes, while biospecific fishing takes advantage of an affine interaction between analyte molecules spread over a large volume of biomaterial and ligand molecules immobilized on the chip surface. Fishing may be conducted in Kd-dependent reversible mode and in Kd-independent irreversible mode. In this study, the DLs of two previously applied proteomic approaches were determined and compared to the DL of a newly developed analytical method. The first approach, based on MS analysis of biomaterial after 2-DE or LC separation of proteins, attained a DL at the level of 10^{-8}–10^{-10} M. The second approach, based on the optical biosensor analysis of molecular interactions in the format of proteomic microarrays, had a DL of 10^{-9}–10^{-10} M. Our proposed method which combines biospecific fishing with AFM allowed us to attain DL values of 10^{-11} M under reversible binding conditions and 10^{-16} M under irreversible binding conditions.

Keywords:
Atomic-force microscopy / Biospecific fishing / Detection limit / Molecular detectors / Optical biosensor

1 Introduction

Undoubtedly, much technological progress has been made in life science since the year 1990, when the Human Genome Project began. However, the progress of proteomic technologies has been constrained by the absence of a reaction analogous to PCR [1] whose application has given a major impetus to the rapid development of genomics [2]. PCR has enabled researchers to surmount the limitations on sensitivity inherent in procedures for identification of genetic material [3]. It allows multiple copies of nucleic acids to be made and thus removes the detection limit (DL) constraints for DNA-encoded [4] and RNA-encoded [5] genetic material. By applying PCR, it is possible to identify even single copy of nucleic sequence occurring in the sample [6]. Current proteomic technologies are far from such successful solutions [2, 7]. Their present state is well expressed by the phrase: "we began to run before we could walk" [8]. The most often used proteomic approach involves fractionation and separation of biomaterial with subsequent identification of protein molecules by MS. The separation of protein molecules is commonly performed by using electrophoresis or high-pressure LC [7].

Protein molecules occur in biological material at various concentrations, *i.e.*, there exists a dynamic concentration

Correspondence: Professor Alexander Archakov, 119121 Pogodinskaya Str. 10, Moscow, Russia
E-mail: alexander.archakov@ibmc.msk.ru
Fax: +7-4952450857

Abbreviations: AB, amido black; **AFM,** atomic force microscopy; **anti-HBsAg,** antibody to HBsAg; **anti-HCVcore,** antibody to HCVcoreAg; **cTrC,** cardiotroponin C; **Cy5-sat,** Cy5 saturated labeling; **DL,** detection limit; **DTBPA,** 4,4′-dithiobisphenylazide; **EDC,** 1-ethyl-3-(3-dimethyl-1-aminopropyl)carbodiimide; **GA,** glutaric aldehyde; **HBsAg,** hepatitis B antigen; **HCVcoreAg,** hepatitis C core antigen; **im,** immobilized; **MW,** molecular weight; **OB,** optical biosensor; **PBS/t,** PBS/Tween-20 buffer; **ST,** silver thiosulfate

range. The problem lies in detecting low-abundant molecules, whose concentrations are below 10^{-10}–10^{-12} M, against a background of molecules occurring at much higher concentrations such as 10^{-3}–10^{-6} M [9]. Although the importance of dynamic range has generally been recognized [10–12], the problems associated with DL have not yet received much attention [13, 14].

Methods for protein separation have a rather low DL. For instance, in the case of MS-compatible 2-DE, the DL is equivalent to 10^{-9} M for postelectrophoretic staining and 10^{-11} M for pre-electrophoretic labeling with Cy5 under saturating conditions [15]. The MS protein identification stage is characterized by a higher concentration sensitivity. Usage of triple quadrupole makes it possible to go down to the femto- and attomolar sensitivity levels [16]. For the typical volume (0.5–1.0 μL) of sample injected into the mass spectrometer from the probe vial such sensitivity corresponds to the DL range of 10^{-11}–10^{-14} M (hereinafter the averaged molecular weight (MW) of 50 kDa is used to convert to the common units of moles *per* liter (M)). Combining electrophoresis with MS lowers the sensitivity of the entire technology to the picomolar level [7], *i.e.*, down to 10^{-10} M, because of the loss of protein molecules during the separation step. In addition, concentration limitations arise as a result of the limited volumes of biomaterial employed in the analytical procedure. Thus, for MS-oriented shotgun proteomics there exist objective limitations that are determined by a variety of factors, including detector sensitivity, separation losses, and sampling constraints [7, 17].

The second group of proteomic technologies is based on the application of protein microarrays, which are microchips with surface-immobilized ligands for specific proteins [18]. Antibodies or aptamers are used as ligands [2, 19], and a microchip could carry up to 400 sites of various immobilized ligands (Biacore Flexchip Product Information, 2007, GE Healthcare, USA). When biomaterial is loaded onto the microchip, the chip-immobilized ligands bind the appropriate protein molecules (analytes). The complexes thus formed are registered either by ELISA-based methods [18] or through label-free direct detection of the molecular interactions in real time [20–24].

With respect to proteomic microarrays, the DL is determined not only by the technical restrictions of the detecting devices but also by the existence of a thermodynamic barrier to the reversible binding of a pair of biomolecules. The number of observed protein complexes depends on two parameters: the dissociation constant and the concentration of analyte molecules in the biomaterial. Despite the intensive development of proteomic microarray technology, the DL has not attained 10^{-12} M in the standard ELISA format [18]; however, it may be lowered to the level of 10^{-13} M by the use of fluorescent labels and waveguide nanodetectors [23].

A rather complex situation has evolved in proteomic research: apparently, by using standard technologies it is impossible to detect protein molecules at concentrations below 10^{-12}–10^{-13} M [9, 25]. On the other hand, at lower concentrations a greater number of proteins are found in biomaterial. If the DL exceeds 10^{-4} M, only several dozens of various protein species may be detected in plasma but with a decrease in the DL to 10^{-12} M, the number of proteins grows to several thousands [26]. Extrapolation of this dependence suggests that, at a concentration level of 10^{-18} M, over 50 000 protein species are likely to be found in plasma [13]. Theoretical calculations based on combinatorial analysis showed that up to 1 million proteins may occur in a living organism [12], of which minimally 500 000 are most likely present in plasma [27].

Nowadays, new approaches are being developed to avoid, partially or wholly, the deadlock of DL in proteomics [21, 28]. One such technology employs a combination of atomic force microscopy (AFM), used to detect individual molecules and their complexes, with a procedure for determining the preliminary concentration of proteins on an AFM-chip surface [22]. The AFM-chip is fabricated by means of covalent immobilization on the atomically smooth surface of ligand molecules which are able to specifically bind target proteins from analyte solutions. The sampled biomaterial is deposited onto the fabricated AFM-chip. Unbound molecules are removed by washing; the complexes formed between ligand and target protein (analyte) are then counted in the scanning area of AFM.

In contrast to proteomic microarrays, which are restricted to limited sample volumes, the AFM-chip is able to capture sought-after proteins from large volumes of biomaterial. This property of AFM-chips is of basic importance for the detection of extra-low protein concentrations.

Previously we demonstrated the possibilities of using AFM as a molecular detector [29, 30]. During the biospecific fishing procedure the components of the biomaterial are concentrated from a large sample volume onto the AFM-chip surface, whereupon individual molecules and their complexes are counted on the chip using AFM. This approach is based on readily available technology; its efficient application is only limited by the absence of devices for high-throughput AFM scanning of the surfaces [22]. High-throughput scanning awaits technologies based on high-speed AFMs [31], including those operating in a noncontact mode [24], or "Milliped" multicantilever scanners [32].

Unfortunately, owing to the existence of the thermodynamic barrier, such molecular detectors do not eliminate the concentration constraints [13]. This problem may be solved by means of conversion of reversible analyte/ligand binding into irreversible binding. For this purpose, the ligands are modified by photochemically active labels. Upon UV-illumination of the chip the formation of covalent complexes takes place, enabling the analyte molecule to be irreversibly bound to the ligand. Such complexes would not dissociate, thereby surmounting the thermodynamic barrier. That opens the way to counting the number of individual protein molecules and their complexes at concentration levels that approximate the reverse Avogadro number [13].

This study was undertaken with the aim of comparing proteomic technologies that are based on different principles of identification of protein molecules. On the one hand, the potentialities of two traditional approaches were investigated. The first approach combines electrophoretic/chromatographic separation of proteins with their subsequent identification by use of mass-spectrometric analysis. The second approach employs proteomic microarrays where antigen–antibody interactions are registered using an optical biosensor (OB). In addition to these commonly adopted approaches, a new one was proposed that involves biospecific fishing followed by AFM-based counting of molecular complexes. We showed that the DL of MS-oriented proteomics was 10^{-7} M while application of the OB enabled us to reach values as low as 10^{-10} M. Even lower DL values of 10^{-11} and 10^{-16} M were achieved by using reversible and irreversible AFM fishing, respectively.

2 Materials and methods

2.1 Reagents

For covalent immobilization of antibodies, the following reagents provided by Sigma were used: 1-ethyl-3-(3-dimethyl-1-aminopropyl)carbodiimide (EDC) and *N*-hydroxysuccinimide (NHC). Photoactivated coupling was carried out using 4-[*p*-azidosalicylamido]butylamine (ASBA) from Pierce and 4,4′-dithiobisphenylazide (DTBPA) from Sigma.

Urea, Cy5, CHAPS, and DTT were obtained from GE Healthcare; BCA Protein Assay kit was obtained from Pierce; IPG strips (nonlinear, pH 3–10, 5 cm long), SDS, and acrylamide were from BioRad; ampholines were from Amersham Biosciences; sequencing grade modified trypsin was from Promega (Madison, USA). ACN (HPLC grade) and water (HPLC grade) were from Merck. Other reagents were of analytical grade. For AFM experiments, Milli-Q bidistilled deionized water was filtered through 1-nm centrifuge Millipore filters.

The following protein preparations were used: antibodies to HCVcore, clone 1 E5, further referred to as anti-HCVcore (Vitrogen, USA); recombinant antigens HBsAg (Aldevron, USA); cardiotroponin C (cTrC) and anti-cTrC (USBiological, USA); antibodies NE3 to antigen HBsAg (anti-HBsAg) were a gift from Dr. A. Khramtsov and Dr. A. Filatov (Institute of Immunology, Moscow, Russia); recombinant antigen HCVcore (HCVcoreAg) was kindly presented by Dr. O. Yastrebova (Vector-Best, Novosibirsk, Russia).

2.2 Analysis of plasma

2.2.1 Immunoaffine depletion

The removal of six major proteins from plasma, including serum albumin, immunoglobulins IgG/IgA, transferrin, antitrypsin, and haptoglobin, was carried out according to Agilent Multiple Affinity Removal System (MARS) protocol (Agilent Technologies).

2.2.2 2-DE

Samples (10 μL each) were prepared by dilution of depleted plasma in the concentration range of 10^{-6}–10^{-5} M. The protein content in the samples was determined by BSA assay. Samples were solubilized by the addition of 190 μL buffer containing 8 M urea, 2% CHAPS, and 0.1% ampholine Bio-Lyte, pH 3–10. Isofocusing strips were rehydrated overnight in the presence of 200 μL solubilized sample and IEF was performed using IEF-cell (BioRad) according to the manufacturer's protocol, gradually increasing the voltage up to 250 V for 20 min, then from 250 to 8000 V for 5 h, followed by 45 kV·h at 8000 V. After IEF the strips were equilibrated for 30 min in 4 mL buffer containing 6 M urea, 0.5 M Tris-HCl, pH 6.8, 0.2% SDS, and 2% DTT. Focused strips were transferred to the MiniProtean III electrophoresis (BioRad). Separation in the second dimension was conducted in 12% SDS-PAGE at a constant voltage of 200 V.

Pre-electrophoretic staining of the gels was performed in minimal and saturation labeling modes with Cy5-dye according to the manufacturer's protocols (GE Healthcare). Detection was carried out in the fluorescent mode using the Typhoon 9410 scanner (Amersham Biosciences). Postelectrophoretic staining of gels was performed using amido black (AB) [33], CBB [34], silver thiosulfate (ST) [35], and glutaric aldehyde (GA) [36], according to the published protocols. Obtained gels were scanned by GS-800 densitometer (BioRad). The number of protein spots on the images was detected automatically using Melanie III software (GeneBio, Geneva, Switzerland).

2.2.3 Identification of SDS-PAGE separated plasma proteins using LC-MS/MS

A 10 μL aliquot of depleted plasma was loaded *per* lane. Separation of proteins was carried out on 12% SDS-PAGE using MiniProtean III at a constant voltage of 200 V. Each lane was divided into 24 peer bands (~2 mm thick). The bands were washed thrice in deionized water. Proteins were subjected to in-gel trypsinolysis: after three washes in water, the gel pieces were incubated in 50% (v/v) ACN and 100 mM ammonium bicarbonate (pH 8.9) for 20 min, then in 100% ACN for 20 min. The pieces were dried for 1 h, then 10 μL of trypsin solution (25 ng/μL trypsin in 50 mM ammonium bicarbonate) were added to the gel and protein hydrolysis was carried out at 37°C overnight.

After trypsinolysis, formic acid was added to the digests up to 0.1%. Then the preparations were evaporated using Vacuum Concentrator 5301 (Eppendorf), dissolved in 10 μL of 0.1% formic acid and analyzed by LC-MS/MS. RP nano-LC-MS/MS was performed using an Agilent 1100 nanoflow HPLC-chip cube system coupled to an Agilent XCT Ultra Series MSD Trap (Agilent Technologies). Tryptic peptides

were separated on an HPLC-chip (40 nL trap column, 75 μm × 43 mm analytical column, 5 μm C-18SB-ZX, Agilent technologies) using a linear gradient of 5–80% ACN in 0.1% formic acid over 60 min at the flow rate of 300 nL/min and detected by IT in the 300–1800 m/z range according to the supplier's recommendations. Mass spectra were acquired in a positive-ion mode with automated data-dependent MS/MS of the five most intensive ions selected from the precursor MS scan.

Protein identification was performed using SpectrumMill MS Proteomics Workbench Rev A.03.03.078 (Agilent Technologies) by comparison of experimental data to the Swiss-Prot human subset database. The following search parameters were applied: trypsin as the cutting enzyme, mass tolerance for the monoisotopic peptide window was set to ±1.5 Da, the MS/MS tolerance window was set to ±0.5 Da and two missed cleavages were allowed. Cysteine acrylamide modification and oxidized methionine were chosen as variable modifications. The criteria for positive identification were set as follows: 70% minimum scored peak intensity, δ-forward-reverse score >2; at least two peptide identifications with a confidence score >7 and summarized protein score >14; at least three positive identifications from three different gels.

2.3 Experiments using the OB as a concentration detector

A two-channel IAsys+ "resonance mirror" OB (Affinity Sensors, UK) was used as a prototypic microarray detector. Complex formation between the ligand and the analyte was measured by an increase in the refractive index in the sensitive layer of OB in real time [22]. For this purpose, the molecules of ligand proteins were immobilized in the working channel of the carboxylated OB cuvette *via* their amino groups with the use of EDC/NHS chemistry in accordance with the Methods Guide to the IAsys+ (IAsys Methods Guide (Manual AI1), Affinity Sensors, 1996, SaxonWay, Bar Hill, Cambridge, England).

In the first set of experiments, anti-HBsAg (MW = 150 kDa) antibodies were immobilized in the working channel of the OB cuvette filled with analyte containing the corresponding HBsAg antigen (MW = 25.5 kDa). In the second set of experiments, the 18-kDa cTrC protein was immobilized, and 150-kDa anti-cTrC antibody was introduced as analyte. In the following text, the immobilized ligands are designated as anti-$HBsAg_{im}$ and $cTrC_{im}$. Immobilization of anti-HBsAg and cTrC was conducted in 10 mM Na-acetate buffer (pH = 5.0) for 11 and 20 min, respectively. After immobilization, the unreacted carboxyl groups were deactivated with 1 M ethanolamine at pH 8.5. Control channels of the cuvettes were modified analogously; however, the immobilization buffer did not contain ligand molecules.

Analyte samples were prepared by diluting the stock in 10 mM PBS/Tween-20 buffer (PBS/t; pH 7.4) to final concentrations ranging from 10^{-7} to 10^{-10} M for reversible binding and 10^{-7} to 10^{-14} M for irreversible binding.

2.3.1 Experimental procedure for reversible binding in OB cuvette

Measurements started after addition of the analyte in PBS/t, pH 7.4, to the measuring cuvette. Ligand-to-analyte binding was detected within 7–10 min, then the cuvette was rinsed with fresh buffer and dissociation was monitored for the next 3–5 min. The association rate constant k_{on} was calculated from the first-order exponential equation approximated to the binding part of the experimental curve:

$$R(t) = R(0) + \Delta R\left(1 - e^{-(k_{on}C + k_{off})t}\right)$$

where $R(t)$ is the biosensor response as a function of time, C the ligand concentration, $\Delta R = R(\infty) - R(0)$, and $R(0)$ and $R(\infty)$ designate the responses at the initial and equilibrium states, respectively.

The dissociation rate constant k_{off} was calculated from the approximation of the dissociation part of the curve by the equation: $R_t = \Delta R e^{k_{off} t}$. The dissociation constant (Kd) was calculated as the ratio of k_{off} to k_{on}.

2.3.2 Experimental procedure for irreversible binding in OB cuvette

The measuring cuvette with immobilized ligand was loaded with 50 μL of the appropriate analyte solution in PBS/t, pH 7.4. For the initial 7–8 min the reversible binding was observed, whereupon 1 μL of 2×10^{-5} M solution of DTBPA in DMSO was added. For the next 9–10 min UV irradiation under a mercury lamp (λ>300 nm) was conducted as previously described [37]. After irradiation, the cuvette was rinsed with fresh buffer to wash out the unbound complexes, and the descent of the dissociation curve was followed to equilibrium. The signal amplitude at the achieved level of dissociation was used to assess the surface concentration of the covalently bound complexes.

2.4 Experiments using the atomic force microscope (AFM) as a molecular detector

The antibodies to HCVcore were covalently immobilized onto aminosilanized mica support [38]. For this purpose, 0.5 μL solution of 50 mM K-phosphate buffer with 10 mM NaCl, pH 7.4, containing 0.1 μM antibodies and the EDC/NHS mixture (ratio 0.16:0.04 M), was deposited onto the mica surface where it was incubated for 8 min. After that the support was washed in bidistilled deionized water for 30 min. The density of immobilized antibodies was 15 000 molecules *per* 400 μm^2.

The samples of analyte (HCVcore antigen) in 50 mM PBS buffer, pH 7.4, were prepared in the concentration range of 10^{-8}–10^{-17} M.

2.4.1 Procedure for reversible AFM fishing

The AFM-chip was incubated for 60 min at 37°C in a 1-mL volume of the analyte. After incubation the chips were washed in water for 90 min, then dried and scanned by AFM. The AFM NTEGRA (NT-MDT, Zelenograd, Russia) was employed. The measurements were made in a tapping mode. The NSG10 cantilevers (NT-MDT) were used as probes. The typical radius of the probe tips was 10 nm. The resonance frequency was in the range of 190–325 kHz. The size of one scan was 5×5 μm. The entire scanning area comprised 16 scans (400 μm^2). Each measurement was repeated at least three times.

2.4.2 Procedure for irreversible AFM fishing

The crosslinking protocol for irreversible fishing was adapted based on previous report [39] and Pierce's instructions. The antibodies immobilized onto the AFM-chip surface were chemically modified by photocrosslinker. For this purpose, the prepared AFM-chip was incubated in the activation buffer for 10 min at 25°C in the dark. The activation buffer was prepared by addition of 1 μL of stock solution (2 mM ASBA in DMSO) to 100 μL of 50 mM PBS buffer, pH 7.4, containing 0.04 M NHS and 0.04 M EDC. The AFM-chip was washed once with 1 mL of 2% DMSO solution in water and then twice (each time for 30 min) with pure water. Photocrosslinking was carried out by incubation of the modified AFM-chip either with a 1 or 50-mL sample for 30 min at 30°C with simultaneous irradiation under a mercury lamp (λ>300 nm). Further preparation and scanning of the AFM-chip were performed as described before (Section 2.4.1) in three independent replicates.

2.4.3 Recognition of objects on the AFM image and height histograms

Objects were detected on the image using a combination of nonlinear filtering with watershed algorithm of segregation. Selected objects were additionally verified by analysis of their characteristic shape and height distributions. The maximal heights of objects were taken to produce the histogram of the objects' height distributions (height histogram). Details of the implementation and trial version of the software are available at http://projects.ibmh.msk.su/afmparser/.

2.5 Equations for calculation of DL

The formation of complexes between analyte A and ligand B (immobilized onto the chip) is described as the reaction A + B = AB. By introducing [A], [B], and [AB] as the concentrations of analyte, ligand, and their complex, respectively, the dissociation constant for the above reaction may be expressed by the equation:

$$\mathrm{Kd} = \frac{[\mathrm{A}][\mathrm{B}]}{[\mathrm{AB}]} = \frac{([\mathrm{A_0}] - [\mathrm{AB}])([\mathrm{B_0}] - [\mathrm{AB}])}{[\mathrm{AB}]} \quad (1)$$

where $[A_0]$ and $[B_0]$ are the initial concentrations of analyte and ligand, respectively.

Having solved Eq. (1) relative to [AB] and having designated $G = {}^1/_2\{[B_0] + [A_0] + Kd\}$, we obtain the expression for calculating the concentration of the complexes formed:

$$[\mathrm{AB}] = G \pm \sqrt{G^2 - [\mathrm{A_0}][\mathrm{B_0}]} \quad (2)$$

It should be noted that if the complex formation reaction is irreversible, *i.e.*, if Kd = 0, then Eq. (2) may be written as:

$$[\mathrm{AB}] = \left\{ \begin{array}{ll} [\mathrm{B_0}], & \text{if } [\mathrm{A_0}] > [\mathrm{B_0}] \\ [\mathrm{A_0}], & \text{if } [\mathrm{A_0}] < [\mathrm{B_0}] \end{array} \right\} \quad (3)$$

Proceeding from Eq. (1), the DL may be expressed as a minimal analyte concentration $[A_{min}]$ detectable by a registering device:

$$\mathrm{DL} = [\mathrm{A_{min}}] = \mathrm{Kd}\left(\frac{[\mathrm{B_0}]}{[\mathrm{AB}]} - 1\right)^{-1} + [\mathrm{AB}] \quad (4)$$

If a molecular detector is used as the registering device, we may convert from volume molar concentrations to the real number of chip-immobilized ligand molecules N_{BO} and the number of complexes between ligands and analytes N_{AB}. Then, taking into account that $G = {}^1/_2\{N_{BO}/(V \times N_A) + [A_0] + Kd\}$, Eq. (2) may be written as:

$$N_{\mathrm{AB}} = N_{\mathrm{A}} V[\mathrm{AB}] = N_{\mathrm{A}} V\left(G - \sqrt{G^2 - A_0 \frac{N_{\mathrm{B0}}}{V N_{\mathrm{A}}}}\right) \quad (5)$$

where N_A is the Avogadro number and V is the sample volume. Analogously, Eq. (4) may be rewritten as:

$$\mathrm{DL} = \mathrm{Kd}\left(\frac{\mathrm{N_{B0}}}{\mathrm{N_{AB}}} - 1\right)^{-1} + \frac{\mathrm{N_{AB}}}{V\mathrm{N_A}} \quad (6)$$

Equation (6) is applicable in cases where the binding of analyte to ligand is reversible. With irreversible binding Kd = 0 and Eq. (6) takes the form:

$$\mathrm{DL} = [A_{\mathrm{min}}] = \frac{N_{\mathrm{AB}}}{V N_{\mathrm{A}}} \quad (7)$$

2.5.1 Calculation of DL for OB as a concentration detector

In the case of chip-based detectors, Eq. (4) should be expressed using surface concentrations of molecules (C) on the chip area (S):

$$\mathrm{DL} = [\mathrm{A_{min}}] = \mathrm{Kd}\left(\frac{C_{\mathrm{B0}}}{C_{\mathrm{AB}}} - 1\right)^{-1} + \frac{C_{\mathrm{AB}} S}{V} \quad (8)$$

The surface concentration of immobilized ligands (C_{BO}) and the surface concentration of ligand/analyte complexes (C_{AB}) were calculated using calibration curves relating the MW of a protein immobilized on the specific area $S = 1$ mm^2 to the OB response (600 arcsec/μm^2), according to the technical specification for the FCC-0301 OB cuvette:

$$C = \frac{RN_A \times 10^{-9}}{600 \times MW} \quad (9)$$

where R is the increment in the signal amplitude. If Eq. (9) was used for the calculation of the number of immobilized ligands, *i.e.*, if $C = C_{BO}$, the MW of the ligand was substituted into the formula. In the case of calculating the number of complexes ($C = C_{AB}$), MW was the MW of the analyte. The minimal surface concentration of the complexes C_{AB} can theoretically be determined as 2×10^7 molecules/mm^2 by introducing into Eq. (9) the manufacturer-declared minimal sensitivity of OB ($R = 1$ arcsec) and the averaged MW = 50 kDa.

As a rule, the surface concentration of the immobilized ligand is $C_{BO} = 10^{10}$ molecules/mm^2 [40], while the volume of the FCC-0301 cuvette is $V = 100$ μL. By assuming that the highest affinity limit for a protein–protein interaction is determined by Kd = 10^{-12} M, we are able to calculate the DL for reversible binding from Eq. (8):

$$DL = Kd\left(\frac{C_{B0}}{(2 \times 10^7)} - 1\right)^{-1} + \frac{(2 \times 10^7)S}{V} = 1.5 \times 10^{-15} + 0.3 \times 10^{-11} = 3 \times 10^{-12}\ M \quad (10)$$

In case of irreversible binding, the corresponding theoretical value of DL will be:

$$DL = [A_{min}] = \frac{(2 \times 10^7)S}{V} = 3 \times 10^{-12}\ M \quad (11)$$

The equality of DL values obtained in Eqs. (10) and (11) demonstrates that the contribution of the Kd-dependent item to Eq. (4) is negligible, *i.e.*, DL is mainly dependent on analyte concentration.

2.5.2 Calculation of DL for an AFM-based molecular detector

By substituting the values into Eq. (6), we may estimate the theoretical DL for our AFM-chip containing $N = 15\,000$ immobilized antibodies. We assume that Kd = 10^{-12} M, and the result is reliable if no less than 1000 complexes are observed, *i.e.*, when $N_{AB} \geq 1000$. In this case, the DL value becomes 10^{-12} M:

$$DL = 10^{-12}\left(\frac{15000}{1000} - 1\right)^{-1} + \frac{1000}{V \times 6.02 \times 10^{23}} \approx \left[0.7 \times 10^{-13} + \frac{1.6 \times 10^{-18}}{V}\right] \approx 10^{-12} M \quad (12)$$

Upon reversible binding, the value of the second term in Eq. (6), in whose denominator the sample volume V is included, is negligibly small compared to the value of the first summand; therefore, the DL value does not depend on the volume of biomaterial. However, when we proceed to the case of irreversible binding with Kd = 0, the first term becomes equal to zero and the DL value is determined solely by the second term, Eq. (7). Therefore, for an AFM-chip with the above-indicated characteristics the LOD potential appear to be dependent on the sample volume: the larger sample volume, the higher the sensitivity of the method:

$$DL = \left\{ \begin{array}{l} 1.6 \times 10^{-18} M,\ V = 1\,mL \\ 1.6 \times 10^{-19} M,\ V = 10\,mL \\ 3.0 \times 10^{-20} M,\ V = 50\,mL \end{array} \right\} \quad (13)$$

3 Results and discussion

Proteomics routinely employs concentration detectors which are able to register the aggregate signal from a multitude of molecules. At the same time recent studies demonstrate the advantages of molecular detectors in recognizing signals from individual molecules and their complexes [13, 22]. Among the molecular detectors now in use are AFMs [41] and nanowire- [42] and nanopore-based [43] detectors.

The use of molecular, instead of concentration detectors makes it possible to switch from a regime of concentration measurements to a regime of counting individual molecules and their complexes. The calculated dependence presented in Fig. 1 illustrates the superiority of molecular detectors over concentration devices. Plotted on the abscissa are the values of the dissociation constants for various protein complexes in the concentration range of 10^{-1}–10^{-15} M. For instance, the Kd value is 10^{-13} M for the protein pair barnase/barstar [44]. Apparently, it is the lowest attainable threshold of dissociation constant for proteins whereas, in the case of antigen–antibody interactions, the Kd values vary over a range from 10^{-6} to 10^{-12} M [45]. If the concentrations of interacting proteins are equal to 10^{-7} M (curve I in Fig. 1),

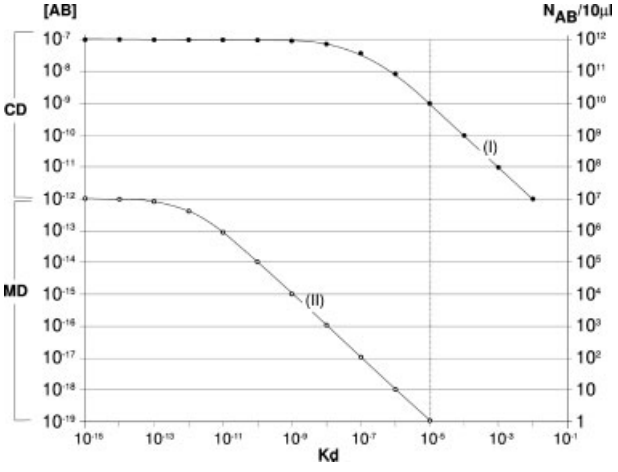

Figure 1. Calculated dependence of concentration [AB] and complex number N_{AB} on the dissociation constant Kd of a protein complex; the dependence is constructed according to Eqs. (2) and (5) on the assumption that the concentration of either of the interacting proteins is 10^{-7} M (I) or 10^{-12} M (II). "CD" indicates the restriction zone of concentration detectors; "MD" indicates the operational domain of molecular detectors.

then the concentration of the complex is also limited to 10^{-7} M (see plateau of curve I in Fig. 1). The registration of molecular complexes at concentrations as low as 10^{-7} M presents certain difficulties for proteomics [17]. However, as shown on the right axis of the graph, a concentration 10^{-7} M corresponds to 10^{12} complexes in a 10 μL of sample. If the Kd exceeds 10^{-8} M, its further growth proportionally diminishes the concentration of complexes [AB] and their number N_{AB}. With Kd = 10^{-5} M, the number of complexes in 10 μL is 10 billion. Detection of this number of complexes presents no problem for a molecular detector whereas the measurement of appropriate concentrations in the region of 10^{-9} M is at the limit of possibility of proteomic technologies.

Below 10^{-12} M (see curve II in Fig. 1) the application of concentration detectors becomes impossible. Although such concentration levels are formally within the limits of ELISA sensitivities, they are hardly attainable in the format of high-throughput proteomic experiments. Even given the assumption that concentration detectors with a sensitivity of 10^{-12} M may be constructed, the area of their application will be limited to Kd values below 10^{-13} M. Curve II shows that if Kd exceeds 10^{-13} M, the complex concentration drops out of the sensitivity domain of concentration detectors. Kd values of 10^{-9}–10^{-7} M are characteristic of many antigen/antibody complexes [18]; however, as seen from corresponding sector curve I, [AB] values as low as 10^{-15}–10^{-17} M are beyond the potential limits of concentration detectors.

On the other hand, the potentialities of molecular detectors are far from being exhausted in the above-indicated concentration range. By capturing individual objects with a height of several nanometers such detectors are able to register 100 and 10 000 molecular complexes for the concentrations 10^{-17} and 10^{-15} M, respectively. Thus, a problem that is principally intractable for concentration detectors may well be solved by the use of molecular detectors. In this study, we would like to demonstrate that application of nanotechnologies will help solve the problem of registration of biomolecules and their complexes at concentrations below 10^{-15} M.

The problem of detection of low-abundant proteins is fundamentally related to the primary aim of proteomics, which is to compile the complete set of all proteins existing in the human body. This raises the question as to how many different protein species occur in biomaterials. In other words, how many proteins would become available for inventory if proteomics was equipped with molecular detectors?

3.1 Dependence of the number of proteins on their concentration in plasma

To demonstrate the role of DL in proteomics, the dependence of the number of proteins detected on their concentration in depleted plasma was constructed. The analysis of such a dependence has been performed previously [26], but the region of interest was restricted to the concentration range of 10^{-9}–10^{-12} M. In our earlier study [13], the dependence obtained by Rose and coauthors was extrapolated down to the level of 10^{-18} M. At such a concentration, over 50 000 proteins were expected to be present in plasma.

In this study, another experimental approach is proposed: assessing the number of proteins as a function of their concentration in biomaterial. The function is conditional on two parameters of applied proteomic technology, the first of which, DL (*i.e.*, the lowest boundary of concentration), has already been described in Section 2.5. The second parameter, here called Q, relates to the loading capacity of a particular proteomic method and reflects the technological upper-boundary limit of protein content in a probed analyte.

The proposed parameters DL and Q offer, in fact, other insights into the yet unresolved problems of sensitivity and dynamic range. Consider the number of proteins N as a function of parameters DL and Q: $N = f(\mathrm{DL}, Q)$. Increasing the sensitivity of a procedure means that $\mathrm{DL} \to 0$ and, accordingly, the number of proteins detected in biomaterial is also increased: $N \to N_{max}$. Oppositely, if the total protein content in the sample is lowered, *i.e.*, $Q \to 0$, then $N \to 0$. Another possibility for determination of the total number of proteins is to increase the amount of biomaterial, as $N \to N_{max}$ if $Q \to \infty$. Such an approach has been taken [26]: a large amount (2.5 L) of plasma was loaded into the proteomic pipeline. Both $Q \to \infty$ and $\mathrm{DL} \to 0$ conditions have to be fulfilled to attain the concentration region of extremely low-abundant proteins.

Different methods for gel staining are assigned certain DL values. To be able to vary DL, we selected six different methods of in-gel protein visualization. These included: staining with AB and CBB as the least sensitive procedures; ST staining and silvering with GA as moderately sensitive procedures. In addition to visible staining, the fluorescent labeling of proteins with Cy5 dye in minimal and saturated (Cy5-sat) modes was used to attain the lowest values of DL for 2-DE-based technology [15]. In Fig. 2, there are six fragments of 1-DE, showing the BSA band area of the gel. Different concentrations of BSA were applied to each lane of 1-DE. For example, in the case of the topmost AB-stained gel fragment, BSA concentrations from 1 μg to 50 ng were probed. The lane loaded with 100 ng of protein was the last one where the software "find band" function indicated the presence of the BSA band. Thus, the appropriate concentration 1.2×10^{-7} M was selected as DL for AB staining.

In the same way, other methods of staining and labeling were assigned appropriate DLs. The next (from the top) fragment of CBB-stained gel revealed $\mathrm{DL} = 6.7 \times 10^{-8}$ M, whereas the ST and GA staining methods exhibited a one-order lower $\mathrm{DL} = 1.3 \times 10^{-9}$ M. Fluorescent labeling with Cy5 dye in minimal and saturation modes allowed even lower DL values to be achieved: 3.7×10^{-10} and 7.4×10^{-11} M, respectively.

Using the above-indicated staining methods, 2-DE of depleted plasma was performed for samples of different concentrations. For each staining method and each protein

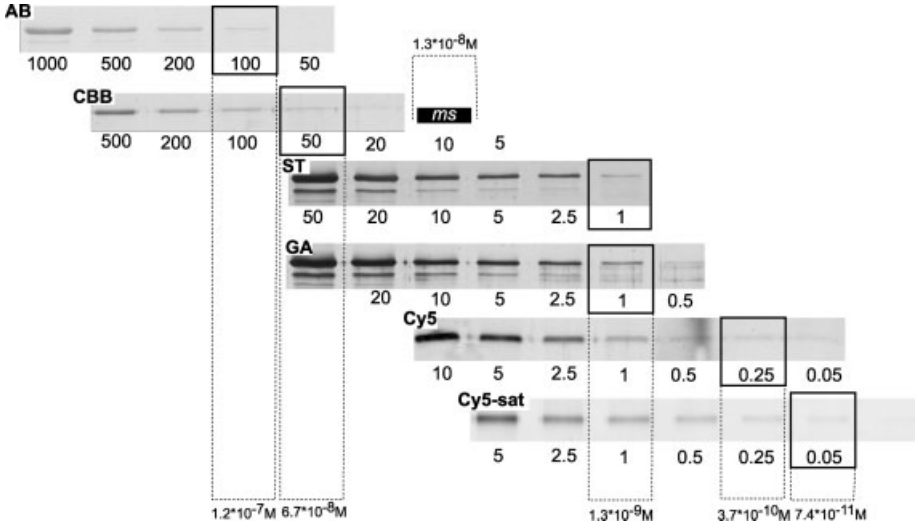

Figure 2. Determination of DL for various methods of gel staining: AB, amido black; CBB, – Coomassie brilliant blue; ST, silver thiosulfate; GA, silvering with glutaric aldehyde; Cy5, minimal Cy5 fluorescent labeling; Cy5-sat, saturated Cy5 labeling. Under each lane the protein content in ng BSA is indicated. At the bottom of the figure the amount of loaded BSA is given in concentration units. Solid squared frames indicate the last lanes where the BSA band was determined using the "find band" function in QuantityOne (BioRad) or FragmenterNT (Amersham Biosciences); "ms" denotes the band on the gel that was cut out for MS-identification. Electrophoresis was carried out in miniProtean III format in 12% SDS-PAGE at a constant voltage of 150 V. Onto each lane 10 μL of probe were deposited.

concentration the number of spots on the 2-DE image was calculated. For example, there were 101 spots on 2-DE of 1 μM depleted plasma stained by ST. The number of spots detectable on the image of 2-DE was considered to be an estimation of the number of proteins available for the corresponding staining method.

Dependencies of protein number on sample concentration are shown in Fig. 3. The shapes of these dependencies were similar for different methods of in-gel protein visualization *e.g.* with Cy5 labeling in the range of 2.5–20 μg the number of protein spots increased with increasing total protein content in the sample. At the level of 40 μg saturation was achieved. A further increase in protein content led to a decrease in the number of observed protein spots because the large spots overlapped the smaller ones.

As seen in Fig. 3a, the concentrations of saturation maximums were different, depending on the method of protein visualization. In the case of Cy5 labeling, the saturation maximum occurs at 25 μg of protein *per* 200 μL; the maxima of the GA and ST staining methods were reached at the level of about 50 μg, while the CB maximum corresponds to 200 μg. The largest amount of protein is required upon staining with AB reagent, for which saturation occurs at a load of 260 μg of total protein *per* sample.

The rectilinear portions of the curves displayed in Figs. 3a and b are the most important. These portions correspond to the concentration ranges within which direct proportionality is obeyed between the protein content in the biomaterial loaded onto the gel and the number of spots observed on 2-DE. The proportional dependence is retained until technical limitations bring the curve to saturation level.

Magnified views of the initial regions of the dependencies are given in Figs. 3c and d, corresponding to protein content up to 10 and 200 μg, respectively. These regions are well approximated with linear trends. Significance of the linear correlation coefficients exceeded 0.9 for each protein visualization method.

The trends obtained differed in their slopes. The Cy5-sat trend rose most steeply (see Fig. 3c) while the CBB trend had the smallest slope (Fig. 3d); other trends had intermediate slopes. Based on the trends of direct proportionality for each of the protein visualization methods it was possible to calculate the normalized number of proteins $Z = N/Q$ (here, Z is expressed in prot/μM units), where N is the number of proteins at the appropriate concentration Q. For Cy5-sat and Cy5 dyes, the values of Z were 412 and 230 prot/μM, respectively; for both GA and ST the corresponding values were several times less, ~100 prot/μM. The two weakest methods for gel staining were characterized by lower Z values: 16 prot/μM for AB and 11 prot/μM for CBB.

The proposed Z value is a constant characteristic of the in-gel protein visualization method (or, as shown later on, of the protein identification method). The dependence of Z on the DL, corresponding to each method, may be constructed. The dependence obtained and its linear approximation on a semilogarithmic scale are presented in Fig. 4. The significance of the correlation coefficient was $R^2 = 0.77$.

By extrapolating the approximate trend to the range of low DL values, we find that at DL = 10^{-24} M the value of Z is equal to 2218. This means that if the DL of some experimental procedure was at the level of the reverse Avogadro number, then 2218 proteins could be detected in 1 μM of

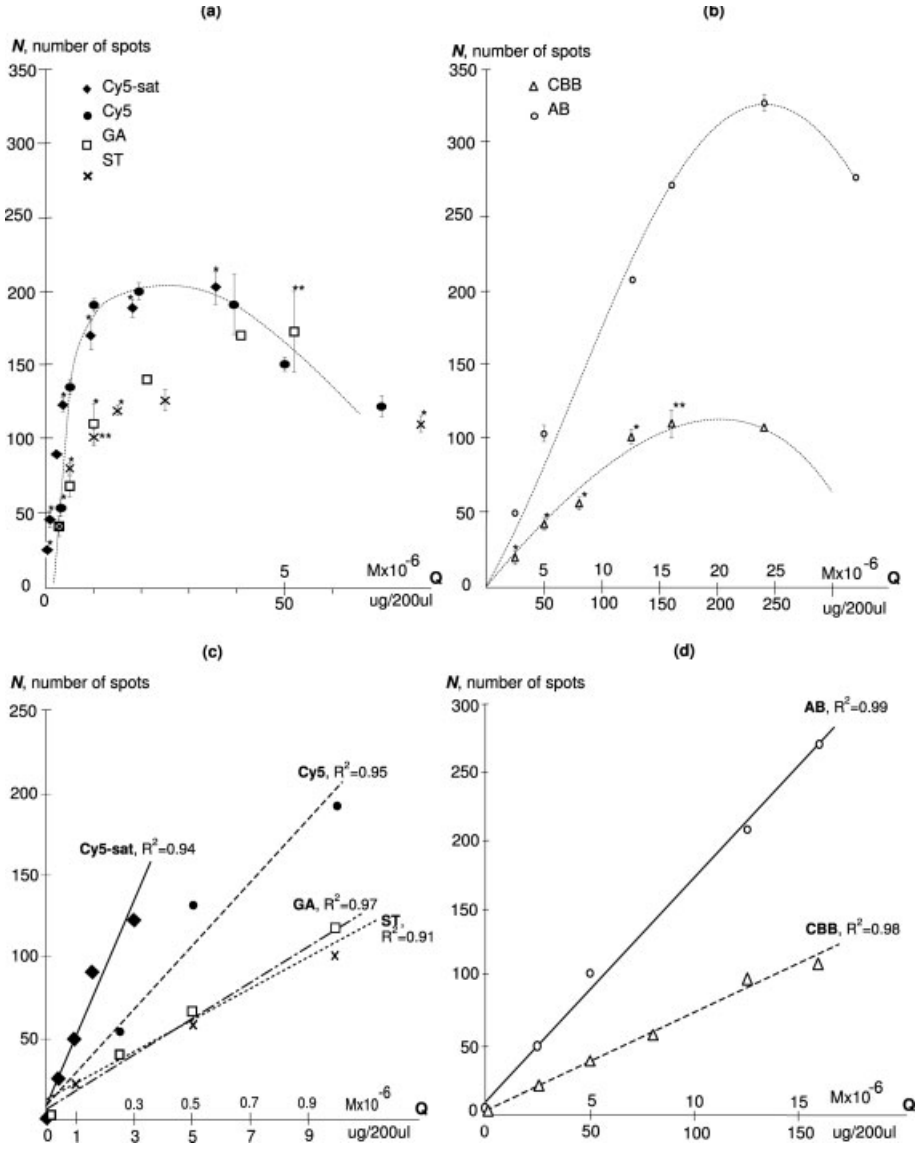

Figure 3. The dependence of the number of 2-DE-revealed spots *N* on the protein content *Q* of depleted plasma: (a,b) general view of the dependencies; (c,d) enlarged sections of the initial regions of proportionality between protein number and total protein content. SD from three (*) or at least four (**) different estimates is presented in (a,b).

depleted plasma. It appears, therefore, that an increase of DL from $\sim 10^{-8}$ M (for CBB and AB) to 10^{-24} M led to a two-order of magnitude increase in *Z*: from ~11–16 prot/μM to 2218 prot/μM. If Q_{max} is taken as the total protein content in plasma (80 mg/mL [46] or 10^{-3} M), the maximal number of proteins N_{max} may be estimated. Given that $Z = N_{max}/Q_{max}$, we obtain:

$N_{max} = Z \times Q_{max} = 2218$ prot/μM $\times 10^3$ μM $\approx 2 \times 10^6$ proteins.

Thus, proteomics is faced with the task of compiling an inventory of over 2 million proteins in plasma. The order of magnitude of this estimation coincides with the speculative expectations of other authors [12, 27] and us [13].

The diagram in Fig. 4 was obtained based on the data of the in-gel protein visualization methods; however, it may be supplemented by the data of the protein identification method. A proteomic experiment was carried out using a combination of protein separation by 1-DE with subsequent identification of the separated proteins by LC-MS/MS (1-D + LC-MS/MS). First, in order to determine the DL value for 1-DE + LC-MS/MS, a range of BSA concentrations was scanned. From CBB-stained 1-DE, the bands corresponding to the putative location of BSA were cut out (even if the protein band was not revealed by OD, see Fig. 2). After in-gel trypsinolysis the protein was identified by LC-MS/MS. It appeared that the statistically significant identification of BSA was possible if the band was cut out from the lane with a protein content of 10 ng (such a band is designated as "ms" in Fig. 2). For the band from the next lane with lower protein content (5 ng), no significant identification was reported by SpectrumMill software. Thus, the DL attained by 1-DE + LC-MS/MS was estimated as 1.3×10^{-8} M.

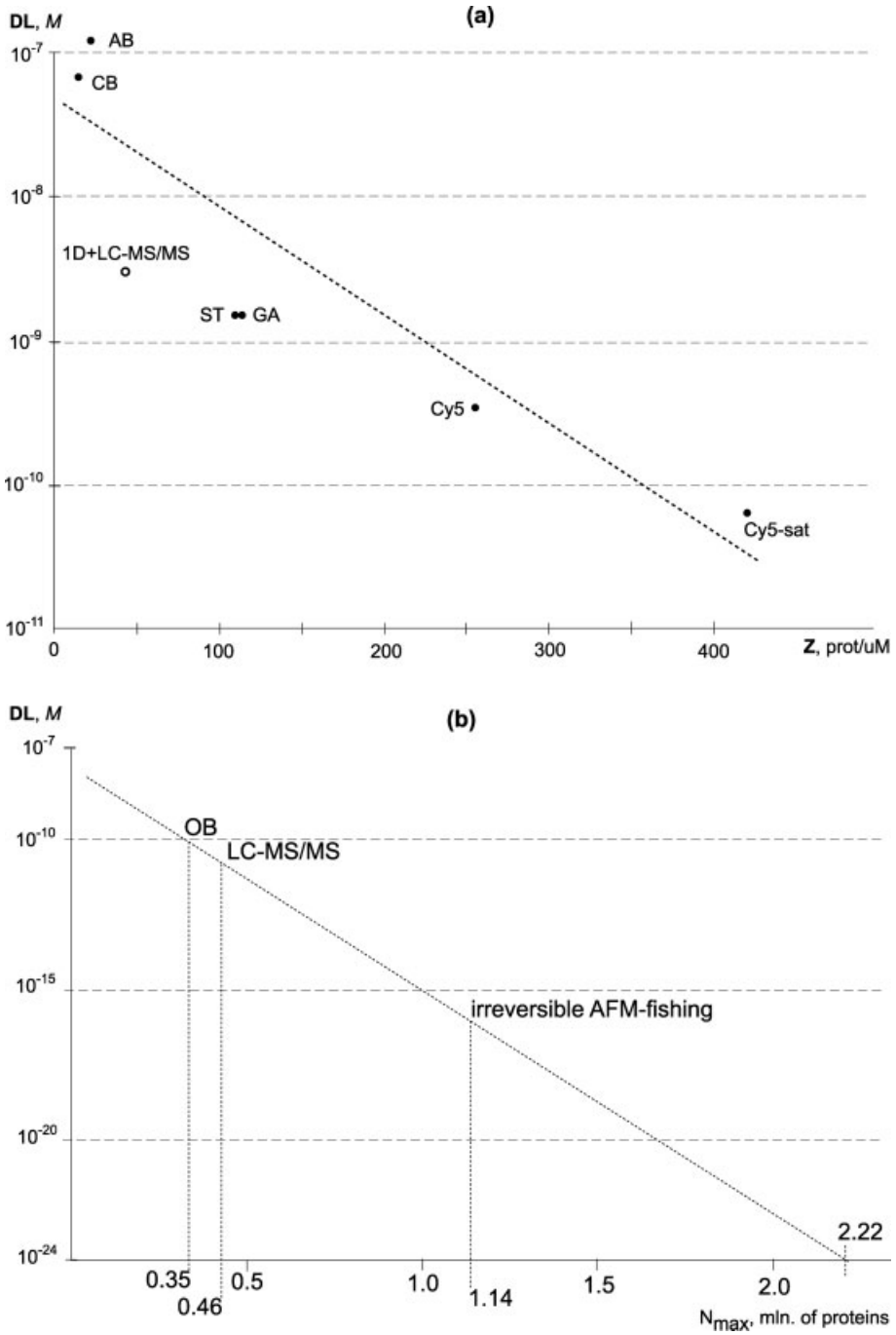

Figure 4. Dependence of DL on the normalized number of proteins Z in the range from 10^{-7} to 10^{-11} M (a) and its approximation (b) down to 10^{-24} M according to the formula $DL = 3 \times 10^{-8} \times e^{-0.017 \times Z}$. Correlation coefficient of approximation is $R^2 = 0.77$. $N_{max} = Z \times Q_{max}$, where Q_{max} is the protein plasma concentration, 10^{-3} M. See Fig. 2 for gel staining abbreviations.

In reality, the DL obtained by a combination of 1-DE with LC-MS/MS is below 1.3×10^{-8} M. The reason is that the proteolysis protocol recommends dissolving the band in 10 μL of trypsin solution (see Section 2.2.3) but afterwards only a restricted portion of the solution is sampled to LC. The sample volume constitutes 2.5 μL, *i.e.*, one-fourth of the initial amount. Therefore, the more reliable $DL = 3.3 \times 10^{-9}$ M was obtained by dividing by 4 the initially determined value. Even this DL appears to be too high as our amendments do not account for the effectiveness of digestion and peptide extraction from the gel.

After establishing DL for the BSA protein, we have performed the experiments with depleted plasma samples using the same 1-DE + LC-MS/MS method. Onto the lanes of 1-DE the 10-μL samples were deposited with protein content varying in the range of 1.25–40 μg (Fig. 5a). Each lane was cut into 24 strips which were analyzed by LC-MS/MS after in-gel trypsinolysis. The identification of proteins

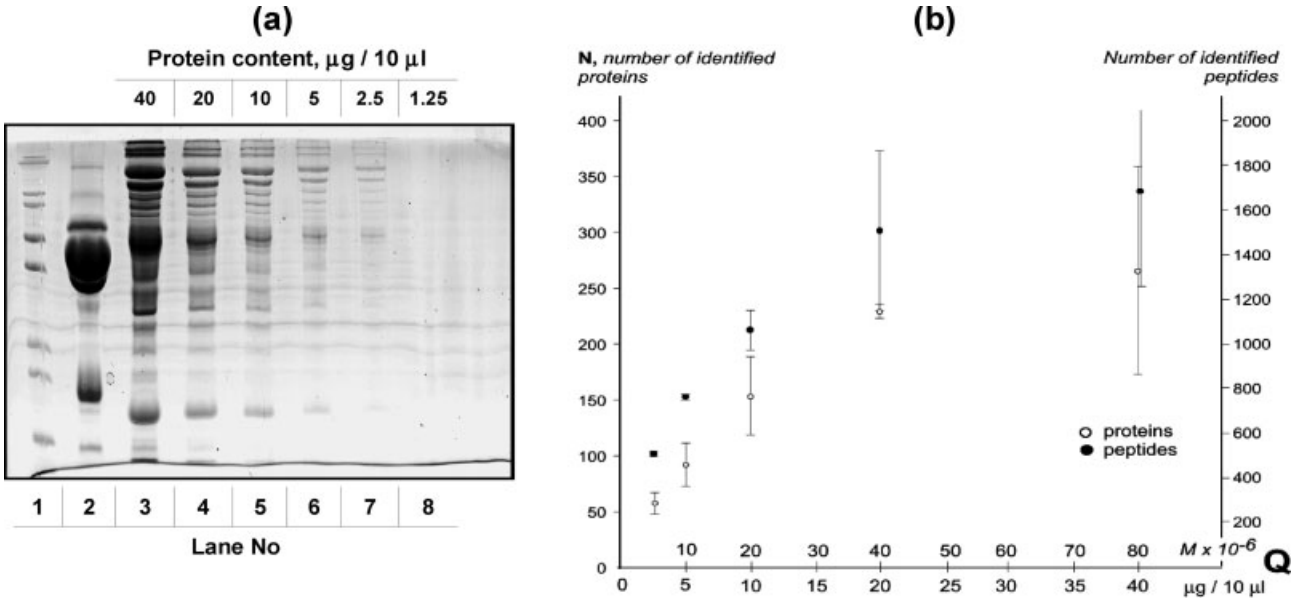

Figure 5. (a) SDS-PAGE of depleted plasma obtained at various concentrations of total protein (lanes 3–8). Lane 1 contains markers of MWs; lane 2 contains the fraction of high-abundant proteins, removed at the depletion stage. Staining with CBB. (b) Concentration dependence of the number of identified proteins and peptide fragments used for their identification in depleted plasma samples. SD from three different estimates is presented.

and their peptide fragments was carried out using SpectrumMill software.

Figure 5b demonstrates the changes in the numbers of proteins and their peptide fragments that occur with increasing total protein content in the sample. The initial virtual linear regions of the dependencies corresponding to the concentration range from 0 to 20 μM for proteins and from 0 to 40 μM for peptides are clearly seen. With a further increase in concentration, both the curves for peptides and proteins reached the saturation maximum (Fig. 5b).

Based on the proportionality between the number of identified proteins and the total protein concentration in the sample (curve I) the Z value was determined as $10 \times 4 = 40$ prot/μM (multiplication by 4 is necessary because of the above-mentioned restrictions of probe sampling). This means that by using such an MS-based proteomic technology it is theoretically possible to determine in blood plasma as many as $N_{max} = Z \times Q_{max} = 40$ prot/μM $\times 10^3$ μM = 40 000 proteins. When these new data were plotted in Fig. 4a, a slight decrease of the correlation significance from $R^2 = 0.77$ to 0.73 was observed.

The results obtained with the use of the 1-DE + LC-MS/MS approach make it possible to investigate the reason why proteomics faces the challenge of DL. In experiments with pure BSA preparations (see Fig. 2), it was established that DL for the most sensitive Cy5-sat was 7.4×10^{-11} M. Approximately the same DL was obtained under conditions in which trypsinolized BSA solution was directly sampled into LC-MS/MS, without prior 1-DE separation (Table 1). However, the combination of 1-DE with LC-MS/MS led to two-orders of magnitude loss in sensitivity compared to the individual method: DL increased up to 1.3×10^{-8} M (see the "ms" mark in Fig. 2). This occurred due to the loss of biomaterial upon electrophoretic separation and, also, upon in-gel trypsinolysis and subsequent extraction of peptides. Application of LC-MS/MS for plasma proteomics without preliminary in-gel fractionation could theoretically detect 465 000 proteins (Table 1) at DL = 10^{-11}, but under real experimental conditions the method is shackled by the complex peptide mixture and thus no more than 150 proteins could be identified [47].

Table 1. Expected total number of proteins available for proteomic technology with different DLs

Method of proteomic analysis	DL[a] (M)	Z (prot/μM)	Total number of proteins in plasma (N_{max})
1-DE (Cy5-sat labeling)	7.4×10^{-11}	412[b]	412 000
1-DE + LC-MS/MS	1.3×10^{-8}	40[c]	40 000
LC-MS/MS	2.0×10^{-11}	~465[d]	465 000

Total number of proteins $N_{max} = Z \times 10^3$ μM, where 10^3 μM is the concentration of proteins in plasma [46].

a) DL was determined in experiments with pure BSA.

b, c) Values of Z were determined based on the direct proportionality trends estimated in Figs. 3c and 5b, respectively.

d) Value of Z is predicted from the approximating trend in Fig. 4a.

The results above show that, in practical terms, the DL of MS-based proteomic technologies enables the detection of less than 1% of all the proteins occurring in plasma.

3.2 Experimental determination of DL for OB

The second increasingly spreading proteomic approach is based on protein microarray technology [2]. OB can possibly be considered as one of the universally applicable microarray readers [20, 21].

We have analyzed DL for the IAsys$^+$ "resonance mirror" biosensor by taking as an example the interactions of hepatitis B virus antigen and troponin C with their respective antibodies. The chosen pair of interacting proteins and the adopted immobilization schemes made it possible to accomplish the registration of analytes with relatively small and relatively large MWs. In the case of the anti-HBsAg/HBsAg pair, the 150-kDa antibodies were immobilized onto the carboxylated surface of the cuvette. The lighter HBsAg antigen with an MW of 25.5 kDa was introduced into the analyte solution. In contrast, in the second set of experiments the immobilized troponin C ligands had a small MW, 18 kDa, whereas the analyte molecules were the antibodies to troponin C with an MW of 150 kDa. In choosing two such pairs of interacting proteins two goals were pursued: (i) to assess the efficiency of the immobilization of ligands with different MWs; and (ii) to determine whether the MW of the analyte affects the DL of the method.

After the immobilization procedure (see Section 2), the OB response (R) increased by 500 and 110 U relative to baseline for the anti-HBsAg antibody and troponin C, respectively. By substituting the numerical values for R and MW into Eq. (9), the surface concentration of immobilized ligand molecules C_{BO} was deduced. It amounted to $C_{BO} = (7.2 \pm 0.7) \times 10^9$ molecules/mm^2 for anti-HBsAg antibody and to $C_{BO} = (6.2 \pm 0.6) \times 10^9$ molecules/mm^2 for troponin C. Despite the five-fold difference in the molecular masses of the ligands, one and the same number of ligands was immobilized in the same type of OB cuvette. We conclude that in our case the efficiency of immobilization does not depend on the molecular mass of the immobilized molecule.

Figures 6a and b present typical curves of the reversible binding for the anti-HBsAg$_{im}$/HBsAg and cTnC$_{im}$/anti-cTnC pairs, respectively. The amplification of the OB-response after addition of the analyte molecules was due to the formation of HBsAg$_{im}$/HBsAg and cTnC$_{im}$/cTnC complexes. The response increased for 7–10 min, then the cuvette was rinsed with buffer solution. After rinsing the process of complex dissociation was observed as evidenced by the descending part of the curve. Based on the concentration dependencies of the binding curves, the association and dissociation rate constants of the complexes formed as well as their affinities were calculated (see Table 2). The Kd values obtained were approximately the same for both protein pairs under study.

As shown in Figs. 6a and b, the amplitude of the OB response was diminished in a concentration-dependent manner depending on the concentration of HBsAg or anti-cTnC in the analyte solution. The DL values were determined as minimal analyte concentrations, where the S/N ratio of the OB response exceeded 1 (Table 2).

Experimentally measured DL values for the reversible interaction were dependent on the MW of the analyte. When the analyte solution contained 25-kDa HBsAg, the DL was 10^{-9} M. For the higher MW analyte molecule, which was 150-kDa anti-cTrC, a one-order lower DL = 10^{-10} M was estimated. By comparing the experimentally measured DL values to those obtained from Eqs. (10) and (11), it appears that the experimental values were two orders of magnitude lower than the theoretically calculated DLs.

The results of irreversible binding are presented in Figs. 6c and d. First, the solution of appropriate analyte was added to the cuvette with immobilized ligand and the interaction between ligand and analyte was recorded for 7–10 min. After that the UV-crosslinking of the ligand/analyte complexes was carried out as described in Section 2.4.2. As shown by the dashed arrows in Figs. 6c and d, after photocrosslinking an abrupt increase in OB response was registered. The gain in response amplitude under photocrosslinking conditions significantly exceeded the response of reversible binding, which has been extrapolated to the

Table 2. The values of k_{on}, k_{off}, Kd, and DL measured using an OB for the anti-HBsAg$_{im}$/HBsAg and cTnC$_{im}$/anti-cTnC pairs

Interacting proteins	k_{on}[a] ($\times 10^5$ M^{-1} s^{-1})	Reversible binding				Irreversible binding	
		k_{off}[a], ($\times 10^{-4}$ s^{-1})	Kd[b] ($\times 10^{-10}$ M)	DL$_{rev}$ (M)	S/N	DL$_{irrev}$ (M)	S/N
anti-HBs$_{im}$/HBsAg	1.50 ± 0.10	1.0 ± 0.3	6.7 ± 2.2	10^{-9}	2.0	10^{-10}	1.3
cTnC$_{im}$/anti-cTnC	0.70 ± 0.06	0.27 ± 0.07	3.9 ± 1.4	10^{-10}	3.0	10^{-10}	5.0

Sensograms were measured for the range of analyte concentrations down to 10^{-11} M and the DLs (DL$_{rev}$ and DL$_{irrev}$) were determined as the concentration of analyte at which the signal of the biosensor still exceeded the level of noise, *i.e.*, S/N was above 1.

a) k_{on} and k_{off} are association and dissociation rate constants, respectively.

b) Kd is the dissociation constant: Kd = k_{off}/k_{on}.

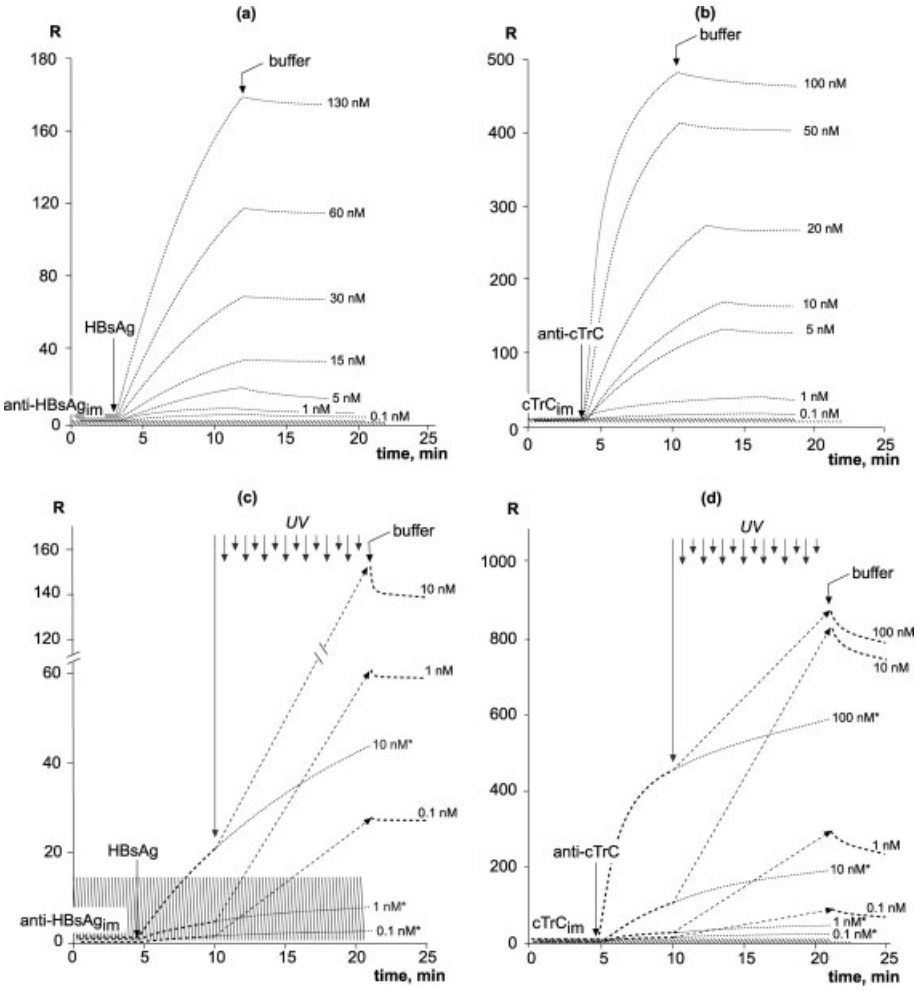

Figure 6. The concentration dependencies of the reactions of complex formation and decay, obtained by using of an OB. (a,b) Reversible complex formation: (a) anti-HBsAg$_{im}$/HBsAg and (b) cTrC$_{im}$/anti-cTrC. The incubation medium contained PBS/t, pH 7.4. Arrows indicate addition of analyte and replacement of the incubation medium by buffer solution (PBS/t). (c,d) Irreversible complex formation: (c) anti-HBsAg$_{im}$/HBsAg and (d) cTrC$_{im}$/anti-cTrC. The incubation medium contained PBS/t, pH 7.4. Arrows indicate time points of addition of the analyte molecules and cuvette rinsing. To the cuvette was added 1 μL of DTBPA photocrosslinker dissolved in 2×10^{-5} M DTBPA in DMSO. UV irradiation was conducted for 10 min (λ>300 nm). Asterisks show extrapolation of the curves for reversible complex formation. Images were vectorized from the original sensograms. Shaded areas designate the noise level.

same time point (see asterisks in the figure). This distinction in response levels indicates that covalent binding actually takes place.

Since the covalent binding is time-dependent, the efficiency of the process is not absolute and there always exists some portion of noncovalently bound molecules. When rinsing the cuvette after UV-irradiation, the dissociation of such complexes was observed and registered as descending parts of the curves in Figs. 6c and d. From Fig. 6c, it is seen that the signal for 0.1 nM concentration of HBsAg exceeded by 1.3 times the level of noise. Decreasing the analyte concentration below 0.1 nM did not allow us to register the OB response. For anti-cTrC it was also 0.1 nM – the lowest analyte concentration where S/N achieved 5. Therefore, 10^{-10} M appeared to be the DL for both investigated protein pairs (Table 2).

The experiments with the use of the OB have shown that:

(i) under reversible conditions the DL was 10^{-9} M for the anti-HBsAg$_{im}$/HBsAg pair and 10^{-10} M for the cTnC$_{im}$/anti-cTnC pair, with Kd values $(6.7 \pm 2.2) \times 10^{-10}$ M and $(3.9 \pm 1.4) \times 10^{-10}$ M, respectively;

(ii) removal of the thermodynamic limitations by switching to irreversible binding did not lead to a decrease in DL below 10^{-10} M. This situation was apparently determined by two factors: the restricted volume of biomaterial in the cuvette and/or the limited sensitivity of the concentration detector;

(iii) from the approximation in Fig. 4b one can see that by using a biosensor-based microarray with DL = 10^{-10} M it is possible to detect 350 000 proteins. This number accounts for ~20% of the total number of proteins in plasma.

3.3 Determination of DL by a combination of AFM and biospecific fishing

The application of molecular detectors enables us to remove the limitations on sensitivity. The theoretical possibility of DL reduction by the use of molecular detectors was discussed based on the data displayed in Fig. 1. In this study, AFM was employed as a molecular detector. However, the application of molecular detectors for the registration of molecules and molecular complexes, spread over the sample

volume, requires first their concentration during the biospecific fishing procedure.

The distinguishing feature of fishing compared to seemingly analogous procedures based on specific antibody/antigen interactions in microarrays is that the AFM-chip can be incubated in a larger volume of biomaterial. In the course of fishing, the analyte molecules that are spread throughout the whole sample volume are concentrated on the relatively small surface of the AFM-chip. In this study, procedures for the reversible and irreversible fishing of HCVcoreAg (MW = 18 kDa) from analyte solutions in the concentration range of 10^{-8}–10^{-17} M were conducted. The MW of HCVcoreAg is less than that of the immobilized HCVcore antibodies (the latter have MW = 150 kDa). This makes it possible to demonstrate the potentialities of AFM-based revelation of the complexes in cases when the mass and therefore the physical dimensions of the analyte molecule are less than those of the ligand.

Control images of the surface with immobilized HCVcore antibodies (MW = 150 kDa) were obtained by incubation of AFM-chip in antigen-free solution (Fig. 7a). The antibodies were visualized on the image of the AFM-chip as objects with characteristic heights below 2 nm (see inset in Fig. 7a). The distribution of the heights of the objects shown in Fig. 7a has its maximum at 1.5±0.2 nm.

Reversible fishing was carried out using a solution containing HCVcoreAg at concentrations of 10^{-8}–10^{-12} M. Figures 7b and c display the histograms of height distributions for AFM-detected objects, obtained after chip incubation in 10^{-9} and 10^{-11} M analyte solutions, respectively. One can see (insets in Figs. 7b and c), on the surface of the AFM-chip incubated in analyte solution, the appearance of the objects with greater heights (above 2.0 nm). The proportion of objects with heights above 2 nm increased depending on concentration [A_0]: this may be seen upon comparison of the height distributions for analyte concentrations of 10^{-9} and 10^{-11} M (Figs. 7b and c) with the histogram in Fig. 7a. We conclude that the number objects higher than 2 nm constituted a quantitative measure of the amount of ligand/analyte complexes. Comparison of the areas encompassed between control and experimental histograms (Figs. 7b and c) shows that the lower concentration 10^{-11} M corresponded to the lower square of the area, *i.e.*, to the smaller amount of complexes.

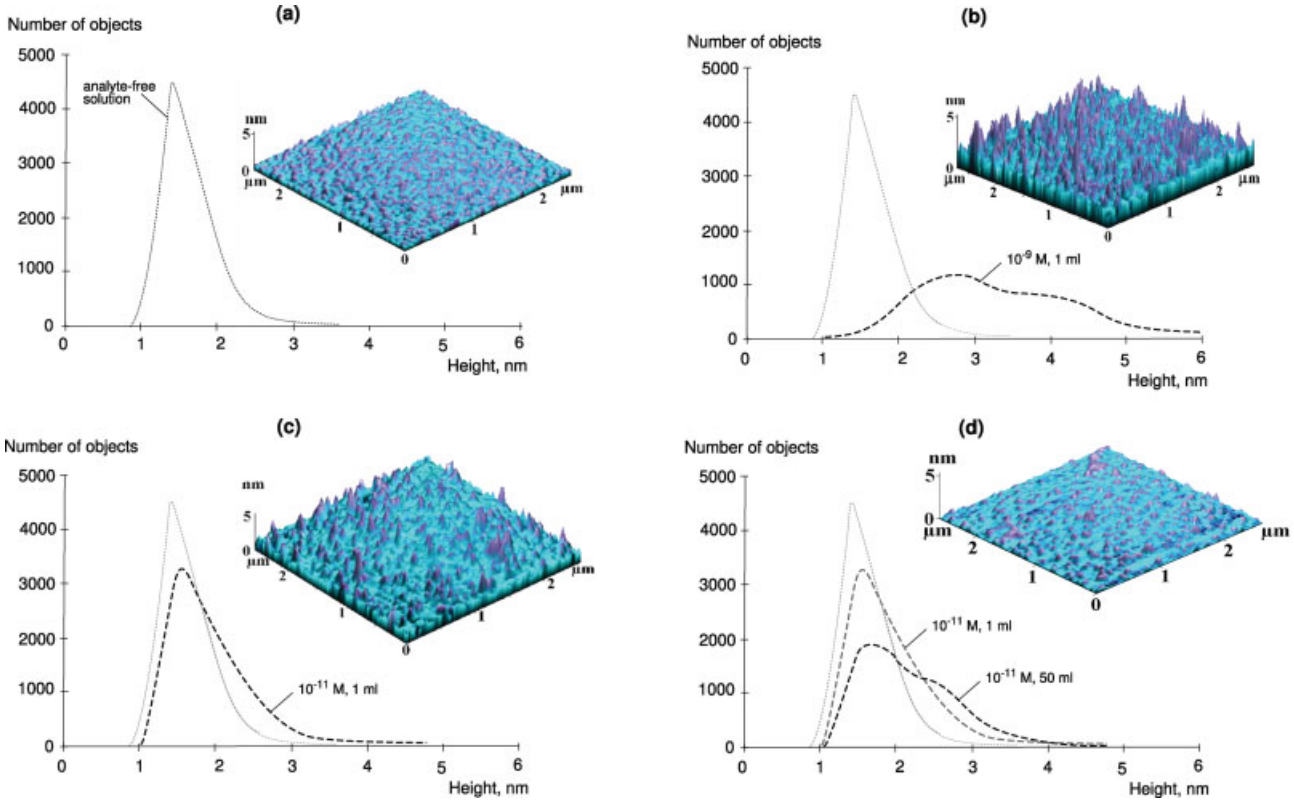

Figure 7. Height histograms after reversible fishing of HCVcore antigen onto the surface of the AFM-chip and the corresponding AFM-images. 400 μm^2 AFM-chip with 15 000 ± 1000 immobilized anti-HCVcore antibodies was incubated for 1 h at 37°C in 1 mL of 50 mM PBS buffer solution of HCVcore at concentrations of: (a) 0 M – analyte-free incubation mixture, (b) 10^{-9} M, and (c) 10^{-11} M; (d) AFM-chip was incubated in 50 mL of 10^{-11} M analyte solution. After washing twice with pure water and drying in air, the chip area was scanned using NTEGRA (NT-MDT) in a tapping mode.

In case of reversible binding, the increase of analyte volume did not affect on complex formation, as shown in Fig. 7d. The AFM-chip was incubated in 50 mL of 10^{-11} M analyte solution, however, the resulting heights' histogram did not differ significantly from the same histogram, obtained for the 1 mL of analyte (Fig. 7c). This observation corresponded to the theoretical calculations given in Section 2.5.2, where we have shown, that the effect of increased sample volume was indistinguishable due to the reversibility of Kd-dependent complex formation.

Figure 8a presents the analyte-free case for irreversible fishing, when the analyte concentration [A_0] was zero. The results of fishing under irreversible conditions for analyte concentrations 10^{-13}–10^{-16}M are presented in Figs. 8b, c, and d. In Figs. 8b and c, the images were obtained as a result of incubation of the AFM-chip in a 1-mL volume, whereas in Fig. 8d one can see the image obtained for a concentration of 10^{-16} M after AFM-chip incubation in a 50-mL volume of analyte. Analogously to results shown in Figs. 7, 8b, c, and d demonstrate the rise of the right wing of height distribution in the region of objects higher than 2 nm. The degree of this rise was easily noticeable for the analyte concentration 10^{-13} M (Fig. 8b) and became almost negligible for the lower concentration 10^{-16} M (Fig. 8c). Since under reversible conditions objects above 2 nm practically disappeared at the concentration level 10^{-12} M, it was concluded that under irreversible conditions objects higher than 2.0 nm corresponded to covalently bound complexes between chip-immobilized antibodies and HCVcoreAg.

Figure 9 presents the dependencies of the number of anti-HCVcore$_{im}$/HCVcoreAg complexes on the analyte titrated in the concentration range of 10^{-8}–10^{-12} M. This figure shows experimental as well as theoretical dependencies for a specific AFM-chip. The theoretical curves for reversible (Fig. 9, curve I) and irreversible (Fig. 9, curve II) binding were obtained according to Eq. (5) by substituting the following numerical values: the number of immobilized molecules $N = 15\,000$; sample volume $V = 1$ mL; and dissociation constant for the anti-HCVcore$_{im}$/HCVcoreAg complex was either Kd = 10^{-12} M (the lowest Kd value between any antigen and any antibody) for reversible binding or Kd = 0 for the irreversible case.

The theoretical dependence between analyte concentration and the number of antigen/antibody complexes is represented in Fig. 9 by curve I. Curve I shows that under reversible conditions a decrease in HCVcore antigen concentration from 10^{-8} to 10^{-11} M theoretically leads to a decrease in the number of objects on the AFM-chip surface. In accordance with the calculated dependence, the experimentally detected number of complexes decreased from

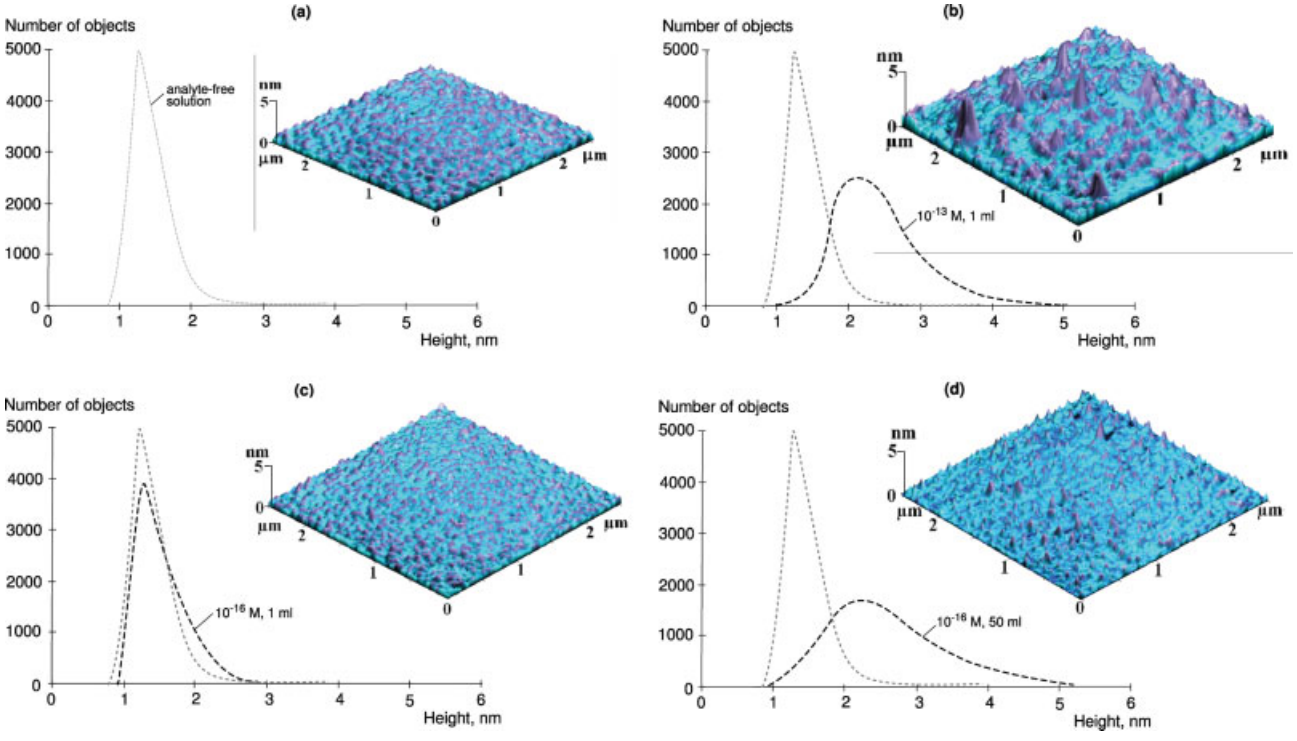

Figure 8. Height histograms after irreversible fishing of HCVcore antigen onto the surface of the AFM-chip and the corresponding AFM-images. 400 μm^2 AFM-chip with 15 000 $\pm$ 1000 immobilized photoreactive anti-HCVcore antibodies was incubated under UV irradiation (λ>300 nm) for 1 h at 37°C in 1 mL of 50 mM PBS buffer solution of HCVcore at concentrations of: (a) 0 M – analyte-free incubation mixture, (b) 10^{-13} M, and (c) 10^{-16} M. (d) Incubation of the chip in 50 mL of 10^{-16} M analyte solution. After washing twice with pure water and drying in air, the chip area was scanned using NTEGRA (NT-MDT) in a tapping mode.

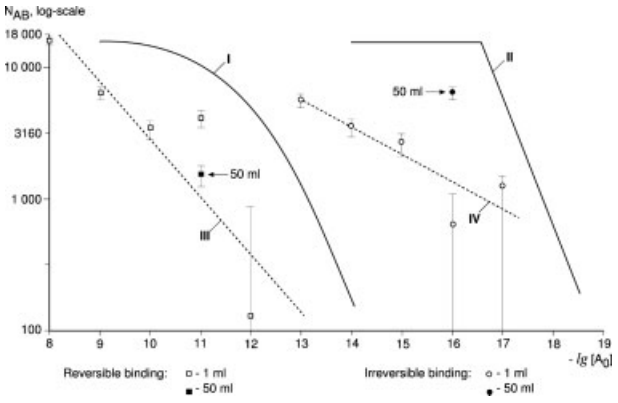

Figure 9. Dependence of the number of antigen/antibody complexes N_{AB} on the HCVcore antigen concentration [A_0]. I and II – theoretical curves obtained from Eq. (6) assuming that Kd = 10^{-12} M and Kd = 0, respectively. III and IV – linear approximations of experimental data for the cases of reversible and irreversible AFM fishing, respectively. Deviation from the mean was taken as 1000 objects to indicate the measuring error of the experiment.

15 000 ± 1000 to 4000 ± 1000 with decreasing concentration of HCVcore antigens in the sample from 10^{-8} to 10^{-11} M. With decreasing concentrations below 10^{-11} M, the number of complexes detected on the AFM-chip became less than the measurement error, which was taken as 1000. Thus, the DL for reversible fishing on the AFM-chip surface was determined as 10^{-11} M, *i.e.*, it was one order of magnitude higher than was expected from Eq. (12).

A 50-fold increase in sample volume with the same analyte concentration 10^{-11} M was also tried for incubating the AFM-chip. Consistent with the theoretical calculations according to Eq. (5), no corresponding increase in the number of complexes was observed on the AFM-chip after its incubation in such a large sample volume under conditions of reversible fishing.

The theoretical dependence for the case of irreversible binding (curve II in Fig. 9) was obtained by taking Kd = 0 for Eq. (5). Comparison of curves I–II in Fig. 9 shows that under irreversible conditions the complexes should be present at a lower concentration (below 10^{-16} M) than under reversible conditions.

In a set of experiments on irreversible binding, the AFM-chip was incubated in 1 mL of analyte solution, starting from a concentration 10^{-13} M and ending with 10^{-17} M (curve IV in Fig. 9). There were 6000 ± 1000 objects on the image of the AFM-chip incubated in 1 mL of 10^{-13} M HCVcoreAg solution. A further decrease in analyte concentration down to 10^{-15} M led to a proportional reduction in the number of recorded objects to 3000 ± 1000. At 10^{-16} M and, also, at 10^{-17} M concentrations the number of ligand-analyte complexes was less than the accepted error level of 1000 objects *per* AFM-chip. Therefore, the DL for biospecific irreversible AFM fishing was determined as 10^{-15} M.

Previous calculations have shown that, in contrast to reversible binding, under irreversible binding conditions the DL is dependent on sample volume. According to Eq. (13) by increasing the sample volume it is possible to lower the DL. The corresponding experiment was performed for the analyte concentration 10^{-16} M using a 50-mL sample volume instead of the usual 1-mL volume. Incubation of the AFM-chip in the 50-fold larger volume allowed detection of a total of 6000 ± 1000 objects (in three replicates of the experiment). The number of objects obtained significantly exceeded the statistical threshold; therefore for the 50-mL sample volume the concentration 10^{-16} M is considered to be the DL of biospecific AFM fishing.

Experiments on the biospecific fishing of HCVcoreAg onto the AFM-chip surface allow the following conclusions to be drawn:

(i) under reversible binding conditions it is impossible to lower the DL of biospecific AFM fishing below 10^{-11} M. In addition, in this case the volume of sample does not affect the number of complexes observed on the AFM-chip surface;

(ii) irreversible binding conditions enable lowering the DL down to the level of 10^{-16} M. As the DL is dependent on the volume of the analyte (see Eq. 12), increasing the volume from 1 to 50 mL decreases the DL accordingly by an order of magnitude.

4 Concluding remarks

Each technological platform applied to proteomic research has a specific DL. For the most sensitive Cy5 saturated method of gel staining, the DL was determined at a level of $\sim 10^{-10}$ M (see Fig. 2). According to the approximation provided in Fig. 4, this DL enables detection of at most 400 000 proteins, *i.e.*, no more than 20% of the total number of plasma proteins. Even more limited are the capabilities of 1-DE + LC-MS/MS technology, where the DL was established as being equal to 10^{-8} M. At this DL only 40 000 proteins constituting approximately 2% of the total proteome would be available for investigation (Table 1).

Presently the capabilities of microarray technology coupled with an OB reader are also at the level of DL = 10^{-10} M. With the AFM-based method we achieved a reduction in DL of four orders of magnitude compared to the lowest levels available using the usual proteomic technologies. Such a breakthrough towards DL = 10^{-16} M became possible due to the combination of biospecific irreversible fishing of analyte molecules from large sample volumes onto the chip-immobilized ligands with subsequent counting of ligand/analyte complexes by an AFM molecular detector.

At the technically attainable level of 10^{-16} M there are more than 1 million proteins available for analysis, constituting approximately half of the total number of proteins existing in plasma. In our opinion, the ability to distinguish between 1 million different individual proteins in a

high-throughput mode would essentially facilitate the requirement for reliable biomarkers of the earlier disease stages.

The authors have declared no conflict of interest.

5 References

[1] Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J. *et al.*, Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 1988, *239*, 487–491.

[2] Cutler, P., Protein arrays: The current state-of-the-art. *Proteomics* 2003, *3*, 3–18.

[3] Samatov, T. R., Chetverina, H. V., Chetverin, A. B., Expressible molecular colonies. *Nucleic Acids Res.* 2005, *33*, 300–302.

[4] Nakano, M., Komatsu, J., Matsuura, S., Takashima, K. *et al.*, Single-molecule PCR using water-in-oil emulsion. *J. Biotechnol.* 2003, *102*, 117–124.

[5] Palmer, S., Wiegand, A. P., Maldarelli, F., Bazmi, H. *et al.*, New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type 1 RNA in plasma. *J. Clin. Microbiol.* 2003, *41*, 4531–4546.

[6] Li, H., Cui, X., Arnheim, N., Direct electrophoretic detection of the allelic state of single DNA molecules in human sperm by using the polymerase chain reaction. *Proc. Natl. Acad. Sci. USA* 1990, *87*, 4580–4584.

[7] Garbis, S., Lubec, G., Fountoulakis, M., Limitations of current proteomics technologies. *J. Chromatogr. A* 2005, *1077*, 1–18.

[8] Wilkins, M. R., Appel, R. D., Eyk, J. E. V., Chung, M. C. M. *et al.*, Guidelines for the next 10 years of proteomics. *Proteomics* 2006, *6*, 4–8.

[9] Anderson, N. L., Anderson, N. G., The human plasma proteome: History, character, and diagnostic prospects. *Mol. Cell. Proteomics* 2002, *1*, 845–867.

[10] Corthals, G. L., Wasinger, V. C., Hochstrasser, D. F., Sanchez, J. C., The dynamic range of protein expression: A challenge for proteomic research. *Electrophoresis* 2000, *21*, 1104–1115.

[11] Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., Aebersold, R., Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Natl. Acad. Sci. USA* 2000, *97*, 9390–9395.

[12] Nielsen, M. L., Savitski, M. M., Zubarev, R. A., Extent of modifications in human proteome samples and their effect on dynamic range of analysis in shotgun proteomics. *Mol. Cell. Proteomics* 2006, *5*, 2384–2391.

[13] Archakov, A. I., Ivanov, Y. D., Lisitsa, A. V., Zgoda, V. G., AFM fishing nanotechnology is the way to reverse the Avogadro number in proteomics. *Proteomics* 2007, *7*, 4–9.

[14] Johnson, C. J., Zhukovsky, N., Cass, A. E. G., Nagy, J. M., Proteomics, nanotechnology and molecular diagnostics. *Proteomics* 2008, *8*, 715–730.

[15] Miller, I., Crawford, J., Gianazza, E., Protein stains for proteomic applications: Which, when, why? *Proteomics* 2006, *6*, 5385–5408.

[16] Martin, S. E., Shabanowitz, J., Hunt, D. F., Marto, J. A., Subfemtomole MS and MS/MS peptide sequence analysis using nano-HPLC micro-ESI Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* 2000, *72*, 4266–4274.

[17] Anderson, L., Candidate-based proteomics in the search for biomarkers of cardiovascular disease. *J. Physiol.* 2005, *563*, 23–60.

[18] Varnum, S. M., Woodbury, R. L., Zangar, R. C., A protein microarray ELISA for screening biological fluids. *Methods Mol. Biol.* 2004, *264*, 161–172.

[19] Rad'ko, S. P., Rakhmetova, S. I., Bodoev, N. V., Archakov, A. I., Aptamers as perspective affine reagents for clinical proteomics. *Biomed. Khim.* 2007, *53*, 5–24.

[20] Yu, X., Xu, D., Cheng, Q., Label-free detection methods for protein microarrays. *Proteomics* 2006, *6*, 5493–5503.

[21] Wingren, C., Borrebaeck, C. A. K., Progress in miniaturization of protein arrays – A step closer to high-density nanoarrays. *Drug Discov. Today* 2007, *12*, 813–819.

[22] Ivanov, Y. D., Govorun, V. M., Bykov, V. A., Archakov, A. I., Nanotechnologies in proteomics. *Proteomics* 2006, *6*, 1399–1414.

[23] Ghatnekar-Nilsson, S., Dexlin, L., Wingren, C., Montelius, L., Borrebaeck, C. A. K., Design of atto-vial based recombinant antibody arrays combined with a planar wave-guide detection system. *Proteomics* 2007, *7*, 540–547.

[24] Sinensky, A., Belcher, A., Label-free and high-resolution protein/DNA nanoarray analysis using Kelvin probe force microscopy, *Nat. Nanotech.* 2007, *2*, 653–659.

[25] Omenn, G. S., States, D. J., Adamski, M., Blackwell, T. W. *et al.*, Overview of the HUPO Plasma Proteome Project: Results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* 2005, *5*, 3226–3245.

[26] Rose, K., Bougueleret, L., Baussant, T., Boehm, G. *et al.*, Industrial-scale proteomics: From liters of plasma to chemically synthesized proteins. *Proteomics* 2004, *4*, 2125–2150.

[27] Hunsucker, S. W., Duncan, M. W., Is protein overlap in two-dimensional gels a serious practical problem? *Proteomics* 2006, *6*, 137–145.

[28] Archakov, A. I., Ivanov, Y. D., Analytical nanobiotechnology for medicine diagnostics. *Mol. Biosyst.* 2007, *3*, 336–342.

[29] Kiselyova, O. I., Yaminsky, I. V., Ivanov, Y. D., Kanaeva, I. P. *et al.*, AFM study of membrane proteins, cytochrome P450 2B4, and NADPH-cytochrome P450 reductase and their complex formation. *Arch. Biochem. Biophys.* 1999, *371*, 1–7.

[30] Kuznetsov, V. Y., Ivanov, Y. D., Archakov, A. I., Atomic force microscopy revelation of molecular complexes in the multiprotein cytochrome P450 2B4-containing system. *Proteomics* 2004, *4*, 2390–2396.

[31] Hansma, P. K., Schitter, G., Fantner, G. E., Prater, C., Applied physics. High-speed atomic force microscopy. *Science* 2006, *314*, 601–602.

[32] Vettiger, P., Despont, M., Drechsler, U., Duerig, U. *et al.*, The Millipede – More than one thousand tips for future AFM data storage. *IBM J. Res. Develop.* 2000, *44*, 323–339.

[33] Wirth, P. J., Romano, A., Staining methods in gel electrophoresis, including the use of multiple detection methods. *J. Chromatogr. A* 1995, *698*, 123–143.

[34] Neuhoff, V., Arold, N., Taube, D., Ehrhardt, W., Improved staining of proteins in PAGEs including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 1988, *9*, 255–262.

[35] Shevchenko, A., Wilm, M., Vorm, O., Mann, M., Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* 1996, *68*, 850–858.

[36] Wray, W., Boulikas, T., Wray, V. P., Hancock, R., Silver staining of proteins in polyacrylamide gels. *Anal. Biochem.* 1981, *118*, 197–203.

[37] Mikkelsen, R. B., Wallach, D. F., Photoactivated cross-linking of proteins within the erythrocyte membrane core. *J. Biol. Chem.* 1976, *251*, 7413–7416.

[38] Mattson, G., Conklin, E., Desai, S., Nielander, G. *et al.*, A practical approach to crosslinking. *Mol. Biol. Rep.* 1993, *17*, 167–183.

[39] Qureshi, N., Perera, P.-Y., Shen, J., Zhang, G. *et al.*, The proteasome as a lipopolysaccharide-binding protein in macrophages: Differential effects of proteasome inhibition on lipopolysaccharide-induced signaling events. *J. Immunol.* 2003, 1515–1525.

[40] Ivanov, Y. D., Usanov, S. A., Archakov, A. I., Optical biosensor studies on the productive complex formation between the components of cytochrome P450scc dependent monooxygenase system. *Biochem. Mol. Biol.* 1999, *47*, 327–336.

[41] Hinterdorfer, P., Schuetz, G., Kienberger, F., Schindler, H., Detection and characterization of single biomolecules at surfaces. *J. Biotechnol.* 2001, *82*, 25–35.

[42] Patolsky, F., Zheng, G., Hayden, O., Lakadamyali, M. *et al.*, Electrical detection of single viruses. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 14017–14022.

[43] Li, J., Gershow, M., Stein, D., Brandin, E., Golovchenko, J. A., DNA molecules and configurations in a solid-state nanopore microscope. *Nat. Mater.* 2003, *2*, 611–615.

[44] Hartley, R. W., Directed mutagenesis and barnase-barstar recognition. *Biochemistry* 1993, *32*, 5978–5984.

[45] Archakov, A. I., Govorun, V. M., Dubanov, A. V., Ivanov, Y. D. *et al.*, Protein-protein interactions as a target for drugs in proteomics. *Proteomics* 2003, *3*, 380–391.

[46] Horvatovich, P., Govorukhina, N., Bischoff, R., Biomarker discovery by proteomics: Challenges not only for the analytical chemist. *Analyst* 2006, *131*, 1193–1196.

[47] Liu, T., Qian, W.-J., Mottaz, H. M., Gritsenko, M. A. *et al.*, Evaluation of multiprotein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry. *Mol. Cell. Proteomics* 2006, *5*, 2167–2174.

