

Quality assurance in immunoassay performance—comparison of different enzyme immunoassays for the determination of caffeine in consumer products

Julia Grandke · Lidia Oberleitner · Ute Resch-Genger ·
Leif-Alexander Garbe · Rudolf J. Schneider

Received: 17 August 2012 / Revised: 19 November 2012 / Accepted: 21 November 2012 / Published online: 7 December 2012
© Springer-Verlag Berlin Heidelberg 2012

Abstract Enzyme immunoassays with optical detection are amongst the most widely used bioanalytical tools. We defined seven parameters for the quality assessment of immunoassays that were addressed in a systematic study of direct and indirect immunoassays, using the enzymes horseradish peroxidase (HRP) and alkaline phosphatase (AP), the chromogenic substrates 3,3',5,5'-tetramethylbenzidine (TMB) and *para*-nitrophenyl phosphate, and the fluorescent substrates 3-(4-hydroxyphenyl)propionic acid and 4-methylumbelliferyl phosphate. The same monoclonal antibody against caffeine was used throughout the study. The four quality parameters regarding the standard curve were the test midpoint (sensitivity), the measurement range, the relative dynamic range of the signal, and the goodness of fit of the adjusted four-parameter logistic function. All HRP immunoassays showed a higher sensitivity compared to the AP assays. On the basis of all four criteria, it was established that the direct assay format is superior to the indirect format, the immunoassay using HRP TMB fulfilling all requirements best. In a second step, caffeine concentrations in 24 beverage and cosmetics samples were determined and three more quality parameters were assessed with this application. The direct

HRP TMB assay showed one of the best intra- and inter-plate precisions and the best accuracy, defined by the correlation of results with those from the chosen reference method liquid chromatography tandem mass spectrometry (LC-MS/MS). Considering all criteria, HRP TMB seems to be the enzyme substrate system of choice preferably used in the direct assay format.

Keywords Enzyme immunoassay · Caffeine · Enzyme substrate · Competitive assay · Assay format · Direct · Indirect

Introduction

High-quality immunoassays possess an excellent selectivity, sensitivity, precision, accuracy, and practicality. In order to obtain such performance, optimization of assay conditions is crucial. The size of the analyte determines the possible test principle and the choice of assay format. Large molecules are able to bind at least two distinct antibodies at different epitopes (sandwich immunoassay, non-competitive format). Smaller molecules are unable to bind more than one antibody at the same time, so typically, a competitive format is used in which labeled antigen competes with unlabeled antigen for a limited number of antibody binding sites [1]. Low molecular weight analytes of interest exist in the fields of food quality, safety testing, drug screening, and represent most environmental pollutants [2–10]. Competitive immunoassays are either performed in the direct (antibody-coated) or the indirect (antigen-coated) format.

The most frequently employed enzymes with enzyme immunoassays (EIAs) are horseradish peroxidase (HRP) and alkaline phosphatase (AP). These labels belong to different enzyme classes, requiring the use of structurally different

Electronic supplementary material The online version of this article (doi:10.1007/s00216-012-6596-0) contains supplementary material, which is available to authorized users.

J. Grandke · L. Oberleitner · U. Resch-Genger ·
R. J. Schneider (✉)
BAM Federal Institute for Materials Research and Testing,
Richard-Willstätter-Str. 11,
12489 Berlin, Germany
e-mail: rudolf.schneider@bam.de

L.-A. Garbe
Technische Universität Berlin, Seestr. 13,
13353 Berlin, Germany

substrates. HRP is a donor hydrogen peroxide oxidoreductase (EC 1.11.1.7). Common EIA substrates include the chromogenic 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS), *ortho*-phenylenediamine, 3,3',5,5'-tetramethylbenzidine (TMB) as well as the fluorescent substrate 3-(4-hydroxyphenyl) propionic acid (HPPA) [11]. AP, a phosphoric monoester phosphohydrolase (EC 3.1.3.1), is employed, e.g., in conjunction with the photometric substrate *para*-nitrophenyl phosphate (*p*NPP) or the fluorescent substrate 4-methylumbelliferyl phosphate (MUP) [12].

Despite thousands of articles being published on EIAs, most reports compare one (favorable) format to another, but do not provide a methodical assessment of different assay formats. For example, Zhang et al. reported in 2006 and 2007 a direct and an indirect assay with a fluorescent label for the detection of the plasticizer dibutyl phthalate. The direct assay yielded a lower detection limit of 0.02 µg/L and a broader linear working range of 0.1–300 µg/L than the indirect assay (0.05 and 0.1–100 µg/L) [13, 14]. Lu et al. described direct and indirect EIAs for the detection of bisphenol A in canned food and beverages. The assay sensitivities and detection ranges were similar in both formats when the HRP substrate TMB was used, but the cross-reactivities were generally better for the indirect format [15]. Furthermore, the two competitive assay formats were compared in the selection of polyclonal or monoclonal antibodies for various analytes [2, 5, 16–25]. Cervino et al. found a better sensitivity for aflatoxin-specific antibodies in the direct assay using the HRP substrate TMB [18]. Manclus et al. detected a slightly lower or equivalent affinity to chlorpyrifos in the direct assay [5]. Deschamps et al. found that the indirect assay was superior to the direct format for picloram detection; however, the HRP substrate ABTS was used. Also, the antibody concentration was not the same for both assay formats with lower concentrations being used in the indirect format [25]. In the analysis of mussels for saxitoxin, the direct assay format was more sensitive and the coating more stable over a longer period of time [22].

In a pioneering work, different enzyme substrates were evaluated by Porstmann et al. The enzyme labels HRP and AP as well as β-galactosidase were coupled to antibodies of the immunoglobulin G type (IgG), and applied in sandwich immunoassays for alpha-1-fetoprotein (AFP) quantification. The enzyme activity decreased after the coupling procedure for all enzymes. Even so, the detection limits for AFP were lower for the fluorescent substrates compared to the chromogenic substrates, suggesting a more sensitive detection for the fluorogenic substrates [26]. The fluorescent HRP substrate HPPA was adapted for use in microtiter plates (MTPs) by Tuuminen et al. HPPA proved to be as sensitive as the chromogenic substrate TMB for protein detection, but the dynamic range was broader for HPPA than for TMB

[27]. Yolken et al. showed that the fluorescent AP substrate MUP enabled the detection of lower concentrations of polyribose phosphate compared to chromogenic *p*NPP within 10 min incubation time. If the incubation time was extended to 4 h, the same sensitivity was reached [28]. Mairal et al. developed various competitive immunoassays for gliadin using a commercial spectrofluorometer for, e.g., colorimetric detection of TMB (I) and fluorescence detection of HPPA (II) and MUP (III) in EIAs as well as direct detection of fluorescein isothiocyanate fluorescence (IV). The assay sensitivities, the limits of detection, width of the linear range and reproducibility decreased from I to IV [29].

Many authors compare newly developed immunoassay formats to previously described ones or develop direct and indirect assays at the same time. Yet, most articles in the field do not base their comparison of these formats on thoroughly optimized assays. Aiming to derive clearer criteria for the design of robust and sensitive immunoassays for small molecules, we performed a systematic study of individually optimized MTP-based formats for the low molecular weight compound caffeine using the same monoclonal antibody throughout. For this purpose, eight competitive immunoassays were compared employing two substrates, one chromogenic (TMB, *p*NPP) and one fluorogenic (HPPA, MUP), per enzyme label (HRP, AP) in the direct and indirect format, respectively. The following evaluation criteria were used to assess the quality of each assay format: sensitivity, measurement range, relative dynamic range of the signal, and the goodness of fit of the standard curve. For a more application-directed assessment, different beverages and cosmetics were analyzed with respect to intra- and inter-plate precision and correlation with the reference method Liquid chromatography tandem mass spectrometry (LC-MS/MS). The (potentially) caffeine-containing samples studied included five soft drinks, five energy drinks, six coffees (five samples from Arabica beans and one sample from Robusta beans), three teas, one cocoa sample, and four cosmetics (two shampoos and two roll-ons declared to contain caffeine).

Experimental

Reagents and materials

All solvents and chemicals were obtained from Merck KGaA (Darmstadt, Germany), Sigma-Aldrich (Taufkirchen, Germany), Mallinckrodt Baker (Griesheim, Germany) and Serva (Heidelberg, Germany) in the best available quality. The proteins HRP and AP were both EIA-grade and obtained from Roche (Mannheim, Germany). Ovalbumin (OVA) was purchased from Protea Biosciences (Morgantown, WV, USA). A Synthesis A10 Milli-Q® water purification system from Millipore (Schwalbach, Germany) was used to obtain ultrapure reagent water for the preparation of buffers and solutions.

All high-binding MTPs with 96 flat-bottomed wells were purchased from Greiner Bio-One (Frickenhausen, Germany). Black Fluotrac 600 plates and black μ Clear plates with clear bottoms were employed for fluorescence measurements whereas clear Microlon 600 plates were used for colorimetric assays. The caffeine reference standard was purchased from Sigma-Aldrich and used for the preparation of calibrators. The anti-mouse IgG whole molecule antibody (polyclonal, sheep, lot 21481) and the anti-mouse IgG whole molecule AP antibody (polyclonal, rabbit, lot 22315) were obtained from Acris Antibodies (Herford, Germany). The anti-caffeine antibody (monoclonal, mouse IgG2B, clone 1.BB.877, lot L2051502M) was purchased from US Biological (Swampscott, MA, USA). The anti-mouse IgG (γ -chain specific) HRP antibody (polyclonal, goat, lot 087K6014) was obtained from Sigma-Aldrich. The beverages, coffees, teas, and cosmetics were bought in a local supermarket.

Sample preparation

Soft drinks and energy drinks (for a list, see Electronic supplementary material, [ESM](#)) were degassed by shaking and subsequently using an ultrasonic bath for approximately 15 min. Masses of 7.000 ± 0.005 g ground coffee powder per sample (100 % Arabica coffee beans for the coffee types 1–5 and 100 % Robusta coffee beans for type 6) were brewed with 250 mL water in a filter coffee machine from Severin (Sundern, Germany). The same protocol was used for the preparation of the cocoa sample (cocoa powder, oil depleted, 7.000 g); 250 mL of boiling hot water were poured onto one bag of each tea (Ceylon-Assam black tea 1.75 g per bag, green tea 1.5 g per bag, and apple fruit tea 2.25 g per bag) and an infusion time of 10 min allowed. The cosmetics samples were prepared by dissolving 5.00 ± 0.05 g of product in 1 L ultrapure reagent water. Sample dilutions were chosen for the resulting sample concentrations in the range of 5 μ g/L for the LC-MS/MS measurements and near the immunoassay test midpoints (~ 0.15 μ g/L) (Table S1, [ESM](#)), respectively.

Methods

Preparation of the protein conjugates and coupling ratio analysis

The synthesis of the caffeine derivative (7-(5-carboxypentyl)-1,3-dimethylxanthine) and the HRP conjugate was performed as described by Carvalho et al. [4]. Analogously to the caffeine–HRP conjugate, the *N*-hydroxysuccinimide/*N*, *N'*-dicyclohexylcarbodiimide activated ester method was used to obtain a caffeine–AP conjugate for the direct assays and a caffeine–OVA conjugate for the indirect assays. Protein concentrations were determined by the Bradford method: 10 μ L diluted sample were added to 200 μ L Coomassie Plus

Protein Assay Reagent from Pierce (ThermoFisher Scientific, Rockford, IL, USA) and vortexed for 30 s. After a 30-min incubation period, the absorbance was measured at 595 nm in a microplate reader [30].

Mass spectra of the unconjugated proteins and the conjugates were obtained by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) on a Bruker Reflex III instrument (Bruker-Daltonik, Bremen, Germany) using a sinapic acid matrix containing 50 % acetonitrile and 0.1 % trifluoroacetic acid after passing through a Zeba™ Spin Micro Desalting Column. A Gaussian function was fitted to the mass signal distribution and the maxima of the fitted curves were used to calculate the mass difference of the protein and the protein conjugate to determine the mean coupling ratio.

Immunoassays

General procedures and optimization

The immunoassays were performed in the direct and indirect competitive format using either the enzyme HRP or AP as label and the chromogenic substrates TMB and *p*NPP as well as the fluorogenic substrates HPPA and MUP, respectively. The immunoassays described below were thoroughly optimized. Only the methods yielding the best results regarding signal intensity and sensitivity over a wide measurement range are reported for comparison. The concentration of the monoclonal anti-caffeine antibody was kept constant after checkerboard titrations showed that only the concentration of the caffeine–enzyme conjugate in the direct and caffeine–OVA conjugate in conjunction with the labeled anti-mouse IgG antibody in the indirect assay influenced the sensitivity significantly. Other parameters exploited for optimization included buffer composition and pH as well as the concentrations and incubation times of the substrates and antibodies used. All incubation steps were performed at room temperature [31].

For all pipetting steps, a 96-channel pipette Liquidator⁹⁶ from Steinbrenner Laborsysteme (Wiesenbach, Germany) with matching tips from Mettler-Toledo (Giessen, Germany) was used. Between individual incubation steps, the plates were washed with an automatic 96-channel plate washer (BioTek Instruments, ELx405 Select™, Bad Friedrichshall, Germany). Three-cycle washing steps of all HRP immunoassays were carried out with a PBS-based washing buffer (0.75 mM potassium dihydrogen phosphate, 6.25 mM dipotassium hydrogen phosphate, 0.025 mM sorbic acid potassium salt, 0.05 % (v/v) Tween™ 20, pH 7.6) whereas a diethanolamine (DEA)-based washing buffer (100 mM DEA, 0.05 % (v/v) Tween™ 20, pH 9.8) was used for all AP assays.

MTPs were covered with Parafilm® M and shaken at 750 rpm on Titramax 101 plate shakers from Heidolph (Schwabach, Germany). All signals (absorbance, fluorescence) were acquired using the SpectraMax M5 multi-mode reader from Molecular Devices (Biberach an der Riss, Germany) at ambient temperature. The instrument was controlled by the Softmax® Pro software version 5.4. All measurements were performed in the top reading mode with adapter; however, when black MTPs with clear bottoms were used an additional measurement in the bottom reading mode was performed without adapter.

Direct competitive format

All 96 wells were coated with 200 µL of 1 mg/L anti-mouse IgG in PBS (10 mM sodium dihydrogen phosphate, 70 mM disodium hydrogen phosphate, 145 mM sodium chloride, pH 7.6) and incubated for 18 h on a plate shaker. After a three-cycle washing step, 200 µL anti-caffeine antibody in TRIS buffer (13.7 µg/L; 10 mM tris(hydroxymethyl)aminomethane (TRIS), 150 mM sodium chloride, pH 8.5) were added and incubated for 1.5 h. Following another washing step, 150 µL of the calibrators in the range of 0 to 33 µg/L were added to each well. A ~1 g/L methanolic stock solution of caffeine was prepared gravimetrically and calibrators were obtained by sequential dilution with ultrapure water. A randomized application pattern with 16 calibrators ($N=6$) was used to obtain a standard curve with precision profile [31]. For the sample determination, 8 calibrators and 24 samples were measured ($N=3$) with identical distribution twice on two separate plates. 8 min later, 50 µL of caffeine enzyme conjugate was added. The caffeine–HRP conjugate (4 µg/L) was diluted in sample buffer (1 M glycine, 3 M sodium chloride, 2 % (w/v) ethylenediaminetetraacetic acid disodium salt, pH 9.5). The caffeine–AP conjugate was diluted in TRIS buffer. Here, the concentration varied depending on the substrate used: 15 µg/L for *p*NPP and 3.2 µg/L for MUP. After a 30-min incubation period and another washing step, 200 µL of the respective substrate solution were added.

1. A protocol according to Frey et al. was used for the preparation of the HRP substrate TMB [32]. For one plate, 21 mL citrate buffer (220 mM potassium dihydrogen citrate, 0.5 mM sorbic acid potassium salt, pH 4.0) with 8.1 µL H₂O₂ (30 %) and 525 µL TMB solution (40 mM TMB, 8 mM tetrabutylammonium borohydride, in *N,N'*-dimethylacetamide) were mixed and 200 µL added to each well. Following a 30-min incubation step, the reaction was stopped by adding 100 µL 1 M H₂SO₄. Absorbance was measured at 450 nm and referenced to 620 nm.

2. The fluorogenic HRP substrate HPPA was prepared similar to Tuuminen et al. [27]: 21 mL of a 3.5 g/L HPPA solution in 0.1 M TRIS buffer (pH 8.5, stable for at least 2 months at 4 °C) with 4.29 µL H₂O₂ (30 %, 2 mM) were freshly prepared. Two hundred microliters substrate solution were added to each well and after 45 min the reaction was stopped adding 100 µL of 0.2 M glycine/sodium hydroxide solution (pH 10.3). Fluorescence was excited at 320 nm (325 nm cutoff filter) and detected at 415 nm [33].
3. The chromogenic AP substrate *p*NPP was prepared as a 1 mg/mL *p*NPP solution in DEA substrate buffer (1 M DEA, 0.1 mM magnesium chloride hexahydrate, pH 9.8). Two hundred microliters were added to each well and the reaction was stopped after 45 min by adding 100 µL 2 M sodium carbonate. Absorbance was measured at 400 nm using 620 nm as reference [26].
4. The AP substrate MUP was used for quantification of relative fluorescence intensities. A 1 mg/mL MUP solution in DEA substrate buffer was prepared and 200 µL were pipetted in each well. The reaction was stopped after 45 min by adding 100 µL 2 M sodium carbonate. The fluorescence was excited at 360 nm (420 nm cutoff filter) and detected at 450 nm [29].

Indirect competitive format

Each well was coated with 200 µL of caffeine–OVA conjugate in PBS and incubated for 18 h on a plate shaker. The concentrations depended on the substrate used: 3.1 µg/L for TMB, 2.5 µg/L for HPPA, 1.8 µg/L for *p*NPP, and 0.8 µg/L for MUP. Following a three-cycle washing step, the remaining binding sites on the MTP were blocked with 200 µL casein solution (0.1 %) in PBS buffer. After incubating for 1 h, the plates were washed again and 150 µL calibrator solution (randomized, 0 to 33 µg/L) or sample were added. 50 µL anti-caffeine antibody (54.8 µg/L) in TRIS buffer were added 8 min later. The concentration of the anti-caffeine antibody in each well in the indirect assay equals the concentration in the direct format as the solution was further diluted by a factor of 4 by the samples or calibrators.

After a 30-min incubation period and another washing step, enzyme-labeled anti-mouse IgG was added and incubated for 1 h. Two hundred microliters of 100 µg/L anti-mouse IgG HRP conjugate diluted in PBS buffer were added for both HRP substrates. Two hundred microliters of 27.5 µg/L anti-mouse IgG-AP conjugate for *p*NPP or 13.8 µg/L anti-mouse IgG-AP conjugate for MUP diluted in TRIS buffer were added for the AP substrates. Following another washing step, the addition of the substrate solution, the stopping solution and the detection was completed as described above.

Immunoassay quality assessment

Calibrators (and samples) were assayed twice (2 MTPs) in triplicate and subjected to a Grubbs outlier test ($\alpha=0.01$). A four-parameter logistic function (4PL) was fitted to the mean values of the calibrators using the Origin 8G software (OriginLab, Northampton, USA) [34]. Parameter B was set to 1 as previously described [31]. The resulting sigmoidal standard curve is characterized by the highest signal A at minimal caffeine concentration, the background signal D at highest caffeine concentration, and the test midpoint C (or point of inflection, $\approx IC_{50}$) as a measure for assay sensitivity. Additionally, the coefficient of determination R^2 for the curve fitting was obtained as a measure for goodness of fit. The signal dynamic range (DR) of each assay was calculated by subtracting the background signal D from the maximum signal A. A normalized, relative dynamic range (RDR) was calculated by dividing the DR by the signal intensity A. Standard deviations of the mean signals were used to obtain the precision profile according to Ekins by calculating the relative error to each calibrator caffeine concentration [35]. The range with a relative error of the concentration below 30 % was assigned the measurement range of the respective assay. Intra- and inter-plate precision data for the concentration measurements was determined for a series of caffeine samples of different types. For these experiments, eight samples were determined in sextuplicate per plate and on four different plates (6×4). Intra-plate variation is reported as the range between the lowest and the highest coefficient of variation (CVs, in %) for the sample groups. Inter-plate variation was derived as the range of the CV values of the individual replicates ($N=6$, the same well on all four plates).

LC-MS/MS method

The measurements were performed with an Agilent 1100 Liquid Chromatography system (Agilent Technologies, Waldbronn, Germany) including a degasser, binary pump, auto sampler, and a column heater. The instrument was coupled to an API 4000 mass spectrometer from Applied Biosystems (Darmstadt, Germany) and controlled by the Analyst software 1.4.1 for data acquisition and analysis. The Turbo VTM electrospray ion source was operated in the positive mode.

A reversed-phase C18 Ultrasep ES Phen column (250×3 mm, $5 \mu\text{m}$, SepServ, Berlin, Germany) with a guard column (10×3 mm) was used for chromatographic separation. A binary gradient consisting of 10 mM ammonium acetate and 0.1 % acetic acid in water (A) and methanol (B) was used: starting with 80 % A, isocratic for 3 min, linear decrease to 5 % A within 20 min, kept at 5 % A for 10 min. The flow rate was constant at 0.5 mL/min, the column oven

was set to 40 °C and 50 μL sample were injected. The acquisition was done in the Multiple Reaction Monitoring (MRM) mode. The first transition MRM1, m/z 195 \rightarrow 138, was used for the quantification of the peak area in duplicates and the second one, MRM2, m/z 195 \rightarrow 110, for confirmation.

Results and discussion

Protein conjugate characterization

High-quality protein conjugates are necessary to develop sensitive immunoassays. In order to characterize the conjugates prepared with the proteins OVA, HRP, and AP, coupling ratios and conjugate concentrations were determined. The masses, obtained by MALDI-TOF-MS, of the HRP conjugate and the unmodified HRP were 44,903 and 44,124 Da, respectively. This leads to a mass difference of 779 Da. The mass of the caffeine derivative minus water equals 276 Da, so an average of 2.8 molecules of caffeine derivative had coupled per HRP molecule. The protein concentration of the HRP conjugate was determined to 2.0 mg/mL. For the caffeine-AP conjugate, masses of 58,134 Da for the pure AP and 61,593 Da for the conjugate were determined; the caffeine-AP conjugate had a mean number of 12.5 molecules caffeine per AP molecule and a protein concentration of 0.45 mg/mL. Masses of 48,651 and 44,368 Da were assigned to the caffeine-OVA conjugate and the pure protein, corresponding to a mean of 15.5 molecules of caffeine derivative per OVA molecule (2.5 mg/mL). Data shows that conjugates of good quality had been obtained.

Immunoassay quality assessment and comparison: calibration curves

Sensitivity

The test midpoints C (Table 1) were used to compare assay sensitivities. The midpoints did not vary significantly between the direct and indirect assay set-ups of the individual substrates; however, it is noteworthy that the test midpoints were systematically lower in the direct format for the HRP label whereas the opposite was observed for the AP label. Use of the HRP label yielded better assay sensitivity compared to AP.

Measurement range

For instrumental methods, the limit of detection (LOD) is the lowest concentration that can be distinguished from a blank value within an established confidence limit, estimated from

Table 1 Characteristic parameters of the standard curves: the test midpoint *C* [concentration, in micrograms per liter], measurement range [concentration, in micrograms per liter], the signal dynamicrange (*A*–*D*), the relative dynamic range ($((A-D)/A)$), and the coefficient of determination R^2 as a measure for goodness of fit (parameters *A*, *D*, and *C* are results from 4PL fitting)

Assay and detection	<i>C</i> [μg/L]	Measurement range [μg/L]	Dynamic range (DR)	Relative DR (RDR)	R^2
Direct HRP TMB (Abs.)	0.095	0.033–33 cont ^a	0.76	1.00	0.999
Indirect HRP TMB (Abs.)	0.184	0.083–11.750	0.66	0.98	0.998
Direct HRP HPPA (Flu.)	0.075	0.105–0.623	2090	0.81	0.997
Indirect HRP HPPA (Flu.)	0.219	0.106–2.197	2260	0.90	0.998
Direct AP <i>p</i> NPP (Abs.)	0.817	0.128–33 cont ^a	0.60	0.87	1.000
Indirect AP <i>p</i> NPP (Abs.)	0.555	0.233–7.193	1.37	0.86	0.996
Direct AP MUP (Flu.)	0.890	0.191–33 cont ^a	69300	0.92	1.000
Indirect AP MUP (Flu.)	0.668	none	93100	0.72	0.998

Abs. absorbance, *Flu.* fluorescence

^a The measurement range extends beyond the concentration of the calibrators

the mean of the blank and $\sigma=3$ times its standard deviation (the limit of quantitation, LOQ, is obtained for $\sigma=10$). In contrast to linear standard curves for instrumental methods, sigmoidal standard curves are commonly used for immunoassay evaluation. It is under discussion whether the definitions of the LOD and LOQ can be transferred to immunoassays [36, 37]. As an alternative, the precision profile (a response error relationship) can be used to define a working range with a certain confidence limit as suggested by Ekins [35]. In allusion to the “three-sigma-criterion” described above for the LOD, the relative error of the concentration threshold was set to 30 %.

Generally, a wide measurement range is desired. As specific requirement for this criterion, the range should comprise three orders of magnitude. The direct HRP TMB assay and potentially the direct assays for AP *p*NPP as well as AP MUP showed a measurement range of three orders of magnitude. The widest measurement range was found for the direct HRP TMB assay followed by the direct AP *p*NPP assay and the direct format for AP MUP. A narrower measurement range was obtained for the indirect HRP TMB assay. The smallest measurement ranges were found for both HRP HPPA assays. For the indirect AP MUP assay, no measurement range could be determined within the 30 % threshold.

Dynamic range and relative dynamic range

The DRs of the signals were usually larger for the indirect assay. The values of the DR depend on the detection method used: usually a value of up to 1.5 is realistic for absorbance measurements whereas for fluorescence measurements the DR is arbitrary. However, if the background signal *A* (4PL) is taken into account and the DR normalized, this unified scale can then be used to compare the assays performed with colorimetric and fluorometric substrates. For the RDR, a value of at least 0.90 is desired. This condition was met by four

assay formats in our study: the direct and indirect HRP TMB, the indirect HRP HPPA, and the direct AP MUP assay. The direct assay is superior to the indirect assay in three out of four tested formats with the exception of the HRP HPPA assay. The HRP TMB assays achieved the best RDR values with 0.98 for the indirect format and 1.0 for the direct format.

Goodness of fit

The goodness of fit is a measure of how well the fitting function adjusts to the measured data points. As a measure the coefficient of determination R^2 of the 4PL was chosen. It should be very close to 1 but at least 0.990. The values for R^2 of all assays exceeded this requirement. When comparing both assay formats for the different substrates, the R^2 values were better for the direct assays; this supports the superiority of the direct format. Although the goodness of fit is a helpful indicator for the quality of the standard curve, the standard deviations of the mean for each calibrator should be considered, too.

The standard curves for the optimized direct and indirect HRP TMB assay (Fig. 1) show that the curve characteristics are very similar for both assays with the maximum absorbance being around 0.7. This leads to comparable test midpoints (0.095 and 0.184 μg/L). However, the standard deviations of the means were higher for the indirect assay. As a consequence, a reduced measurement range was obtained for the indirect assay. Taking all these criteria into account, the direct format proves superior to the indirect format for the caffeine immunoassay.

Application-focused quality assessment

Precision and uncertainties

In an application-oriented assessment of the different immunoassays, 24 different beverages and cosmetics were studied.

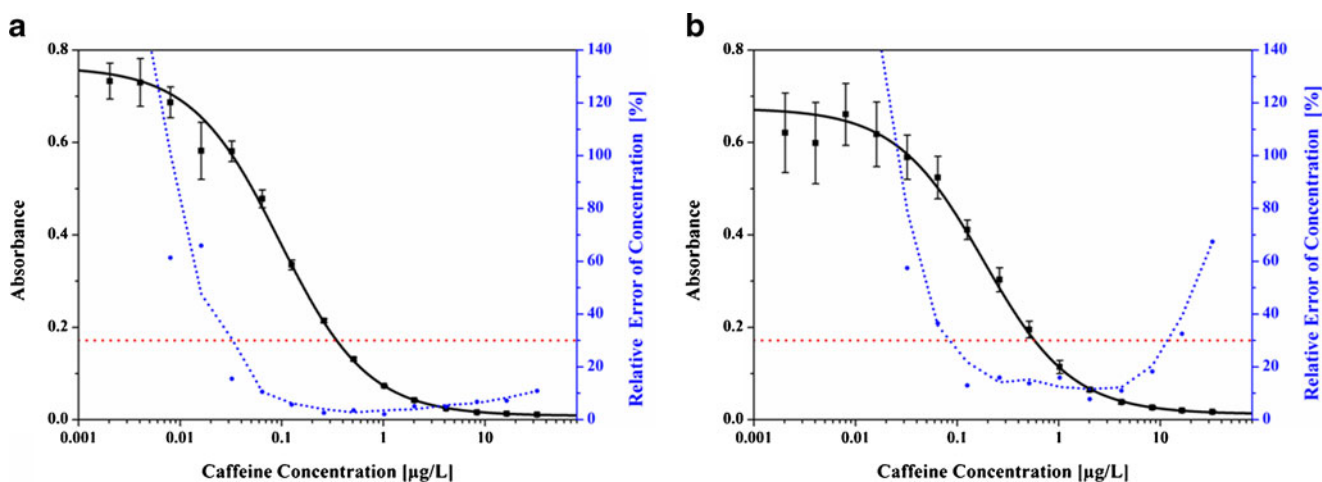


Fig. 1 Standard curves of the direct (a) and indirect (b) caffeine immunoassay using the HRP substrate TMB including the precision profiles ($N=6$, moving average of two adjacent points is shown as dotted blue line)

Here, calibration curves with 8 caffeine calibrators per assay were run on the same MTP as the samples. The distinctive parameters of these curves were reviewed and proved equivalent to the standard curves obtained for 16 calibrators. The results for the chromogenic HRP substrate TMB revealed that the measurement uncertainty is higher in the indirect assay compared to the direct assay, because the CVs [%] of the concentrations were higher for the majority of the samples (Fig. 2; caffeine concentrations and standard deviations are available in the ESM, Table S2).

The results for the HRP HPPA, AP *p*NPP, and AP MUP assays used for the caffeine measurement in beverages and cosmetics were similar: the indirect assay lags behind the direct assay with respect to the dynamic range. A possible explanation for the smaller RDR in the direct HRP assays

for HPPA could be that the standard curve for the direct assay is shifted towards higher RFUs; therefore, a higher background fluorescence reduces the RDR. The CVs for the caffeine samples were smaller for the direct HRP HPPA assay. When focusing on the AP assays for *p*NPP and MUP, the direct format was advantageous again and yielded smaller standard deviations and hence an extended measurement range (Fig. 3a, b). The larger measurement uncertainty obtained for the indirect assay may be attributed to the additional washing and pipetting step. In all cases, the direct format was superior to the indirect set-up for the analyte caffeine. The CVs were smaller for almost all caffeine-containing samples, the results showed higher precision which leads to smaller overall uncertainties.

The intra- and inter-plate CVs [%] of the direct formats were determined for different sample groups (soft drinks, energy drinks, coffees, teas, and cosmetics; Table 2). For the intra-plate variation, the lowest CV values were found for soft drinks, energy drinks, coffees, and teas using the HRP HPPA assay whereas for cosmetics the HRP TMB assay yielded the best results. Regarding inter-plate variation, the HRP HPPA assay gave the smallest values for soft drinks, energy drinks, and teas while the HRP TMB assay was best for coffees and cosmetics.

The intra- and inter-plate CVs [%] for each assay should not exceed 10 and 20 %, respectively. The intra-plate CVs for the HRP assays were similar in the range of 1.0–9.9 % for TMB and 1.8–9.4 % for HPPA (Table 3). The CV values for the AP assays were significantly higher: 6.7–41 % for *p*NPP and 6.2–52 % for MUP. The inter-plate variations were 0.9–18 % for HRP TMB, 0.4–19 % for HRP HPPA, 1.5–29 % for AP *p*NPP and 5.3–50 % for AP MUP. The intra- and inter-plate CVs of the HRP immunoassays are smaller than 10 and 20 %, respectively and therefore fulfill these requirements.

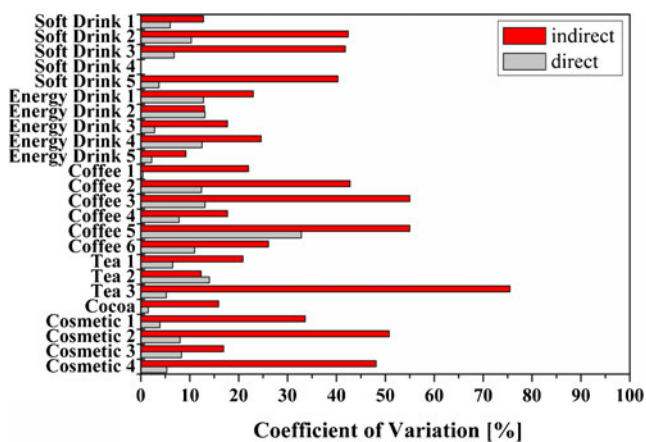


Fig. 2 The direct and indirect caffeine immunoassays were performed using the HRP substrate TMB. The coefficients of variation [%] of the caffeine concentrations for all samples are given for both assay formats (caffeine concentrations and standard deviations are provided in the ESM)

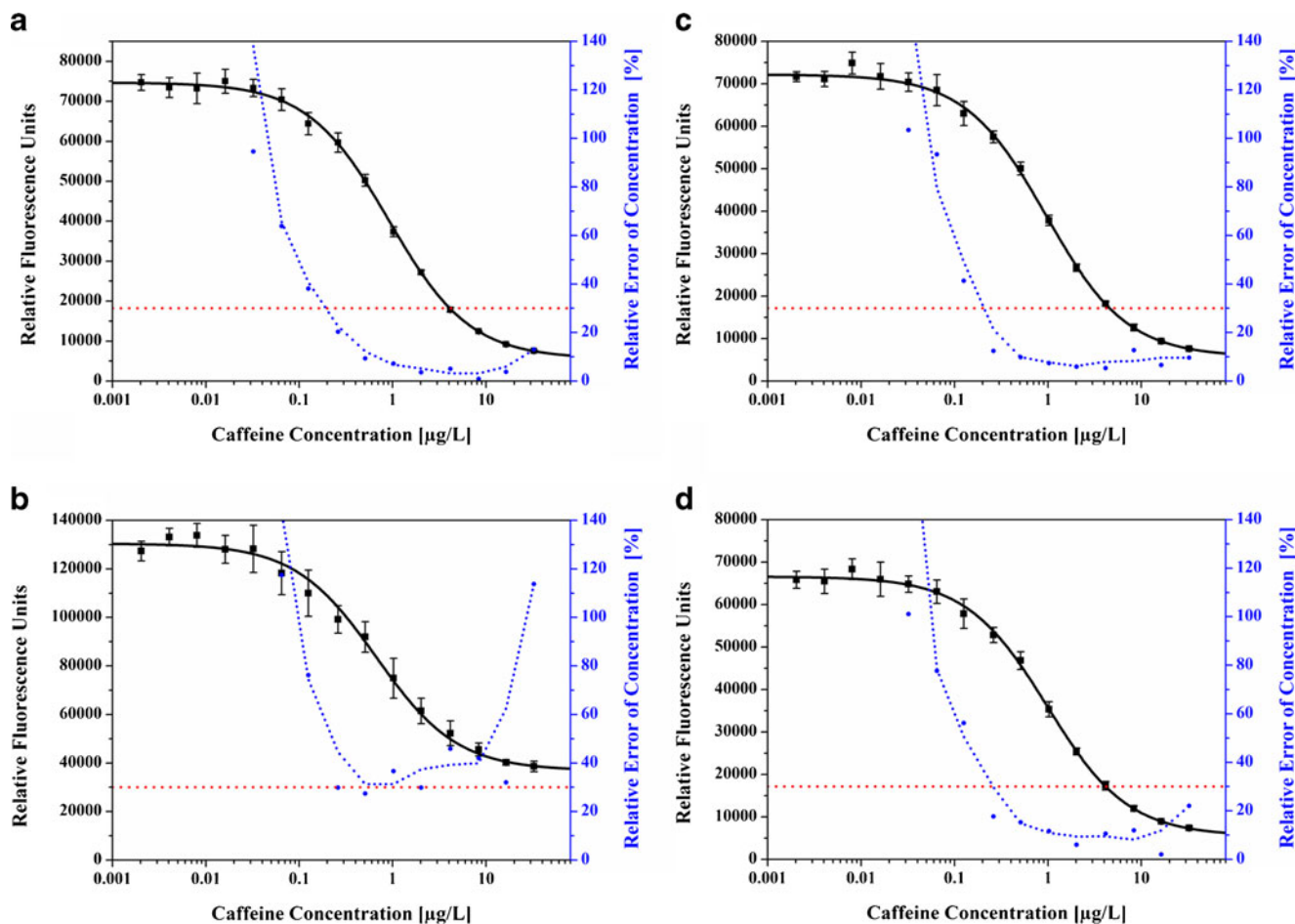


Fig. 3 Standard curves of the direct (a, c, d) and indirect (b) caffeine immunoassay using the AP substrate MUP including the precision profiles ($N=6$, moving average of two adjacent points is shown as

dotted blue line). The assay was performed in black MTPs (a, b) or black MTPs with transparent bottom (c, d) and fluorescence detected from the top (a–c) or bottom (d)

Both AP immunoassays exceed the required values for the intra- as well as the inter-plate repeatability.

Accuracy

In the assessment of accuracy of the immunoassay formats, the caffeine content of soft drinks, cosmetics, coffees, and energy drinks determined by immunoassay was compared to

a reference method. For immunoassay analysis, the direct formats with the substrates TMB, HPPA, *p*NPP, and MUP that provided optimum assay performance were chosen. Reference analysis was performed by LC-MS/MS. Accuracy assessment was obtained via the parameters of the linear regression analysis of the results, immunoassay results being the dependent variable: the slope m , the intercept with the y -axis n , and the coefficient of determination R^2 (squared correlation

Table 2 Intra- and inter-plate coefficients of variation (CVs, %) are given for the four direct formats on the basis of the different caffeine sample groups

Direct assay	HRP TMB		HRP HPPA		AP <i>p</i> NPP		AP MUP	
	Intra-plate	Inter-plate	Intra-plate	Inter-plate	Intra-plate	Inter-plate	Intra-plate	Inter-plate
Soft drinks	1.0–6.2	0.9–7.2	1.8–4.3	0.4–6.9	6.7–20	1.5–18	9.6–28	8.9–25
Energy drinks	1.6–6.3	5.2–12	2.4–5.2	1.3–5.1	16–26	3.7–22	17–52	7.4–40
Coffees	2.7–9.6	1.2–10	2.4–9.4	0.9–19	13–41	3.9–29	14–35	6.3–32
Teas	3.9–9.9	4.6–18	2.2–6.7	0.8–9.1	6.7–20	9.1–20	7.4–31	7.5–50
Cosmetics	1.0–6.9	5.5–7.0	2.1–8.4	1.0–9.8	15–22	3.7–15	6.2–27	5.3–22

Table 3 Intra- and inter-plate coefficients of variation (CVs, %) are given for the four formats. Parameters of the linear regression of immunoassay results and results from LC-MS/MS measurements ($c(\text{Caf})_{\text{assay}} = m \times c(\text{Caf})_{\text{LC-MS/MS}} + n$)

Immunoassay format	Intra-plate CV [%]	Inter-plate CV [%]	Slope m	Intercept n	R^2
HRP TMB	1.0–9.9	0.9–18	0.96	−0.41	0.968
HRP HPPA	1.8–9.4	0.4–19	0.83	0.11	0.966
AP <i>p</i> NPP	6.7–41	1.5–29	1.05	−15	0.914
AP MUP	6.2–52	5.3–50	0.99	−11	0.971

coefficient R) (Fig. 4, Table 3). As a quality criterion, the slope should be 1.00 ± 0.05 and the intercept with the y -axis should be close to 0. The former was the case for all assays except for the HRP HPPA assay. In contrast to the HRP assays, the AP assays showed larger deviations from the coordinate origin. The correlations between the immunoassay results and LC-MS/MS measurements were good, because R^2 was greater than 0.95 for all formats except the AP *p*NPP assay.

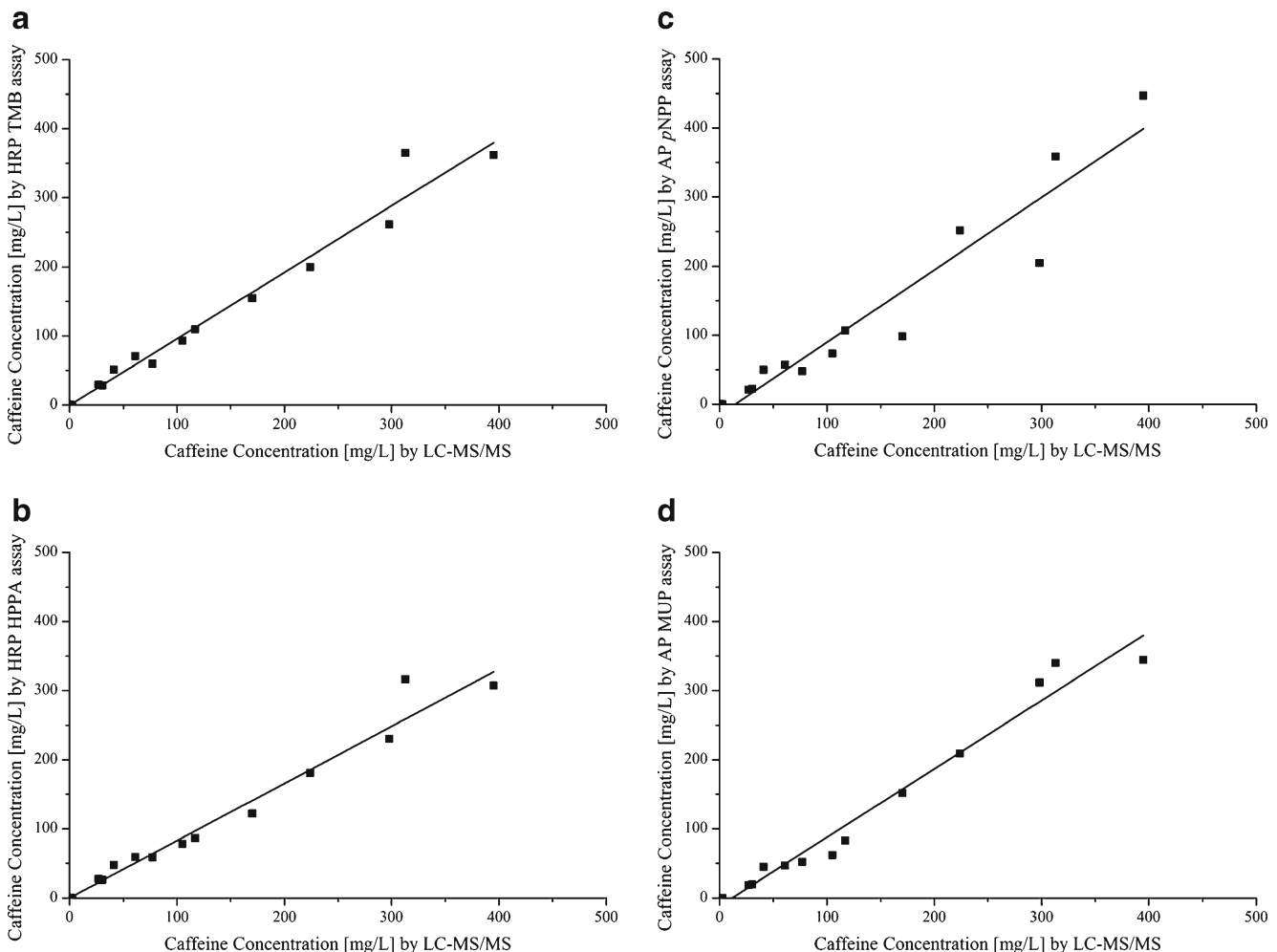


Fig. 4 Correlation between the caffeine concentrations of consumer products measured by four direct immunoassay formats (a HRP TMB, b HRP HPPA, c AP *p*NPP, d AP MUP) and LC-MS/MS, respectively

Additional considerations

For specific applications, other properties, than the ones discussed here, might be taken into consideration for the choice of format.

An important benefit of the direct assays are the shorter incubation times, saving valuable time for possible high-throughput screenings.

A possible advantage of the indirect assay can be the fact that enzyme and sample are not directly in contact with each other, so no enzyme inactivation by matrix compounds can occur. Matrix validation is hence crucial for generating a robust application.

As an alternative for peroxidase, which as a common enzyme in plants cannot be used for some applications, AP can be employed as enzyme label. In this case, both substrates can be employed depending on the sample group, e.g., for energy drinks the AP *p*NPP assay showed lower intra- and inter-plate variations compared to the AP MUP assay.

Fluorescence can be detected in the top as well as in the bottom read-out mode with the multi-mode reader Spectra-Max M5. Detection from the bottom requires black plates with clear-bottomed wells, whereas detection from the top requires a plate adapter. Detection from the top can be realized in plain black plates or black plates with clear bottoms. The AP MUP assay was performed in both types of MTPs and measured in top and bottom read-out mode (Fig. 3c and d). The signal intensities, RDR and CVs of the samples were slightly smaller in top compared to bottom detection. These findings were confirmed with the second fluorescent substrate HRP HPPA. All in all, no significant differences were detected. However, the costs for plain black plates are only half of the clear-bottomed plates. Moreover, most plate readers are equipped for top rather than bottom detection. Thus, we recommend to measure fluorescence in the top reading mode.

Conclusion

Seven parameters were defined for the quality assessment of caffeine enzyme immunoassays using different formats, enzymes, substrates, and detection methods, yet the same monoclonal antibody: four parameters for the standard curves and another three parameters for the application of these assays for the caffeine quantification in consumer products like beverages and cosmetics. These criteria should not be considered as stand-alone evaluation tools, but used in combination considering all criteria to assess fitness for purpose.

When comparing the different caffeine immunoassays, the direct assay formats led to wider measurement ranges as well as to larger relative dynamic ranges of the curves and better goodness of fit. For these four criteria defined for the standard curves, the direct formats are superior to the indirect assays, independent of the enzyme label and the substrate used. Additionally, the direct format saves time. Therefore, the direct format should be used, provided the enzyme conjugate is available.

In regard to the test midpoints, the HRP assays are more sensitive than the AP assays. The HRP substrate TMB is already widely used in laboratory routine. The direct format for this substrate fulfills all requirements defined by us for the calibration. Additionally, it yields one of the lowest test midpoints. Considering these aspects, the direct HRP TMB immunoassay is the best format of the immunoassay formats studied here. These findings are supported by the measurements of the caffeine concentration of different consumer products. The criteria for this assessment were intra- and inter-plate precision and accuracy derived by the linear regression of concentrations with results by a reference method, LC-MS/MS. Here, the direct HRP TMB assay reveals the best performance, complying with all requirements.

The transferability of the findings and the parameters defined and verified for the caffeine immunoassays to other analytes needs to be confirmed. This is the focus of ongoing research.

Acknowledgments J. Grandke thanks BAM Federal Institute for Materials Research and Testing for a grant within its PhD program. We greatly acknowledge Dr. A. Lehmann for LC-MS/MS measurements, A. Grasnack for technical assistance during the immunoassay performance, and S. Flemig for MALDI-TOF-MS measurements (all BAM).

References

1. Schuurs AH, van Weemen BK (1980) Enzyme-immunoassay: a powerful analytical tool. *J Immunoassay* 1(2):229–249. doi:10.1080/01971528008055786
2. Alarcon SH, Palleschi G, Compagnone D, Pascale M, Visconti A, Barna-Vetro I (2006) Monoclonal antibody based electrochemical immunosensor for the determination of ochratoxin A in wheat. *Talanta* 69(4):1031–1037. doi:10.1016/j.talanta.2005.12.024
3. Bahlmann A, Weller MG, Panne U, Schneider RJ (2009) Monitoring carbamazepine in surface and wastewaters by an immunoassay based on a monoclonal antibody. *Anal Bioanal Chem* 395(6):1809–1820. doi:10.1007/s00216-009-2958-7
4. Carvalho JJ, Weller MG, Panne U, Schneider RJ (2010) A highly sensitive caffeine immunoassay based on a monoclonal antibody. *Anal Bioanal Chem* 396(7):2617–2628. doi:10.1007/s00216-010-3506-1
5. Manclus JJ, Primo J, Montoya A (1996) Development of enzyme-linked immunosorbent assays for the insecticide chlorpyrifos. 1. Monoclonal antibody production and immunoassay design. *J Agric Food Chem* 44(12):4052–4062. doi:10.1021/jf960144q
6. Mire-Sluis AR, Barrett YC, Devanarayan V, Koren E, Liu H, Maia M, Parish T, Scott G, Shankar G, Shores E, Swanson SJ, Taniguchi G, Wierda D, Zuckerman LA (2004) Recommendations for the design and optimization of immunoassays used in the detection of host antibodies against biotechnology products. *J Immunol Methods* 289(1–2):1–16. doi:10.1016/j.jim.2004.06.002
7. Pastor-Navarro N, Brun EM, Gallego-Iglesias E, Maquieira A, Puchades R (2009) Development of immunoassays to determine sulfamethoxazole residues in wastewaters. *J Environ Monit* 11(5):1094–1100. doi:10.1039/b818299j
8. Schneider C, Schöler HF, Schneider RJ (2004) A novel enzyme-linked immunosorbent assay for ethynylestradiol using a long-chain biotinylated EE2 derivative. *Steroids* 69(4):245–253. doi:10.1016/j.steroids.2004.01.003
9. Schneider RJ (2003) Environmental immunoassays. *Anal Bioanal Chem* 375(1):44–46. doi:10.1007/s00216-002-1659-2
10. Yu FY, Chi TF, Liu BH, Su CC (2005) Development of a sensitive enzyme-linked immunosorbent assay for the determination of ochratoxin A. *J Agric Food Chem* 53(17):6947–6953. doi:10.1021/jf051392z
11. Schomberg D, Schomberg I (eds) (2006) Class 1, Oxidoreductases X, EC 1.9–1.13, vol 25. Springer handbook of enzymes, 2nd edn. Springer, Berlin
12. Schomberg D, Schomberg I (eds) (2003) Class 3.1, Hydrolases V, EC 3.1.3, vol 10. Springer handbook of enzymes, 2nd edn. Springer, Berlin
13. Zhang MC, Wang QE, Zhuang HS (2006) A novel competitive fluorescence immunoassay for the determination of dibutyl phthalate. *Anal Bioanal Chem* 386(5):1401–1406. doi:10.1007/s00216-006-0703-z

14. Zhang MC, Wang QE, Zhuang HS (2007) Determination of dibutyl *o*-phthalate by antigen-coated competitive fluorescence immunoassay. *Anal Lett* 40(1):127–137. doi:10.1080/00032710600952564
15. Lu Y, Peterson JR, Gooding JJ, Lee NA (2012) Development of sensitive direct and indirect enzyme-linked immunosorbent assays (ELISAs) for monitoring bisphenol-A in canned foods and beverages. *Anal Bioanal Chem* 403(6):1607–1618
16. Abad A, Moreno MJ, Montoya A (1999) Development of monoclonal antibody-based immunoassays to the *N*-methylcarbamate pesticide carbofuran. *J Agric Food Chem* 47(6):2475–2485. doi:10.1021/jf981184s
17. Bekheit HKM, Lucas AD, Szurdoki F, Gee SJ, Hammock BD (1993) An enzyme-immunoassay for the environmental monitoring of the herbicide bromacil. *J Agric Food Chem* 41(11):2220–2227. doi:10.1021/jf00035a075
18. Cervino C, Weber E, Knopp D, Niessner R (2008) Comparison of hybridoma screening methods for the efficient detection of high-affinity hapten-specific monoclonal antibodies. *J Immunol Methods* 329(1–2):184–193. doi:10.1016/j.jim.2007.10.010
19. Esteve-Turrillas FA, Parra J, Abad-Fuentes A, Agullo C, Abad-Somovilla A, Mercader JV (2010) Hapten synthesis, monoclonal antibody generation, and development of competitive immunoassays for the analysis of picoxystrobin in beer. *Anal Chim Acta* 682(1–2):93–103. doi:10.1016/j.aca.2010.09.042
20. Li K, Chen RL, Zhao BT, Liu M, Karu AE, Roberts VA, Li QX (1999) Monoclonal antibody-based ELISAs for part-per-billion determination of polycyclic aromatic hydrocarbons: Effects of haptens and formats on sensitivity and specificity. *Anal Chem* 71(2):302–309. doi:10.1021/ac980765d
21. Maragos C, Busman M, Sugita-Konishi Y (2006) Production and characterization of a monoclonal antibody that cross-reacts with the mycotoxins nivalenol and 4-deoxynivalenol. *Food Addit Contam* 23(8):816–825. doi:10.1080/2652030600699072
22. Micheli L, Di Stefano S, Moscone D, Palleschi G, Marini S, Coletta M, Draisci R, delli Quadri F (2002) Production of antibodies and development of highly sensitive formats of enzyme immunoassay for saxitoxin analysis. *Anal Bioanal Chem* 373(8):678–684. doi:10.1007/s00216-002-1399-3
23. Radoi A, Targa M, Prieto-Simon B, Marty JL (2008) Enzyme-Linked Immunosorbent Assay (ELISA) based on superparamagnetic nanoparticles for aflatoxin M-1 detection. *Talanta* 77(1):138–143. doi:10.1016/j.talanta.2008.05.048
24. Usleber E, Lorber M, Straka M, Terplan G, Märtlbauer E (1994) Enzyme-immunoassay for the detection of isoxazolyl penicillin antibiotics in milk. *Analyst* 119(12):2765–2768. doi:10.1039/an9941902765
25. Deschamps RJA, Hall JC (1991) Enzyme immunoassays for picloram detection: a comparison of assay formats. *Food Agric Immunol* 3(3–4):135–145
26. Porstmann B, Porstmann T, Nugel E, Evers U (1985) Which of the commonly used marker enzymes gives the best results in colorimetric and fluorimetric enzyme immunoassays—horseradish peroxidase, alkaline phosphatase or beta-galactosidase? *J Immunol Methods* 79(1):27–37
27. Tuuminen T, Palomäki P, Rakkolainen A, Welin MG, Weber T, Käpyaho K (1991) 3-*para*-hydroxyphenylpropionic acid—a sensitive fluorogenic substrate for automated fluorometric enzyme immunoassays. *J Immunoassay Immunochem* 12(1):29–46
28. Yolken RH, Leister FJ (1982) Comparison of fluorescent and colorimetric substrates for enzyme immunoassays. *J Clin Microbiol* 15(5):757–760
29. Mairal T, Frese I, Llaudet E, Redondo CB, Katakis I, von Germar F, Drese K, O’Sullivan CK (2009) Microfluorimeter with disposable polymer chip for detection of coeliac disease toxic gliadin. *Lab Chip* 9(24):3535–3542. doi:10.1039/b914635k
30. Bradford MM (1976) A rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *Anal Biochem* 72(1–2):248–254. doi:10.1006/abio.1976.9999
31. Grandke J, Resch-Genger U, Bremser W, Garbe LA, Schneider RJ (2012) Quality assurance in immunoassay performance—temperature effects. *Anal Methods* 4(4):901–905. doi:10.1039/c2ay05918e
32. Frey A, Meckelein B, Externest D, Schmidt MA (2000) A stable and highly sensitive 3,3′,5,5′-tetramethylbenzidine-based substrate reagent for enzyme-linked immunosorbent assays. *J Immunol Methods* 233(1–2):47–56
33. Matsumoto T, Furuta T, Nimura Y, Suzuki O (1984) 3-*(para)*-Hydroxyphenylpropionic acid as a new fluorogenic reagent for amine oxidase assays. *Anal Biochem* 138(1):133–136
34. Dudley RA, Edwards P, Ekins RP, Finney DJ, McKenzie IGM, Raab GM, Rodbard D, Rodgers RPC (1985) Guidelines for immunoassay data-processing. *Clin Chem* 31(8):1264–1271
35. Ekins R (1981) The Precision Profile: its use in RIA assessment and design. *The Ligand Quarterly* 4(2):33–44
36. Findlay JWA, Smith WC, Lee JW, Nordblom GD, Das I, DeSilva BS, Khan MN, Bowsher RR (2000) Validation of immunoassays for bioanalysis: a pharmaceutical industry perspective. *J Pharm Biomed Anal* 21(6):1249–1273. doi:10.1016/s0731-7085(99)00244-7
37. Hayashi Y, Matsuda R, Maitani T, Imai K, Nishimura W, Ito K, Maeda M (2004) Precision, limit of detection and range of quantitation in competitive ELISA. *Anal Chem* 76(5):1295–1301. doi:10.1021/ac0302859