



Standardized LC–MS/MS based steroid hormone profile-analysis

Therese Koal^{a,*}, Diane Schmiederer^a, Hai Pham-Tuan^a, Cornelia Röhring^a, Manfred Rauh^b

^a BIOCRRATES Life Sciences AG, Innrain 66, A-6020 Innsbruck, Austria

^b Kinder- und Jugendklinik, Universitätsklinikum Erlangen, Loschgestr. 15, 91054 Erlangen, Germany

ARTICLE INFO

Article history:

Received 14 September 2011

Received in revised form 1 December 2011

Accepted 4 December 2011

Keywords:

Steroid hormones
High performance liquid chromatography–tandem mass spectrometry (HPLC–MS/MS)
Endocrinology
Kit
Standardization
Metyrapone

ABSTRACT

In order to overcome many limitations of immunoassays, high performance liquid chromatography–tandem mass spectrometry (HPLC–MS/MS) has the potential to find its place in the clinical laboratory medicine for quantification of steroid hormones. A prerequisite for the application of a new analytical procedure in clinical diagnostics is standardization to minimize analytical intra- and interlaboratory variability and inaccuracy. We evaluate a newly standardized HPLC–MS/MS assay in kit-format, developed for routine determination of 16 steroid hormones in human serum samples. Fifteen metabolites can be measured quantitatively, which include aldosterone, androstenedione, androsterone, corticosterone, cortisol, cortisone, 11-deoxycortisol, dehydroepiandrosterone (DHEA), dehydroepiandrosterone sulfate (DHEAS), 17 β -estradiol (E2), estrone (E1), etiocholanolone, 17 α -hydroxyprogesterone (17OHP), progesterone, and testosterone. 11-Deoxycorticosterone is the only compound rated as semi-quantitative in this kit. The sample preparation is performed by solid phase extraction (SPE) on a 96-well plate. The standardized assay has been validated for human serum in terms of lower and upper limit of quantification (LLOQ 0.01–32 ng/mL, ULOQ 5–8000 ng/mL), linear correlation coefficient of calibration ($R^2 > 0.9966$), intra- and inter-day precision (intra-day 1.1–8.8%, inter-day 5.2–14.8% and 8.2–18.6% for 11-deoxycorticosterone), accuracy (intra-day 88.3–115.5% and 109.3–128.2% for 11-deoxycorticosterone, inter-day 91.4–117.2% and 102.3–137.1% for 11-deoxycorticosterone), analytical total error (3.6–17.8%), proficiency test accuracy (85.4–113.4%), recovery (68–99%), and metabolite stability (freeze/thaw stability 95.5–108.1%, short term stability 86.9–107.2%). Inter-assay comparison with a routine reference HPLC–MS/MS assay and seven immunoassays demonstrates the outstanding high performance of this HPLC–MS/MS based kit by improvements in accuracy for progesterone, androstenedione, and 17OHP. Finally, results of two metyrapone tests demonstrate the potential of the standardized HPLC–MS/MS assay for the analysis of a comprehensive steroid hormone profile in clinical diagnostics.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

The reliable and simultaneous analysis of a broad panel of steroid hormones is a powerful tool for the investigation of the hormone status, which is relevant for a variety of clinical questions and diagnoses (e.g. congenital adrenal hyperplasia, polycystic ovary syndrome, adrenal insufficiency and metyrapone test).

The high performance liquid chromatography–tandem mass spectrometry (HPLC–MS/MS) is able to simultaneously analyze multiple target analytes within on sample run. It is shown over the past ten years that HPLC–MS/MS is a reliable standard technique in clinical application fields like therapeutic drug monitoring (TDM), newborn screening, and toxicological diagnostics. Therefore, HPLC–MS/MS has the potential to become a strong competitor for widely used immunoassays for steroid analysis, especially when

more than one steroid hormone has to be analyzed with high selectivity, sensitivity, precision and accuracy. Main disadvantages of immunoassays, particularly relevant for steroid hormone analysis [1–22] are as follows:

- Measurement of one analyte at a time in one processed sample.
- Lack of antibody specificity due to cross reactivities with structurally similar metabolites (poor accuracy) especially at low concentration ranges, which can result in diagnostic errors e.g. in pediatric range of application.
- Limitation in sensitivity (e.g. testosterone in both premenopausal and postmenopausal women) [7].
- Matrix interferences, which cannot be corrected by internal standards.
- Limited dynamic range.
- High variabilities among different immunoassay kits (resulting in significant variations in reference intervals) and among different laboratories employing the same kit due to poor and diverse validation and standardization data.

* Corresponding author. Tel.: +43 512 579 823 4232; fax: +43 512 579 823 4270.
E-mail address: therese.koal@biocrates.com (T. Koal).

Table 1

Analytical parameters of the SteroIDQ[®] kit: 16 analytes and deuterium-labeled internal standards pairs, MRM of analytes, analyte retention times, calibration ranges and calibration correlation coefficients, calibration with linear regression (except androstenedione, corticosterone, 11-deoxycorticosterone, 17OHP, and testosterone with quadratic regression) and 1/x weighting.

No.	Analyte	Internal standard	Analyte MRM	Retention time [min]	Calibration range [ng/mL]	Correlation coefficient (R^2)
1	Aldosterone	d ₇ aldosterone	361.2/343.2	1.77	0.05–5	0.9978
2	Androstenedione	d ₃ androstenedione	287.2/97.1	6.74	0.03–8	0.9994
3	Androsterone	d ₄ androsterone	273.2/255.1	8.05	0.06–6	0.9977
4	Corticosterone	d ₈ corticosterone	347.2/329.1	4.39	0.03–30	0.9989
5	Cortisol	d ₄ cortisol	363.2/345.1	2.34	1–1000	0.9988
6	Cortisone	d ₇ cortisone	361.2/163.1	2.37	0.10–100	0.9968
7	11-Deoxycorticosterone	d ₈ 17OHP	331.2/109.1	6.49	0.03–15	0.9997
8	11-Deoxycortisol	d ₅ 11-deoxycortisol	347.2/109.1	4.75	0.01–10	0.9995
9	DHEA	d ₄ E1	271.2/253.2	6.75	0.12–30	0.9969
10	DHEAS	d ₅ DHEAS	271.2/253.2	2.78	32–8000	0.9997
11	E2	d ₃ E2	255.2/159.1	5.84	0.02–20	0.9985
12	E1	d ₄ E1	271.2/253.2	6.62	0.03–15	0.9984
13	Etiocolanolone	d ₄ androsterone	273.2/255.1	7.86	0.06–6	0.9991
14	17OHP	d ₈ 17OHP	331.2/109.1	6.96	0.05–50	0.9996
15	Progesterone	d ₉ progesterone	315.2/109.1	8.46	0.06–15	0.9966
16	Testosterone	d ₅ testosterone	289.2/97.0	6.19	0.01–10	0.9993

calibration matrix, and three ampules containing liquid additives for the mobile phase. The kit allows the standardized analysis of up to 80 serum samples. The solid phase extraction (SPE) sample preparation is scheduled to be performed manually in 96 well plate format. However, this procedure can be adapted to robotic platforms as well. A sample volume of 500 μ L serum is needed using 4000QTrap[®] or API4000[™] MS/MS instrument. A system suitability test is part of the kit and has to be performed before running the samples to ensure the inter-day performance of the instrument.

A kit specific steroid SPE-plate in 96 well plate design, a HPLC precolumn (Security Guard[™], Phenomenex, Aschaffenburg, Germany), and a kit specific HPLC column are included in delivery of the kit. In addition one 96 deep well plate (2 mL well volume) for the first sample preparation step (addition of internal standard and sample dilution), two 96 capture well plates (1.2 mL well volume) for the subsequent SPE extracts with two appropriate cover mates are also included. Acetone, acetonitrile, dichloromethane, isopropanol, methanol, purified water (Millipore ultrapure water purification system, Millipore, Vienna, Austria) as solvents for sample preparation and HPLC–MS/MS analysis have to be provided by the user.

The 96 well plate based SPE sample preparation procedure needs a vacuum chamber for 96 well plate (e.g. PlatePrep 96 well, Supelco, Sigma Aldrich, Austria), vacuum pump for SPE (e.g. vacuum pump 220 V 50 Hz, Waters, Austria), evaporator system for 96 well plate using nitrogen with thermostat and heating block (e.g. EVA LS1 MT S with stative, LS1V 130, HBS1 MT R, VLM GmbH, Germany), nitrogen supply (minimum 4 bar), shaker for 96 well plates (e.g. Thermomixer comfort, Eppendorf, Germany) in the laboratory.

16 external standards with 13 deuterium-labeled internal standards are used for external seven-point calibration, as shown in Table 1. No individual stable isotope labeled IS were available for 11-deoxycorticosterone, DHEA, and etiocholanolone. Therefore, alternative IS with close retention times were used. The calibration ranges were optimized with reference to expected endogenous concentrations in human serum and analytical feasibility and are summarized in Table 1.

2.2. SteroIDQ[®] kit sample preparation

500 μ L of blank, calibrators, QC and serum samples were placed into individual wells of a 2-mL 96-deep well plate. 10 μ L reconstituted internal standard mixture (except for blank) and 400 μ L purified water were added to each well. The whole contents of the 2-mL 96-deep well plate have to be transferred to the appropriate

positions of the pre-conditioned SPE plate (1 mL methanol followed by 1 mL purified water). The samples on the SPE plate were allowed to pass through the SPE material bed. A subsequent washing step with 500 μ L purified water and a vacuum assisted drying step with nitrogen prepared the SPE plate for the elution procedure. The elution of the steroid hormones was carried out in two subsequent steps using two individual 96 capture well plates. The first SPE extract was obtained by elution twice with 500 μ L dichloromethane, drying of the extract under nitrogen and final reconstitution with 50 μ L 40/60 (v/v) methanol/purified water. The second extract was generated by elution with 600 μ L acetonitrile and dilution with 400 μ L purified water. Both individual extract plates were covered and placed into the autosampler for subsequent HPLC–MS/MS analysis. The first extract contained all 15 steroid hormones except DHEAS, which was present in the second extract. The overall sample processing time for the 96-well plate based assay (80 samples) comprised around 5 h for SPE sample preparation, 32 h for HPLC–MS/MS analyses, and around 1 h for data analysis.

2.3. SteroIDQ[®] kit HPLC–MS/MS

For chromatographic separation a steroid specific HPLC column and precolumn were used. The HPLC system was an HPLC 1100 series (Agilent Technologies, Vienna, Austria) consisting of a binary pump, a micro vacuum degasser, and a column oven (45 °C). A temperature controlled HTC PAL autosampler (CTC Analytics, Zwingen, Switzerland) equipped with a 20 μ L sample loop was used for sample storage at 10 °C and sample injection (injection volume 20 μ L). 500 mL mobile phase A (470 mL purified water plus contents of additive ampules 1, 2, and 3) and 1000 mL mobile phase B (85/10/5, v/v acetonitrile/methanol/purified water) have to be prepared to run the HPLC–MS/MS assay. 4000QTrap[®] triple quadrupole mass spectrometer (AB Sciex, Toronto, Canada) equipped with a ESI-Turbo V source operating in positive ionization mode and controlled by Analyst 1.5.1 software (AB Sciex, Toronto, Canada) was used for detection.

Multiple reaction monitoring (MRM) was applied for highly selective and sensitive detection of the analytes (summarized in Table 1) and internal standards d₇-aldosterone (368.2/350.2), d₃-androstenedione (290.2/100.1), d₄-androsterone (277.2/259.1), d₈-corticosterone (355.2/337.1), d₄-cortisol (367.2/121.1), d₇-cortisone (368.2/169.1), d₅-11-deoxycortisol (352.2/113.1), d₅-DHEAS (276.2/258.2), d₃-E2 (258.2/159.1), d₄-E1 (275.2/257.2), d₈-17OHP (339.2/113.1), d₉-progesterone (324.2/113.1), and d₅-testosterone (294.2/100.1). HPLC–MS/MS and data analysis method

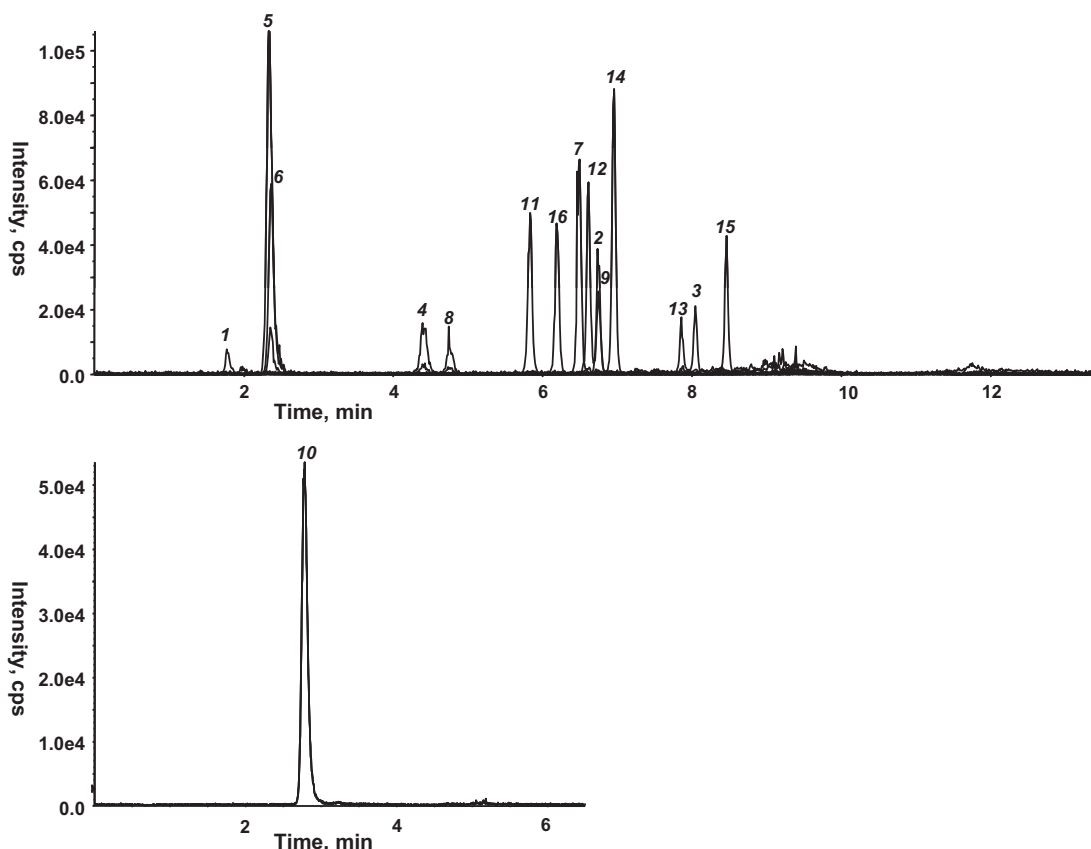


Fig. 2. Representative HPLC–MS/MS chromatogram: A typical extracted ion chromatogram (XIC) is shown for calibrator 3 containing the following concentrations; (1) aldosterone 0.20 ng/mL, (2) androstenedione 0.13 ng/mL, (3) androsterone 0.24 ng/mL, (4) corticosterone 0.60 ng/mL, (5) cortisol 20 ng/mL, (6) cortisone 2.0 ng/mL, (7) 11-deoxycorticosterone 0.60 ng/mL, (8) 11-deoxycortisol 0.20 ng/mL, (9) DHEA 0.48 ng/mL, (10) DHEAS 128 ng/mL, (11) E2 0.40 ng/mL, (12) E1 0.60 ng/mL, (13) etiocholanolone 0.24 ng/mL, (14) 17OHP 1.0 ng/mL, (15) progesterone 0.24 ng/mL, and (16) testosterone 0.20 ng/mL.

files containing all relevant parameters were included in the *SteroIDQ*[®] kit. Both sample extracts (dichloromethane and acetonitrile extract) have to be analyzed in alternating sequence. The HPLC–MS/MS run time is 20 min including system re-equilibration per sample (13.5 min dichloromethane extract, 6.5 min acetonitrile extract). Representative extracted ion chromatograms (XIC) are shown in Fig. 2. Individual retention times are listed in Table 1.

2.4. Reference HPLC–MS/MS assay

The modified online SPE–HPLC–MS/MS assay developed by Rauh et al., allowing quantitative analysis of steroid hormones in 100 μ L serum [2–4], was used as a reference HPLC–MS/MS assay for inter-assay comparison with the *SteroIDQ*[®] kit. 11 steroid hormones (aldosterone, androstenedione, corticosterone, cortisol, cortisone, 11-deoxycortisol, DHEA, DHEAS, 17OHP, progesterone, and testosterone) were covered in this method.

In brief, the assay sample preparation is based on protein precipitation with methanol/zinc sulfate solution (50 g/L, 1/1, v/v). The online SPE was performed by a Chromolith column (50 mm \times 4.6 mm), which was coupled to a Chromolith HPLC column (RP-18e, 100 mm \times 3.0 mm, Merck, Darmstadt, Germany). The autosampler was a HTC PAL (CTC Analytics, Zwingen, Switzerland) fitted with a 250 μ L sample loop. The HPLC consisted of a binary pump (1100 series, Agilent Technologies, Waldbronn, Germany). For MRM based mass spectrometric detection a 4000QTrap[®] triple quadrupole mass spectrometer (AB Sciex, Toronto, Canada) equipped with an APCI source in positive mode was used. The total online SPE–HPLC–MS/MS analysis time per sample was 8 min.

2.5. Immunoassay

Cortisol, DHEAS, progesterone and testosterone were measured by electrochemiluminescence immunoassay on a Cobas e411 immunoassay analyser (Roche Diagnostics, Mannheim, Germany). Chemiluminescent assay (Immulite 2000 Siemens, Healthcare, Marburg, Germany) was used for the analysis of androstenedione. Aldosterone and 17OHP were measured by radioimmunoassay (Siemens Healthcare, Marburg, Germany and IBL, Hamburg, Germany). All assays were applied according to the instructions of the manufacturers.

2.6. Serum samples

Methods were compared by Passing/Bablok regression analysis within specified measuring ranges using left-over routine samples (children, male, female) which were collected at the university hospital of Erlangen. Serum samples were stored at -20°C until randomized analysis. Different n numbers of samples (Table 4, Fig. 3) are either due to missing sample volume or differences in test ranges.

3. Results

3.1. *SteroIDQ*[®] kit validation

The *SteroIDQ*[®] kit was validated for human serum on AB Sciex API4000[™] and 4000QTRAP[®] instruments according to the FDA Guidance for Industry – Bioanalytical Method Validation [24]. The analytical selectivity of the developed assay is based on the

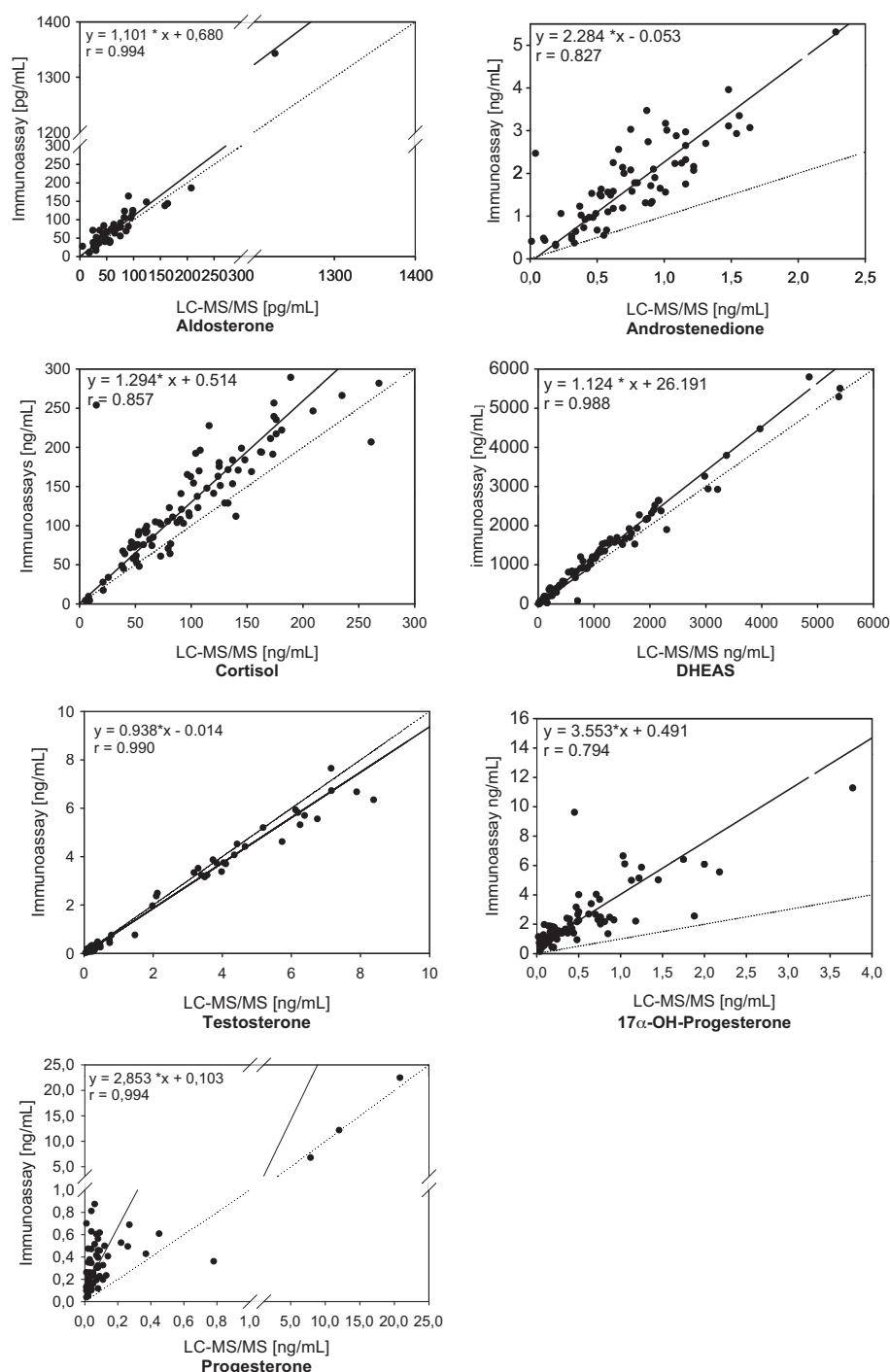


Fig. 3. Interassay comparison to immunoassays: The comparabilities in concentration for seven steroid hormones (aldosterone, androstenedione, cortisol, DHEAS, testosterone, 17OHP, and progesterone) were investigated between SteroIDQ® kit and immunoassays showing acceptable correlation and comparability only for four steroid hormones; aldosterone, DHEAS, cortisol and testosterone (Section 3.3).

complementary combination of SPE (sample clean-up and matrix removal), HPLC (chromatographic separation of isobaric analytes) and MS/MS (MRM as most selective MS/MS detection mode), which is challenging especially for steroid hormones in biological sample matrices. All 16 external and 13 internal standards were investigated. No interferences were observed except for cortisol and cortisone. A 3.1% of the cortisone concentration was detected as cortisol and 0.1% of the cortisol concentration as cortisone. In human serum samples these interferences are negligible due to the fact that cortisol is typically in 3–10 fold higher concentrations than cortisone. In addition, the presence of structurally related steroid

hormone drugs can result in interferences as well. Therefore, the selectivity of synthetic corticosteroid drugs (methylprednisolone, prednisolone, prednisone) were tested and can be confirmed by different retention times [27].

The lower (LLOQ) and upper (ULOQ) limits of quantification correlate with the calibration range (LLOQ 0.01–32 ng/mL, ULOQ 5–8000 ng/mL). The assay performance parameters are summarized in Tables 1–3. Acceptable linear regression correlation coefficients between 0.9966 and 0.9997 were obtained for calibration. Intra-day (within batch, $n=6$) and inter-day (batch to batch, $n=8$) accuracy and precision were investigated for spiked human

Table 2
Assay performance parameters: intra-day and inter-day accuracy and precision for spiked serum pool samples at three different concentrations.

No.	Analyte	Concentration [ng/mL]	Intraday (n = 6)		Interday (n = 8)	
			Accuracy [%]	Precision [%]	Accuracy [%]	Precision [%]
1	Aldosterone	0.22	88.3	3.7	98.3	11.5
		0.46	102.1	4.4	109.3	9.6
		3.1	100.1	4.4	103.4	5.9
2	Androstenedione	1.0	99.8	2.2	98.8	9.9
		1.5	104.2	1.7	107.0	9.6
		5.7	100.5	1.9	106.6	6.1
3	Androsterone	0.38	96.1	3.2	97.4	12.4
		0.66	104.2	4.8	102.1	13.2
		3.7	99.7	3.4	98.9	12.1
4	Corticosterone	2.2	102.6	2.5	100.3	16.4
		4.3	109.0	5.8	102.4	11.7
		19	98.9	2.5	107.3	7.9
5	Cortisol	106	103.0	3.6	98.7	9.0
		180	110.1	1.1	107.4	7.9
		718	105.1	1.5	104.4	5.2
6	Cortisone	18	99.1	4.7	101.1	8.5
		25	103.8	2.3	106.1	9.7
		77	101.8	4.2	106.1	8.0
7	11-Deoxycorticosterone	0.79	122.9	2.5	137.1	18.6
		1.6	128.2	4.4	134.2	13.7
		10	109.3	2.9	102.3	8.2
8	11-Deoxycortisol	0.45	96.6	8.8	98.2	9.5
		1.1	105.0	2.0	108.4	9.7
		6.2	98.4	5.2	101.2	14.8
9	DHEA	2.2	97.8	3.7	93.0	9.6
		4.2	101.9	2.9	98.1	13.4
		18	90.7	2.0	94.3	13.3
10	DHEAS	1860	100.8	1.7	97.5	6.8
		2492	105.7	1.9	101.7	8.4
		7073	108.5	1.4	102.9	7.5
11	E2	0.53	105.7	4.6	106.2	12.3
		1.9	109.9	1.8	117.2	12.3
		13	105.5	5.6	107.6	7.5
12	E1	0.68	102.7	2.9	100.8	11.5
		1.4	110.6	3.2	108.4	5.8
		9.1	100.1	3.5	100.2	6.3
13	Etiocolanalone	0.30	100.4	4.7	101.8	14.6
		0.62	113.3	2.7	117.1	14.3
		4.2	115.5	5.4	109.9	17.4
14	17OHP	1.7	103.0	3.1	102.1	9.4
		5.2	110.4	2.6	110.6	9.6
		33	106.5	3.7	104.0	6.3
15	Progesterone	1.8	95.9	2.1	95.8	8.7
		2.8	97.8	2.5	100.0	9.1
		9.4	88.5	3.9	91.4	10.5
16	Testosterone	0.50	90.2	2.6	103.7	10.6
		1.1	97.3	2.6	110.6	11.2
		6.6	103.3	3.3	105.0	10.5

Table 3
Assay performance parameters: recoveries and stabilities (RT = room temperature).

No.	Analyte	Recovery, n = 4 [%]	Stability at low concentration, n = 6 [%]			Stability at high concentration, n = 6 [%]		
			3 freeze-thaw cycles	Short term 24 h at 4 °C	Short term 4 h at RT	3 freeze-thaw cycles	Short term 24 h at 4 °C	Short term 4 h at RT
1	Aldosterone	78.2 ± 8.5	100.4	107.2	97.0	99.8	100.3	97.7
2	Androstenedione	87.6 ± 2.5	98.5	99.7	100.3	98.1	97.4	97.9
3	Androsterone	70.7 ± 4.9	96.7	91.2	98.6	96.6	86.9	93.4
4	Corticosterone	89.5 ± 3.5	103.7	104.2	105.0	95.5	96.7	97.0
5	Cortisol	93.5 ± 4.7	97.8	100.4	100.2	96.9	97.5	97.6
6	Cortisone	92.5 ± 3.1	99.0	98.6	100.1	96.4	97.1	98.1
7	11-Deoxycorticosterone	83.0 ± 2.0	108.1	102.6	105.1	104.1	99.2	100.9
8	11-Deoxycortisol	87.8 ± 4.3	96.0	95.3	103.3	96.0	95.4	97.7
9	DHEA	95.8 ± 2.5	100.9	98.9	101.3	100.1	98.2	97.9
10	DHEAS	99.2 ± 3.7	97.0	99.2	100.4	97.7	98.3	99.5
11	E2	90.0 ± 4.3	100.5	104.9	104.5	99.4	101.2	97.1
12	E1	88.7 ± 5.9	99.4	100.0	100.6	102.2	99.6	98.1
13	Etiocolanalone	73.7 ± 4.1	98.8	103.5	99.4	99.7	98.0	100.0
14	17OHP	83.1 ± 0.7	97.9	98.1	99.0	97.3	100.1	97.0
15	Progesterone	68.2 ± 5.3	97.5	103.0	102.4	96.6	100.2	98.5
16	Testosterone	87.6 ± 2.7	101.3	101.2	97.3	101.4	100.8	103.6

serum (see Table 2). Inter-day experiments were performed by 4 different operators and with 2 different HPLC–MS/MS instruments. Excellent accuracy and precision results were achieved for all analytes except for 11-deoxycorticosterone. The accuracy of ~130% for low and medium concentrations obtained for 11-deoxycorticosterone did not fulfil the acceptance criteria of $100 \pm 15\%$ set for this assay. Therefore, 11-deoxycorticosterone was rated as a semiquantitative analyte.

The total errors were determined for steroid hormones, which are linked with Westgard's criteria [25]. In principle (biological and analytical) imprecision related with the analytical inaccuracy are under combined consideration, which is named as acceptable total error in % (total error (TE) = analytical inaccuracy + $1.65 \times$ analytical imprecision). For example, the maximum within-subject biological variation for 17OHP is specified by Westgard as 19.6%. The maximum analytical imprecision for 17OHP is 9.8% (equivalent to 50% of within-subject biological variation). The defined limit for analytical inaccuracy is 14%. A maximum total error of 30.2% ($=14\% + 1.65 \times 9.8\%$) is calculated for 17OHP. The analytical imprecision of 17OHP in the newly developed assay is 2.6–3.7% (intra-day) and 6.3–9.6% (inter-day) as listed in Table 2 and fulfils the acceptance criterion of 9.8%. The total errors were between 8.1 and 14.7% and are within the acceptance criterion of 30.2%. Analyte specific acceptance criteria were available for aldosterone (TE 27.1%), androstenedione (TE 22.6%), cortisol (TE 29.8%), DHEAS (TE 10.9%), 11-deoxycortisol (TE 27.1%), E2 (TE 21.6%), 17OHP (TE 30.2%), and testosterone (TE 14%) so far. The calculated total errors of these steroid hormones determined for the newly developed assay (aldosterone 7.4–17.8%, androstenedione 3.8–7.0%, cortisol 7.6–11.9%, DHEAS 3.6–10.8%, 11-deoxycortisol 8.3–17.9%, E2 12.9–14.7%, 17OHP 8.1–14.7%, testosterone 7.0–14.1%) are within the individual criteria.

The accuracy performance of the assay was approved by analysis of reference proficiency test samples (proficiency test program at DGKL (*Deutsche Vereinte Gesellschaft für Klinische Chemie und Laboratoriumsmedizin e.V.*). Seven steroid hormones (aldosterone, cortisol, DHEAS, E2, 17OHP, progesterone and testosterone) are covered in the proficiency test program. The target values were determined by isotope-dilution gas chromatography–mass spectrometry analysis (GC–MS) except for DHEAS (calculated median of all measurements). Excellent accuracies were obtained for all analytes measured in two proficiency test samples (aldosterone 95.2 and 102.8% (valid range 55–144%), cortisol 96.4 and 103.0% (valid range 70–130%), E2 109.2 and 113.4% (valid range 65–135%), progesterone 100.0 and 102.3% (valid range 65–135%), testosterone 86.4 and 98.4% (valid range 65–135%), DHEAS 108.5 and 110.0 (valid range 65–135%), 17OHP 85.4 and 98.1% (valid range 40–161%) demonstrating the excellent comparability of HPLC–MS/MS regarding to the GC–MS reference method. The relatively high acceptance ranges are due to the variety of different immunoassays in this program. Currently less than 2% (mostly less than 0.5%) proficiency test participants are using mass spectrometry methods.

The recovery was determined calculating the ratio between the peak area of the analyte in the final extract after SPE and the appropriate peak area measured without SPE. Highly precise recoveries between 68 and 99% with coefficients of variation between 0.7 and 8.5% (Table 3) were obtained. Strongly related to recovery are matrix effects, which have to be investigated as well. The use of appropriate internal standards ideally corrects (sample individually) matrix effects. The set-up for this validation experiment was post-extraction spiking of the target analytes at 3 different concentrations and internal standards into the final extracts of human pooled serum and buffer as reference ($n=4$). The uncorrected matrix effect was calculated as peak area of the target analyte in human pooled serum compared to buffer and the corrected matrix effect by internal standard as peak area ratio of the target

analyte and internal standard in human pooled serum compared to buffer. Corrected matrix effects less than 7% were obtained (data not shown) demonstrating the benefit of appropriate internal standards, which were not applied in immunoassay analysis.

Furthermore, several stability experiments were performed for pooled human serum at two spiked concentrations investigating the influence of different sample and sample extracts storage. Freeze–thaw stabilities were determined for the recommended 3 freeze (-20°C) and thaw (at room temperature) cycles. Similarly, investigations were carried out for short term temperature stability (sample 24 h at 4°C , 4 h at room temperature) and post-preparative sample extract stabilities (72 h at 10°C in the autosampler, 7 days at -20°C , data not shown). The results demonstrated that no significant analyte losses were obtained for all stability tests.

3.2. Interassay comparison to reference online SPE–HPLC–MS/MS

Disease affected and non-affected serum samples from children, men and women were included for comparison to the reference online SPE–HPLC–MS/MS assay. The samples covering the whole range of relevant concentrations (e.g. congenital adrenal hyperplasia with increased 17OHP up to 85 ng/mL) were processed by both assays. The samples were measured with the reference method in different batches. The results, summarized in Table 4, showed excellent comparability, i.e. slopes between 0.829 and 1.185, intercept values between -4.77 and 2.022 , correlation coefficients >0.884 , and mean differences ranging from -11.4 up to 20.01 for all steroid hormones.

3.3. Interassay comparison to immunoassays

The performance of the SteroIDQ[®] kit was compared to immunoassays for seven steroid hormones in human serum samples (children, men, women). The sensitivities of both analytical technologies were comparable (aldosterone 11 pg/mL, androstenedione 0.3 ng/mL, cortisol <5 ng/mL, DHEAS 1 ng/mL, 17OHP 0.02 ng/mL, progesterone 0.03 ng/mL, testosterone 0.02 ng/mL). The results are shown in Fig. 3. Adequate correlation and comparability were obtained for aldosterone, DHEAS, cortisol and testosterone. Progesterone showed a clear concentration depended comparability with only acceptable comparability for higher concentrations. Approximately 2–3 fold higher concentrations were obtained for progesterone measured by electrochemiluminescence immunoassay at low concentrations up to 1 ng/mL. Finally, inadequate comparability could be detected for androstenedione between HPLC–MS/MS and chemiluminescence assay and 17OHP between HPLC–MS/MS and radioimmunoassay over the whole range of tested concentrations.

3.4. Metyrapone test

Metyrapone (trade name Metopirone) is a drug used in the diagnosis of adrenal insufficiency and occasionally in the treatment of Cushing's syndrome (hypercortisolism) [28–33]. Metyrapone blocks cortisol synthesis (decrease of cortisol) in healthy humans by inhibiting steroid 11 β -hydroxylase in the cortex of suprarenal gland. This stimulates adrenocorticotrophic hormone (ACTH) secretion by negative feedback, which results in an increase of 11-deoxycortisol and 11-deoxycorticosterone levels. 11-Deoxycortisol and cortisol concentrations in serum sample are usually measured for the metyrapone test. Cortisol levels below 220–280 nmol/L (80–102 ng/mL) and 11-deoxycortisol levels above 7 $\mu\text{g}/\text{dL}$ (70 ng/mL) indicate a positive metyrapone test. A negative metyrapone test (insufficient increase of 11-deoxycortisol) can indicate deficiency of the cortex of suprarenal

Table 4
Comparison of 11 steroid hormones between SteroIDQ® kit and reference online SPE-HPLC–MS/MS assay.

Steroid	<i>n</i>	Concentration range (ng/mL)	Slope (95% confidence region, P/B)	Intercept (95% confidence region, P/B)	Correlation coefficient	Mean difference (%)
Aldosterone	43	0.005–1.227	1.136 (0.995–1.383)	0.004 (–0.007 to 0.011)	0.993	15.7
Androstenedione	123	0.001–4.386	0.907 (0.852–0.972)	–0.018 (–0.036 to –0.002)	0.974	–11.4
Corticosterone	118	0.022–13.74	1.185 (1.094–1.309)	–0.074 (–0.156 to 0.020)	0.935	12.5
Cortisol	122	5.24–1011	0.930 (0.897–0.966)	2.022 (–0.203 to 3.677)	0.987	–4.81
Cortisone	124	0.16–51.8	1.035 (0.980–1.098)	0.422 (–0.468 to 1.259)	0.944	7.10
11-Deoxycortisol	93	0.032–6.61	1.183 (1.094–1.330)	0.029 (0.0 to 0.048)	0.968	20.01
DHEA	84	0.066–10.1	0.922 (0.826–1.020)	–0.087 (–0.234 to 0.038)	0.884	–9.8
DHEA-S	108	1.1–5618	0.915 (0.893–0.934)	–4.770 (–17.042 to 1.644)	0.985	–8.8
17-OH-Progesterone	101	0.064–85.58	1.014 (0.957–1.114)	0.016 (–0.009 to 0.032)	0.972	0.1
Progesterone	42	0.008–19.084	0.829 (0.706–0.957)	0.119 (0.103 to 0.134)	0.986	0.1
Testosterone	99	0.044–10.46	0.927 (0.864–0.981)	–0.012 (–0.027 to 0.002)	0.972	–10.0

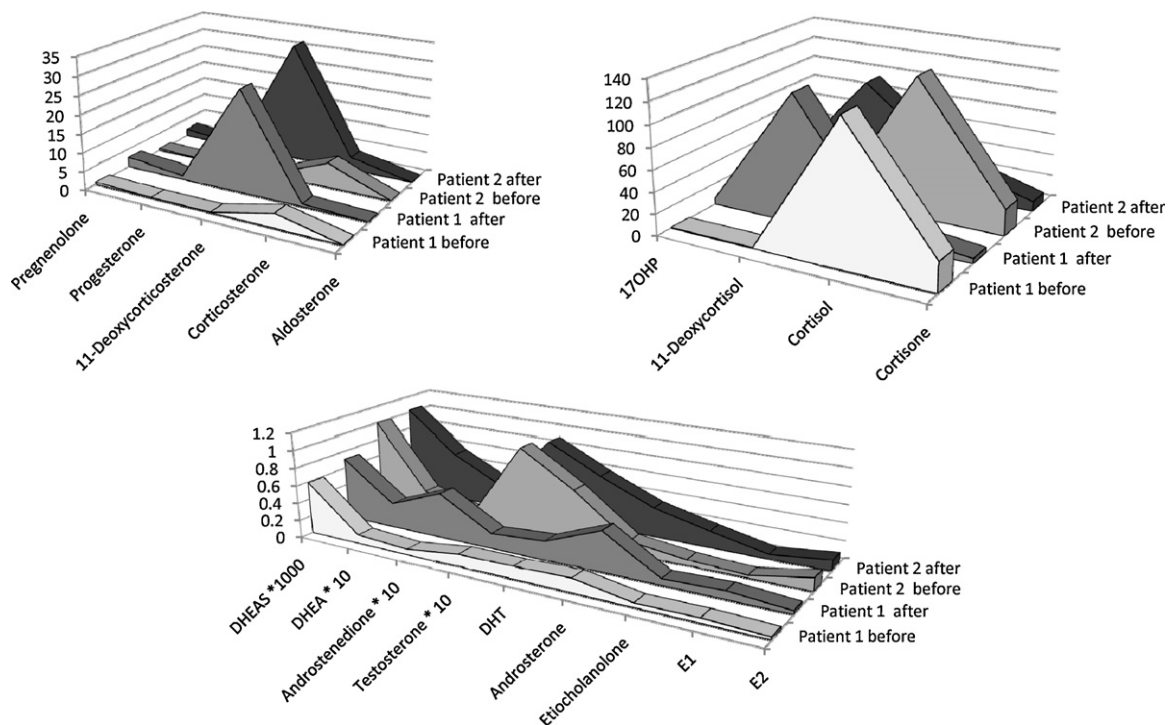


Fig. 4. Steroid hormone profiles of metyrapone test samples (*y*-axis in ng/mL): Two positive metyrapone test results (before and after metyrapone treatment) are pictured demonstrating the same metyrapone effects i.e. increase for and 11-deoxycortisol, 11-deoxycorticosterone, decrease for cortisol and cortisone, and appropriate alleviated effects for the neighbor metabolites. Additionally the relationship and dependencies between the metabolites in the steroid hormone pathway can be followed as well because all three diagrams contain a section of the pathway. (Final DHEAS concentration multiplied by factor 1000; final DHEA, androstenedione, and testosterone concentrations multiplied by factor 10.)

gland but cannot clarify the cause (pituitary or cortex of suprarenal glands).

Fig. 4 shows the steroid hormone profiles of two positive metyrapone tests of two male patients. The analysis and quantification of dihydrotestosterone (DHT) and pregnenolone (not validated metabolites) were included into this study covering eighteen steroid hormones in total. Similar steroid profiles were obtained for both patients. Cortisol, cortisone, corticosterone, and aldosterone showed the expected decrease in concentrations (Fig. 4). Cortisol concentrations before the metyrapone test were 130 ng/mL for both patients and 13 ng/mL in patient 1 and 32 ng/mL in patient 2 after the test. 11-Deoxycortisol was 0.7 ng/mL in patient 1 and 0.6 ng/mL in patient 2 before and 118 ng/mL in patient 1 and 94 ng/mL in patient 2 after the metyrapone test. Steroid 11-beta-hydroxylase is responsible for metabolism of 11-deoxycorticosterone to corticosterone as well. Therefore, metyrapone is blocking the corticosterone synthesis in parallel to the synthesis of cortisol. Aldosterone was only detected in the sample of patient 1 before the metyrapone test with a concentration

of 0.1 ng/mL and was lower than LLOQ in the sample after the metyrapone test. However, no significant effects were detected for E2, testosterone and DHT.

4. Discussion

4.1. SteroIDQ® kit validation

Steroid hormone analyses have always been performed by chromatographic separation due to many known isobaric target metabolites as well as isobaric unknown signals in the chromatogram. Therefore, the selectivity of the assay was one of the challenging validation issues especially for steroid hormones and structure-related xenobiotics (e.g. methylprednisolone, prednisolone, prednisone). Insignificant interferences were only detected for cortisol and cortisone, which plays no role in naturally occurring concentrations. The intra- and inter-day accuracy and precision data fulfilled the acceptance criteria for all steroid hormones except for 11-deoxycorticosterone, which can be

explained by missing a suitable internal standard. However, 11-deoxycorticosterone could be classified as semiquantitative due to acceptable precision results. Furthermore, highly precise recoveries (68–99%), acceptable corrected matrix effects (using stable isotope labeled IS) (<7%) and stabilities, which are relevant for sample and sample extracts storage, were obtained. The total errors for all investigated steroids are in conformance with clinical requirements.

4.2. Interassay comparison

The performance of the developed HPLC–MS/MS assay was compared to a reference online SPE–HPLC–MS/MS. Different sample clean-up strategies were used for both assays; SPE (SteroIDQ[®] kit) and protein precipitation (online SPE–HPLC–MS/MS). The offline SPE procedure itself and the increased sample volume are the major reasons that 16 steroid hormones can be analyzed by the SteroIDQ[®] kit instead of 11 with the reference assay. Excellent comparabilities for all steroid hormones tested were achieved demonstrating the low variability of HPLC–MS/MS technique.

Furthermore, the performance of the newly developed SteroIDQ[®] kit was compared to immunoassays, which are still standard analysis techniques in the clinical environment. Four analytes (aldosterone, DHEAS, cortisol, testosterone) showed an adequate agreement of concentrations between both techniques. Other steroid hormones showed no (androstenedione with chemiluminescent assay Immulite 2000, 17OHP with radioimmunoassay by IBL) or concentration dependent (progesterone with electrochemiluminescence immunoassay) comparabilities.

Fanelli et al. compared immunoassays and HPLC–MS/MS data among others for cortisol, testosterone, progesterone, androstenedione and 17OHP [19]. Similar discrepancies between HPLC–MS/MS and immunoassay based analysis for progesterone, androstenedione, and 17OHP were obtained by Fanelli et al. confirming the limitation of tested immunoassays due to the lack of immunoassay specificity by cross-reactivities with matrix components resulting in overestimation [19]. An analogous concentration depended comparability was obtained for progesterone (ElecysE170). Androstenedione was analyzed by Fanelli et al. with an Immulite 2000 comparable to our investigations [19]. Both RIA assays for 17OHP from IBL and Bridge (Fanelli et al.) showed the same poor comparability compared to the HPLC–MS/MS results [19]. The more accurate quantification by means of HPLC–MS/MS compared to the immunoassay can be attested by proficiency test sample analysis as reference as described in Section 3.1. Cortisol was in agreement (ElecysE170) as well in Fanelli et al. investigations. Fanelli et al. showed that the comparability of testosterone (ElecysE170) was only acceptable in males, not in females at low concentrations. Rothman et al. investigated levels among others for testosterone in premenopausal women by HPLC–MS/MS and compared the results to previously reported immunoassay data [22]. In the result, lower concentrations obtained with HPLC–MS/MS were reported. Various immunoassays for testosterone were compared in the past to HPLC–MS/MS by Moal et al. [34] (5 immunoassays), Wang et al. [35] (4 immunoassays) and to GC–MS by Taieb et al. [36] (10 immunoassays) confirming that testosterone at low concentrations gives no reliable results for all tested immunoassays. Lack of sensitivity, specificity, and accuracy of many testosterone immunoassays has been intensively discussed in an Endocrine Society position statement [37]. The Cobas e411 immunoassay, which was not tested by Fanelli et al. [19], Rothman et al. [22], Moal et al. [34], Wang et al. [35], Taieb et al. [36] showed comparable results even for low concentrations down to 0.01 ng/mL in our study.

Several outlier samples could be identified for the interassay comparison with immunoassays. A single sample was responsible for the cortisol (253 ng/mL immunoassay versus 15 ng/mL

HPLC–MS/MS) and androstenedione (2.5 ng/mL immunoassay versus 0.04 ng/mL HPLC–MS/MS) outlier. Another sample was the source for the DHEAS (80 ng/mL immunoassay versus 705 ng/mL HPLC–MS/MS) and 17OHP (9.6 ng/mL immunoassay versus 0.4 ng/mL HPLC–MS/MS) outlier. No clinical background information is available to help explaining the outlining of these samples. Significant higher immunoassay concentrations of cortisol, androstenedione, and 17OHP compared to HPLC–MS/MS results might be due to the overestimation, caused by cross-reactivities of the target analyte with structurally related metabolites (synthetic glucocorticoids such as prednisone, prednisolone and methylprednisolone used in drug treatment) in the immunoassays. In the sample, where cortisol and androstenedione are outliers, additional peaks could be detected on the cortisol, cortisone, and aldosterone extracted ion chromatogram (XIC) when it was analyzed with SteroIDQ[®] kit. One signal could be identified as methylprednisolone. Similarly, an unknown compound is detected on the XIC of 17OHP in the sample with the 17OHP outlier. These additional signals might indicate the presence of compounds with similar structure leading to the cross-reactivity in immunoassay, and thus results in overestimation of concentration. The discrepancy for the DHEAS outlier could not be explained. However, these results showed clearly the lack of accuracy for many immunoassays for relevant range of concentrations, which can easily be overcome by HPLC–MS/MS with appropriate sample preparation.

4.3. Metyrapone test

Metyrapone is able to differentiate primary adrenal causes of Cushing's syndrome from other causes [30]. Crapo demonstrated that LC–MS/MS is the analysis technique of choice compared to immunoassay when monitoring patients undergoing treatment with metyrapone [31]. Two independent metyrapone tests samples were analyzed (male, sampling before and after test) showing comparable steroid profiles with effects in more or less all steroid hormone metabolite concentrations except E2, DHT, and testosterone. This confirms the fact that a complete overview of steroid hormones profiles can be easily obtained by employing the newly developed SteroIDQ[®] kit. This information is of utmost importance in providing answers for other clinical questions in endocrinology. Fig. 4 visualizes the relationship, dependencies, and metyrapone effects between the neighbor metabolites in the steroid hormone pathway. Convenient presentation of pathway analysis data is mandatory to bring such complex results to a simple data interpretation and finally to establish pathway analysis of endogenous metabolites in the clinical future.

5. Conclusion

The newly developed standardized and validated HPLC–MS/MS assay in kit format (SteroIDQ[®]) allows the reliable, sensitive and simultaneous determination of 16 steroid hormones overcoming known limitations in accuracy and variations of steroid hormone immunoassays and is approved as a CE marked *in vitro* diagnostic medical device in Germany, Switzerland and Austria. Standardized HPLC–MS/MS in combination with an adequate sample preparation is the method of choice for highly accurate, selective, sensitive, and simultaneous analysis of steroid hormones in the clinical laboratory and has the potential to replace immunoassays in the clinical routine in the future. In addition, the application of HPLC–MS/MS in a standardized format like in kits is a big step forward in improvements of analytical accuracy, precision (especially if more restricted acceptance criteria e.g. Westgard's criteria have to be fulfilled in the future) and in terms of international harmonization of steroid hormone reference ranges.

Acknowledgements

The authors would like to thank J. Biskupek-Sigwart and M. Meier in Erlangen and D. Kirchberg, S. Angeben, A. Kania, and P. Hofer in Innsbruck for technical assistance and sample analysis.

References

- [1] T. Koal, H.P. Deigner, Challenges in mass spectrometry based targeted metabolomics, *Curr. Mol. Med.* 10 (2010) 216–226.
- [2] M. Rauh, M. Gröschl, H.G. Dörr, Automated, fast and sensitive quantification of 17 α -hydroxy-progesterone, androstenedione and testosterone by tandem mass spectrometry with on-line extraction, *Steroids* 71 (2006) 450–458.
- [3] M. Rauh, Steroid measurement with LC–MS/MS in pediatric endocrinology, *Mol. Cell. Endocrinol.* 301 (2009) 272–281.
- [4] M. Rauh, Steroid measurement with LC–MS/MS. Application examples in paediatrics, *J. Steroid Biochem. Mol. Biol.* 121 (2010) 520–527.
- [5] U. Ceglarek, C. Shackleton, F.Z. Stanczyk, J. Adamski, Steroid profiling and analytics: going towards sterome, *J. Steroid Biochem. Mol. Biol.* 121 (2010) 479–480.
- [6] C. Shackleton, Clinical steroid mass spectrometry: a 45-year history culminating in HPLC–MS/MS becoming an essential tool for patient diagnosis, *J. Steroid Biochem. Mol. Biol.* 121 (2010) 481–490.
- [7] F.Z. Stanczyk, N.J. Clarke, Advantages and challenges of mass spectrometry assays for steroid hormones, *J. Steroid Biochem. Mol. Biol.* 121 (2010) 491–495.
- [8] U. Ceglarek, M. Werner, L. Kortz, A. Körner, W. Kiess, J. Thiery, J. Kratzsch, Pre-clinical challenges in steroid analysis of human samples, *J. Steroid Biochem. Mol. Biol.* 121 (2010) 505–512.
- [9] H.W. Vesper, J.C. Botelho, Standardization of testosterone measurements in humans, *J. Steroid Biochem. Mol. Biol.* 121 (2010) 513–519.
- [10] J.P. Holst, O.P. Soldin, T. Guo, S.J. Soldin, Steroid hormones: relevance and measurement in the clinical laboratory, *Clin. Lab. Med.* 24 (2004) 105–118.
- [11] B. Hauser, T. Deschner, C. Boesch, Development of a liquid chromatography–tandem mass spectrometry method for the determination of 23 endogenous steroids in small quantities of primate urine, *J. Chromatogr. B* 862 (2008) 100–112.
- [12] H. Noppe, B. Le Bizec, K. Verheyden, H.F. De Brabander, Novel analytical methods for the determination of steroid hormones in edible matrices, *Anal. Chim. Acta* 611 (2008) 1–16.
- [13] M.M. Kushnir, A.L. Rockwood, W.L. Roberts, B. Yue, J. Bergquist, A.W. Meikle, Liquid chromatography tandem mass spectrometry for analysis of steroids in clinical laboratories, *Clin. Biochem.* 44 (2011) 77–88.
- [14] M.M. Kushnir, T. Blamires, A.L. Rockwood, W.L. Roberts, B. Yue, E. Erdogan, A.M. Bunker, A.W. Meikle, Liquid chromatography–tandem mass spectrometry assay for androstenedione, dehydroepiandrosterone, and testosterone with pediatric and adult reference intervals, *Clin. Chem.* 56 (2010) 1138–1147.
- [15] F.Z. Stanczyk, J.S. Lee, R.J. Santen, Standardization of steroid hormone assays: why, how, and when? *Cancer Epidemiol. Biomarker* 16 (2007) 1713–1719.
- [16] S.J. Soldin, O.P. Soldin, Steroid hormone analysis by tandem mass spectrometry, *Clin. Chem.* 55 (2009) 1061–1066.
- [17] J. Holst, O.P. Soldin, T. Guo, S.J. Soldin, Steroid hormones: relevance and measurement in the clinical laboratory, *Clin. Lab. Med.* 24 (2004) 105–118.
- [18] H.W. Vesper, S. Bhasin, C. Wang, S.S. Tai, L.A. Dodge, R.J. Singh, J. Nelson, S. Ohorodnik, N.J. Clarke, W.A. Salameh, C.R. Parker Jr., R. Razdan, E.A. Monsell, G.L. Myers, Interlaboratory comparison study of serum total testosterone measurements performed by mass spectrometry methods, *Steroids* 74 (2009) 498–503.
- [19] F. Fanelli, I. Belluomo, V.D. Di Lallo, G. Cuomo, R. De Iasio, M. Baccini, E. Casadio, B. Casetta, V. Vicennati, A. Gambineri, G. Grossi, R. Pasquali, U. Pagotto, Serum steroid profiling by isotopic dilution–liquid chromatography–mass spectrometry: comparison with current immunoassays and reference intervals in healthy adults, *Steroids* 76 (2011) 244–253.
- [20] M. Vogeser, C. Seger, A decade of HPLC–MS/MS in the routine clinical laboratory–goals for further developments, *Clin. Biochem.* 41 (2008) 649–662.
- [21] M. Vogeser, K.G. Parhofer, Liquid chromatography tandem–mass spectrometry (LC–MS/MS)–technique and applications in endocrinology, *Exp. Clin. Endocrinol. Diabetes* 115 (2007) 559–570.
- [22] M.S. Rothman, N.E. Carlson, M. Xu, C. Wang, R. Swerdloff, P. Lee, V.H. Goh, E.C. Ridgway, M.E. Wierman, Reexamination of testosterone, dihydrotestosterone, estradiol and estrone levels across the menstrual cycle and in postmenopausal women measured by liquid chromatography–tandem mass spectrometry, *Steroids* 76 (2011) 177–182.
- [23] T. Köcher, P. Pichler, R. Swart, K. Mechtler, Quality control in LC–MS/MS, *Proteomics* 11 (2011) 1026–1030.
- [24] Guidance for Industry/Bioanalytical Method Validation, US Department of Health and Human Services, Food and Drug Administration. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070107.pdf>, 2011 (accessed 08.08.11).
- [25] Desirable specifications for total error, imprecision, and bias, derived from biological variations. <http://westgard.com/biodatabase1.htm>, 2011 (accessed 08.08.11).
- [26] J.O. Westgard, Managing quality vs. measuring uncertainty in the medical laboratory, *Clin. Chem. Lab. Med.* 48 (2010) 31–40.
- [27] I.A. Ionita, D.M. Fast, F. Akhlaghi, Development of a sensitive and selective method for the quantitative analysis of cortisol, cortisone, prednisolone, and prednisone in human plasma, *J. Chromatogr. B* 877 (2009) 765–772.
- [28] L.J. Owen, D.J. Halsall, B.G. Keevil, Cortisol measurement in patients receiving metyrapone therapy, *Ann. Clin. Biochem.* 47 (2010) 573–575.
- [29] W.J. Inder, P.J. Hunt, Glucocorticoid replacement in pituitary surgery: guidelines for perioperative assessment and management, *J. Clin. Endocrinol. Metab.* 87 (2002) 2745–2750.
- [30] J. Newell-Price, P. Trainer, M. Besser, A. Grossman, The diagnosis and differential diagnosis of Cushing's syndrome and pseudo-Cushing's states, *Endocr. Rev.* 19 (1998) 647–672.
- [31] L. Crapo, Cushing's syndrome: a review of diagnostic tests, *Metabolism* 28 (1979) 955–977.
- [32] P. Santhanam, S.F. Saleem, T.F. Saleem, Diagnostic predicament of secondary adrenal insufficiency, *Endocr. Pract.* 16 (2010) 686–691.
- [33] K.I. Alexandraki, A.B. Grossman, Medical therapy of Cushing's disease: where are we now, *Front. Horm. Res.* 38 (2010) 165–173.
- [34] V. Moal, Mathieu, P. Reynier, Y. Malhiéry, Y. Gallois, Low serum testosterone assayed by liquid chromatography–tandem mass spectrometry. Comparison with five immunoassay techniques, *Clin. Chim. Acta* 386 (2007) 12–19.
- [35] C. Wang, D.H. Catlin, L.M. Demers, B. Starcevic, R.S. Swerdloff, Measurement of total serum testosterone in adult men: comparison of current laboratory methods versus liquid chromatography–tandem mass spectrometry, *J. Clin. Endocrinol. Metab.* 89 (2004) 534–543.
- [36] J. Taieb, B. Mathian, F. Millot, M.C. Patricot, E. Mathieu, N. Queyrel, I. Lacroix, C. Somma-Delpero, P. Boudou, Testosterone measured by 10 immunoassays and by isotope-dilution gas chromatography–mass spectrometry in sera from 116 men, women, and children, *Clin. Chem.* 49 (2003) 1381–1395.
- [37] W. Rosner, R.J. Auchus, R. Azziz, P.M. Sluss, H. Raff, Position statement: utility, limitations, and pitfalls in measuring testosterone: an Endocrine Society position statement, *J. Clin. Endocrinol. Metab.* 92 (2007) 405–413.