Toxicology in Vitro 29 (2015) 113–123

Contents lists available at ScienceDirect

Toxicology in Vitro

journal homepage: www.elsevier.com/locate/toxinvit

Suitability of skin integrity tests for dermal absorption studies in vitro

Katharina Guth [a], Monika Schäfer-Korting [b], Eric Fabian [a], Robert Landsiedel [a,*], Ben van Ravenzwaay [a]

[a] BASF SE, Experimental Toxicology and Ecology, Carl-Bosch-Str. 38, 67056 Ludwigshafen, Germany
[b] Freie Universität Berlin, Institute for Pharmacy, Königin-Luise-Str. 2+4, 14195 Berlin, Germany

ARTICLE INFO

Article history:
Received 10 April 2014
Accepted 10 September 2014
Available online 2 October 2014

Keywords:
Dermal absorption
Integrity tests
Internal reference standard
Skin barrier function
TEWL
TEER

ABSTRACT

Skin absorption testing in vitro is a regulatory accepted alternative method (OECD Guideline 428). Different tests can be applied to evaluate the integrity of the skin samples. Here, we compared the pre- or post-run integrity tests (transepidermal electrical resistance, TEER; transepidermal water loss, TEWL; absorption of the reference compounds water, TWF, or methylene blue, BLUE) and additionally focused on co-absorption of a ^{3}H-labeled internal reference standard (ISTD) as integrity parameter. The results were correlated to absorption profiles of various test compounds. Limit values of 2 kΩ, 10 g m^{-2} h^{-1} and 4.5 $*$ 10^{-3} cm h^{-1} for the standard methods TEER, TEWL and TWF, respectively, allowed distinguishing between impaired and intact human skin samples in general. Single skin samples did, however, not, poorly and even inversely correlate with the test-compound absorption. In contrast, results with ISTD (e.g. ^{3}H-testosterone) were highly correlated to the absorption of ^{14}C-labeled test compounds. Importantly, ISTD did not influence analytics or absorption of test compounds. Therefore, ISTD, especially when adjusted to the physico-chemical properties of test compounds, is a promising concept to assess the integrity of skin samples during the whole course of absorption experiments. However, a historical control dataset is yet necessary for a potential routine application.

1. Introduction

For compounds which may get in contact with the skin, knowledge of dermal absorption is necessary to estimate the systemic exposure and perform risk assessments. For the determination of the systemic available amount of a compound in contact with the skin in vivo, in vitro and *in silico* methods are established (Schäfer and Redelmeier, 1996b). The in vitro method outlined in the OECD test guideline no. 428 is accepted by many regulatory agencies and is in accordance with the aim to reduce animal testing (OECD, 2004a, 2004b). Excised human or animal skin is mounted on a diffusion chamber, test compound is applied topically and the penetrated and permeated amount is measured in the skin sample and the underlying receptor fluid. The protocol was subject of multicenter validation studies as laid down (van de Sandt et al., 2004) and following specifications of e.g. skin type and handling (Schäfer-Korting et al., 2006, 2008). To avoid unsuitable over-prediction of the dermal absorption by the use of impaired skin preparations, the OECD guideline requires a skin integrity check. This test should ensure the exclusive use of data generated with skin with intact barrier function. In addition to a visual examination of the skin, the guideline proposes measuring the TEER (transepidermal electrical resistance), TEWL (transepidermal water loss) or the absorption characteristics of a reference compound in advance or at the end of an experiment, e.g. ^{3}H-water (TWF, transepidermal water flux), or concurrently by adding an internal reference standard (ISTD) with high specific activity to the test compound preparation, e.g. ^{3}H-sucrose (OECD, 2004a, 2004b).

Widely used standard methods in many laboratories are TWF and TEWL and TEER (Diembeck et al., 1999; Meidan and Roper, 2008). Despite intensive investigations, there is an ongoing debate about experimental performances, limit values and fields of

Abbreviations: AD, potentially absorbable dose; BLUE, methylene blue absorption; BSA, bovine serum albumin; CV, variation coefficient; DMA, dimethylamine; EDTA, ethylenediaminetetraacetic acid; ISTD, internal reference standard as well as integrity test using an internal reference standard; Kp, permeability coefficient (infinite dosing); logP, logarithmic octanol/water partition coefficient; LSC, liquid scintillation counting; maxKp, maximal permeability coefficient (finite dosing); MCPA, 2-ethyl-4-chlorophenoxyacetic acid; MCPA-2EHE, 2-methyl-4-chlorophenoxyacetyl ethylhexylester; MW, molecular weight; R^2, correlation coefficient; SD, standard deviation; SRA, specific radioactivity; TEER, transepidermal electrical resistance; TEWL, transepidermal water loss; TWF, transepidermal water flux.

* Corresponding author at: Department of Experimental Toxicology and Ecotoxicology, GV/TB, Z470, BASF SE, Ludwigshafen, Germany. Tel.: +49 621 60 56203; fax: +49 621 60 51734.

E-mail addresses: katharina.guth@basf.com (K. Guth), Monika.Schaefer-Korting@fu-berlin.de (M. Schäfer-Korting), eric.fabian@basf.com (E. Fabian), robert.landsiedel@basf.com (R. Landsiedel), bennard.ravenzwaay@basf.com (B. van Ravenzwaay).

http://dx.doi.org/10.1016/j.tiv.2014.09.007
0887-2333/

application (Brain et al., 1995; Chilcott et al., 2002; Meidan and Roper, 2008; Netzlaff et al., 2006). For example, TWF is a widely used and established marker for skin barrier function with a large historical dataset (Bronaugh et al., 1986; Meidan and Roper, 2008). Yet, the application of an infinite dose of water and therefore hydration for several hours, followed by the necessary removal and wash, may cause physical deterioration of the skin and higher permeability afterwards (Brain et al., 1995) whereas TWF measurement at the end of the experiment may lead to rejection of previously intact skin samples. Because of most similar treatment of the skin this is conceivable for TEER (Davies et al., 2004; Fasano et al., 2002), too. Also, TEWL is widely used as a marker for skin barrier function in vitro and in vivo. While avoiding physical stress to the skin (Levin and Maibach, 2005), like TEER and TWF, TEWL provides only a snapshot before or after an experiment. However, vehicle ingredients can damage the stratum corneum structure and hydration level; deterioration with time has been reported (Buist et al., 2005; Shah et al., 2008). The same holds true for the integrity test BLUE which utilizes the absorption of methylene blue as a measure for barrier functionality.

In contrast to TWF, TEER, TEWL and BLUE the integrity test ISTD supplies information of the barrier function over the whole experimental period and avoids the elongation of the test period. But the presence of an additional compound in the donor may influence the absorption characteristic of the test compound because of changes in solubility or saturation levels of the test compound and effects of the solvent on the barrier system (Barry, 1987; Dugard and Scott, 1986). Due to this influence the inertness of an ISTD must be proven. ^{3}H-sucrose and phenol red have been used as ISTD in the past, but systematic validation and provision of a sufficient dataset is still missing (Balaguer et al., 2006; Pendlington et al., 1997; Walters et al., 1997).

The purpose of the current work was to investigate the suitability of different skin integrity tests to differentiate impaired and intact human skin. Based on the absorption results of four test compounds (testosterone, caffeine, 2-ethyl-4-chlorophenoxyacetic acid (MCPA) and 2-methyl-4-chlorophenoxyacetyl ethylhexylester (MCPA-EHE)) through human and generally more permeable reconstructed human skin (StrataTest®), the common limit values for the standard integrity methods TEER, TWF and TEWL were assessed. Additionally, results of five skin integrity tests (TEER, TWF, TEWL, ISTD and BLUE) were correlated to absorption results derived with human skin or reconstructed human skin to evaluate their ability to explain minor differences in barrier function. Full-thickness and dermatomed human skin samples were applied to check for a possible effect of the skin preparation. Due to a lower donor dependency, rat skin was used in addition and chosen for a special experiment in which skin samples were systematically damaged to different grades before use. As model ISTD ^{3}H-testosterone was chosen. It was applied in parallel to test compound ^{14}C-MCPA. For human skin experiments two further well-investigated reference compounds with different physico-chemical properties were applied as ISTDs (^{3}H-caffeine and ^{3}H-mannitol) (OECD, 2004a; Peck et al., 1995; Schäfer-Korting et al., 2008; van de Sandt et al., 2004) to get an insight on the effect of ISTD selection. Additional experiments were conducted to check for effects of the present ISTDs on the analytics and absorption characteristics of the test compound.

2. Materials and methods

2.1. Chemicals and reagents

MCPA-2EHE, MCPA, dimethylamine (DMA; 60%), silicone antifoam emulsion (SRE) and ethylenediaminetetraacetic acid (EDTA) were provided by AH Marks and Co, Wyke, Bradford, Great Britain. Testosterone, caffeine, ethanol and methylene blue were purchased from Sigma Aldrich, St. Louis, MO, USA, bovine serum albumin (BSA) was from Roche, Basel, Switzerland, Texapon® N70 from Cognis, Düsseldorf, Germany, NaCl from Merck, Darmstadt, Germany and Soluene 350® and scintillation cocktail Hionic Fluor™ from Perkin–Elmer, Boston, MA, USA. Radiolabeled compounds (radiochemical purity >97%) were supplied by American Radiolabeled Chemicals, St. Louis, MO, USA (^{3}H-caffeine with 2.22 TBq mmol^{-1}), Perkin–Elmer (^{14}C- and ^{3}H-testosterone with 2.1 GBq mmol^{-1} and 6.3 TBq mmol^{-1}, respectively, ^{14}C-caffeine with 1.89 GBq mmol^{-1} and ^{3}H-mannitol with 455.6 GBq mmol^{-1}, ^{3}H-Water with 37 MBq ml^{-1}) or by AH Marks and Co (^{14}C-MCPA with 1.88 GBq mmol^{-1} and ^{14}C-MCPA-2EHE with 1.02 GBq mmol^{-1}). The radioactive isotopes are generally located at stable positions of the molecule: ^{14}C in the A ring of the steroid testosterone, in phenyl ring of MCPA and MCPA-EHE and in the methyl group at N-1 of caffeine; ^{3}H generally at non-acidic groups (testosterone at positions C-1, C-2, C-6, C-7, C-16 and C-17, mannitol at C-1 and caffeine in methyl group at N-1).

2.2. Skin preparations

Split-thickness (450 ± 100 μm) and full-thickness (1000 ± 200 μm) female human skin samples from abdominal surgery were purchased from Biopredic, France. Rat skin was excised from the back of eight-week-old female Crl:WI (Han) rats (Charles River, Germany) after sedation with isoflurane and exsanguination. Split-thickness skin (450 ± 100 μm) was generated with a Dermatome GA 643 (Aesculap, Germany) after hair trimming. For a special investigation various grades of barrier impairment were induced by stressing excised rat skin with chemical or mechanical treatment in advance of experiments using ^{14}C-MCPA as the test substance. Such pretreatment scenarios comprises combinations of water application or application of MCPA formulation (see Table 1) with or without MCPA and one or three washing steps with cotton swabs and 0.7% aqueous Texapon® N70 solution over three consecutive days. The individual treatments are given in Table 2. Experiments 1–3 comprise the 'undamaged' skin and experiments 4–9 the 'damaged' skin. StrataTest® (100–115 μm) purchased from Stratatech Corporation, USA, is a reconstructed human skin model which was added in the current setup as a human skin system with generally lower barrier functionality.

2.3. In vitro dermal absorption study

All studies were conducted following the OECD-Guideline 428 and the corresponding technical guidance document 28 (OECD, 2004a, 2004b). Five skin samples per run, derived from at least two different donors, were mounted on Franz type diffusion cells with a surface area of 1 cm^{2} and receptor volume of 4 ml (Laboratory Glass Apparatus Inc., USA). The water jacket around the receptor compartment was maintained using a water thermostat pump (Thermo Haake, Germany) at a temperature of 32 °C. A finite dose was applied to the surface of the skin under occlusive (Parafilm "M"®, Pechiney Plastic Packaging, USA) or semi-occlusive (Fixomull®, BSN medical, Germany) conditions. The receptor fluid was chosen to provide an adequate solubility of the test compound – at least 10 times higher than the maximal achievable concentration (see Tables 1 and 3). After exposure for 6 or 24 h the compound was washed off with cotton swabs and washing fluid. During the experimental period, samples were taken from the stirred (magnetic stirrers, Variomag Telemodul 20C/40C, H + P Labortechnik, Germany) receptor fluid at distinct time points and replaced with fresh receptor fluid by a fraction collector (222 L, Abimed, Germany) and a multi-channel peristaltic pump (MC 360, Ismatec, Germany). At the end of the run each diffusion cell was dismantled and all parts

Table 1
Test substances (including physico-chemical properties) and experimental conditions (^{14}C-testosterone and ^{14}C-caffeine testing was adapted to literature protocol (van de Sandt et al., 2004)).

Test substance	LogP	MW (g mol^{-1})	Formulation	SRA (MBq ml^{-1})	Applied volume (µl cm^{-2})	Applied dose (µg cm^{-2})	ISTD	Exposure time (h)	Washing fluid	Receptor fluid	Solubility in receptor (g l^{-1})	Occlusion condition
^{14}C-MCPA-EHE	6.8[a]	312.8	16 mg ml^{-1} emulsion in water	3.7	10	160	^{3}H-Testosterone	6	0.7% Texapon® N70 solution	Water	0.40 (32 °C)	Semi-Occlusive
^{14}C-testosterone	3.32[b]	288.4	4 mg ml^{-1} solution in ethanol/water 1/1 (v/v)	1	25	100	^{3}H-Mannitol, ^{3}H-Caffeine	24	Ethanol/water (1/1, v/v)	5% BSA in water	0.49 (32 °C)	Occlusive
^{14}C-caffeine	−0.07[c]	194.2	4 mg ml^{-1} solution in ethanol/water 1/1 (v/v)	1	25	100	^{3}H-Mannitol, ^{3}H-Testosterone	24	Ethanol/water (1/1, v/v)	Water	20 (20 °C)[c]	Occlusive
^{14}C-MCPA	−0.71 (pH 7) 2.75 (pH 1)[a]	200.6	9 mg ml^{-1} DMA salt solution with EDTA and silicone in water, pH 7	3.7	10	90	^{3}H-Testosterone	6	0.7% Texapon® N70 solution	Water	293.9 (25 °C, pH7)[a]	Semi-occlusive

logP: logarithmic octanol/water partition coefficient; MW: molecular weight; SRA: specific radioactivity; EDTA: ethylenediaminetetraacetic acid; DMA: dimethylamine.
[a] Britisch Crop Protection Council (2011).
[b] Yalkowsky et al. (1983).
[c] Merck (2006).

were processed for balancing. Two to six tape strips (Crystal Clear Tape 600, Scotch, France) were used to remove the upper stratum corneum from the skin samples. The tapes with stratum corneum and the remaining skin were digested with Soluene 350®, lasting a minimum of 24 h; cotton swabs as well as the class devices were extracted with ethanol or water – depending on the solubility of the test compounds. All samples were diluted with LSC-Cocktail and measured by Liquid Scintillation Counting (LSC; TriCab 2800TR, Perkin-Elmer, USA; linear range up to 1,000,000 dpm). Absolute and percentage amounts in receptor fluid, skin, tape strips and washing fluids were calculated as well as the total recovery. Only a recovery of 100 ± 10% was assumed to be valid for mean calculations. The sum of content in receptor fluid (including receptor chamber washings) and skin was defined as the potentially absorbable dose (AD); if applicable also the amount recovered from the underlying membrane of the reconstructed human skin was assigned to AD. The cumulative absorbed amount was plotted against time. The steepest slope – the maximal absorption rate in µg cm^{-2} h^{-1} – divided by the applied concentration in µg cm^{-3} provides the maximal permeability constant maxKp in cm h^{-1}. The intercept of the elongated steepest slope line with the *x*-axis represents the lag time (h). Test compound dependent experimental conditions as well as logP and molecular weight are listed in Table 1. All four test compounds were applied to full-thickness and split-thickness human skin, ^{14}C-testosterone, ^{14}C-caffeine and ^{14}C-MCPA were also applied to rat skin and to reconstructed human skin. Unintentionally damaged skin samples were left in the set up and examined along with the intact samples. Intentionally impaired rat skin samples were used for ^{14}C-MCPA experiments.

2.4. Integrity tests

Besides a visual check at least two of the five following integrity tests were conducted in each experiment, the skin being mounted on the Franz cell. TEER, TEWL and TWF were performed in advance, ISTD concurrently and BLUE at the end of the run.

2.4.1. TEER

To measure the transepidermal electrical resistance to an alternating current (impedance), the receptor and donor compartment of the diffusion cell were filled with physiological saline (0.9% aqueous NaCl solution). Electrodes were immersed in each compartment and the impedance was measured via a LCR bridge (LCR400, Thurbly Thandar Instruments, Great Britain) at a frequency of 1 kHz. The standard limit value was 1 kΩ. This value is based on internal experience and experiments to distinguish native and punched human skin samples. A lab-specific limit value is necessary due to limited transferability: The measured resistance is dependent on the device, applied frequency, resulting current, ionic strength of the solution as well as the surface area of the skin sample (Fasano et al., 2002).

2.4.2. TEWL

The transepidermal water loss was measured after minimal 1 h of equilibration and drying of the skin surface. The moisture on the skin surface originating from rehydration of the frozen skin samples or from TEER measurement needs to be evaporated to measure exclusively the water loss through the skin sample. With a VapoMeter (Delfin Technologies Ltd., Finland) the TEWL was determined under closed chamber conditions (Imhof et al., 2009). For this end the donor compartment of the diffusion cell was covered completely with the VapoMeter. The standard limit of 10 g m^{-2} h^{-1} was used (Schäfer and Redelmeier, 1996b).

Table 2
Pretreatment scenarios to obtain various grades of barrier impairment.

Experiment no.[a]	Total experimental time (h)	3 Applications on 3 consecutive days	Washing procedure (mechanical stress)
1	24	None	None
2	96	None	None
3	96	Water	None
4	96	Water	8 h after the last application
5	96	Water	8 h after each application
6	96	Blank formulation[b]	None
7	96	Blank formulation[b]	8 h after each application
8	96	Cold formulation[c]	8 h after the last application
9	96	Cold formulation[c]	8 h after each application

[a] For each experiment 5 replicates were used.
[b] Blank formulation refers to MCPA application solution without MCPA.
[c] Cold formulation refers to MCPA application solution with only unlabeled MCPA.

2.4.3. TWF

To determine the absorption characteristics of tritiated, ^{3}H-labeled, water, the receptor compartment was filled with physiological saline. An infinite dose (300 µl cm^{-2}) with a specific radioactivity of 123 kBq ml^{-1} was applied to the surface of the skin. At distinct time points (0.5, 1, 2, 3, 4 and 5 h) receptor fluid was collected using a syringe. After the last sampling the skin was thoroughly washed with distilled water and cotton swabs. Receptor fluid was diluted with scintillation cocktail, measured by LSC and data were used to calculate the permeability constant (Kp) as described in Section 2.3. A generally accepted limit value of $2.5 * 10^{-3}$ cm h^{-1} was used (Bronaugh et al., 1986). Using TWF as a pre-test, the radioactivity needs to be removed from the system before application of the test compound. Therefore, the receptor fluid was changed several times until the activity in a receptor fluid aliquot declined to 50 dpm (0.8 Bq).

2.4.4. ISTD

A ^{3}H-labeled internal reference standard was added to the ^{14}C-labeled test compound formulation and applied to the skin (see Tables 1 and 3). The concentration was determined by the specific radioactivity of the ^{3}H-ISTD which was chosen to be equal to the specific radioactivity of the ^{14}C-labelled test compound (Table 1). In all samples ^{3}H-activity was measured along with the ^{14}C-activity by LSC. Absorption characteristics (AD and maxKp) were determined analogously, as described in Section 2.3.

2.4.5. BLUE

Following the final washing procedure at the end of the absorption experiment, 250 µl of methylene blue, 0.025% aqueous solution, was applied on top of the skin for 0.5 h and washed off with 0.7% aqueous Texapon® N70 solution. The receptor fluid was tested for permeated dye using a photometer operating at 661 nm. The concentration in the receptor fluid was determined via a calibration curve. Any staining of the epidermis was reported before digestion and processing for LSC measurements.

Table 3
Physico-chemical properties of internal standards.

ISTD	LogP	MW (g mol^{-1})	Solubility in water (g l^{-1})
^{3}H-Mannitol	−3.1[a]	182.17	182.2 (20 °C)[d]
^{3}H-testosterone	3.32[b]	288.4	0.02 (25 °C)[b]
^{3}H-caffeine	−0.07[c]	194.2	20 (20 °C)[c]

logP: logarithmic octanol/water partition coefficient; MW: molecular weight.
[a] Leo et al. (1971).
[b] Yalkowsky et al. (1983).
[c] Merck (2006).
[d] Sigma-Aldrich (2007).

2.5. Assessment of integrity tests

Two methods were used to assess the suitability of the different skin integrity tests. Firstly, the ability for binary differentiation of human skin samples was evaluated for the three standard tests TEER, TEWL and TWF. Therefore, we differentiated valid and invalid excised or reconstructed human skin samples according to the standard limit values for human skin set 1 kΩ, 10 g m^{-2} h^{-1} and $2.5 * 10^{-3}$ cm h^{-1} and for TEER, TEWL and TWF, respectively. In addition one further limit value was used for each test. Based on the outcome, these were more liberal for TEWL (13 g m^{-2} h^{-1}) and TWF ($4.5 * 10^{-3}$ cm h^{-1}), yet more strict for TEER (2 kΩ). The minimum (min), maximum (max) and mean absorption results (maxKp and AD) were calculated separately for the defined valid and invalid groups. Furthermore we plotted the single cell results for the defined valid and invalid skin samples.

Next, the ability of all five integrity tests (TEER, TEWL, TWF, ISTD and BLUE) to detect and explain minor differences in barrier function was investigated by correlation analyses. For this task, rat skin was included, basically, to make use of the in theory lower donor variability of rat skin for the special investigation in which rat skin was systematically damaged to various grades. For the correlation analyses we grouped all experiments using the same test compound (caffeine, testosterone, MCPA or MCPA-EHE) and barrier system (human, rat or reconstructed human skin) together. Groups with at least 10 single data points were used for linear regression analysis of integrity test results (independent variable x) against absorption results (AD and maxKp, dependent variable y). All data points were included independent of valid or invalid classification. Slopes and correlation coefficients (R^2) were reported for evaluation. Min, max and mean values were calculated for each integrity test, but only R^2 from correlations with the correct algebraic sign were used.

To assess the variability of the methods and the effect of the human donor, overall, inter- and intra-donor variabilities were calculated for the different methods. Overall variability is given as the variation coefficient (CV, often referred to as the relative standard deviation (SD)) of all skin samples used, inter-donor variability is given as CV calculated with the mean values for each donor and intra-donor variability which corresponds to the method variability is given as the pooled, average, CV for each donor weighted by the number of replicates. If from one human donor both, full-thickness and dermatomed skin, was used, the underlying means and variabilities were calculated separately. For ISTD and the general in vitro dermal absorption method, only the pooled CV could be calculated due to the various kinds of ISTDs and test compounds used. Underlying means were calculated separately for each ISTD or test compound.

2.6. Verification of ISTD approach

Since energy spectra of ^{14}C and ^{3}H overlap, a LSC method was used that compensates for the influence of the other isotope. The actual independence of ^{3}H- and ^{14}C-analytics was checked by measuring ^{14}C-testosterone standards in the presence and absence of ^{3}H-testosterone and vice versa. High and low levels of matrix isotope were used (^{3}H: 242 and 12667 Bq, ^{14}C: 587 and 1288 Bq, on average). Linear regressions were calculated to evaluate a possible influence. Additionally, the independence of test compound absorption from the presence of an internal reference standard was investigated: The absorption characteristics of ^{14}C-MCPA and ^{14}C-caffeine in presence and absence of ^{3}H-testosterone as well as ^{14}C-testosterone in presence and absence of ^{3}H-caffeine were examined in the identical experimental set-up. Mean and SDs were calculated for each group. Student's *t*-test was performed with Microsoft Office Excel 2003. Significance (*) was set at $p \leqslant 0.05$, high significance (**) at $p \leqslant 0.01$ is indicated, too.

3. Results

3.1. Assessment of integrity tests

Evaluation of binary differentiation of human skin samples by the standard integrity tests TEER, TEWL and TWF is based on the results given in Tables 4–6. Shown are mean, min and max values for the absorption of four test compounds through excised or reconstructed human skin samples separately for valid and invalid skin samples. The integrity or validity of the skin preparations were judged by the standard limit values for human skin of TEER, TEWL and TWF. TEWL and TWF lead to more skin preparations classified as 'invalid' than TEER. In fact, there was almost no need for exclusion with the cut-off level set 1 kΩ. Even the reconstructed human skin samples providing generally a minor barrier function (Schäfer-Korting et al., 2008) and showing apparent higher absorption values for the test compounds, were classified as valid. In general, based on TEWL and TWF the mean absorption values (Kp and AD) for ^{14}C-caffeine, ^{14}C-testosterone and ^{14}C-MCPA were higher in invalid skin preparations compared to the valid skin samples. However, the min–max ranges of absorption values in valid and invalid skin preparations overlapped; this is when high max values for valid and low min values for invalid skin samples were present. The individual maxKp values for the single human skin preparations are visualized in Fig. 1. In this example, classification in valid (open symbols) and invalid (filled symbols) skin samples is based on TEWL, cut-off 10 g m^{-2} h^{-1}. As to be expected from the well-known higher permeability of reconstructed epidermis or reconstructed full-thickness skin compared to human skin (Ackermann et al., 2010; Schäfer-Korting et al., 2008), invalid data are predominantly obtained when testing in the constructs (shown as triangles in Fig. 1). If the constructs were analogously classified as principally invalid by TWF could not be investigated in this study. Due to the observed fragility of the tissue, including the sensitivity to washing steps being part of this pre-test, TWF was waived for the constructs.

Next we tested more liberal cut-off levels. Changing the TEWL limit from 10 to 13 g m^{-2} h^{-1} did not change the distribution significantly for ^{14}C-caffeine and ^{14}C-testosterone (Table 5). For MCPA the valid results increased clearly when applying the higher limit value and the range of valid data even included the range of invalid in full. This effect is mainly due to the inclusion of absorption results obtained with six reconstructed human skin samples which were obviously higher, but based on TEWL cut-off limit 13 g m^{-2} h^{-1} classified as valid. The very slow penetrating test compound ^{14}C-MCPA-2EHE showed no clear difference of absorption values in valid and invalid skin samples. This was observed with all integrity tests (Tables 4–6). Mean, min and max values did not differ significantly for the two different limit values of TWF (Table 6). However, the stricter limit value for TEER (2 kΩ) led to a different distribution (Table 4). Only 2 of 90 skin samples were classified as invalid with 1 kΩ as the limit, but 28 with 2 kΩ. Applying the limit value 2 kΩ, the majority of the reconstructed human skin samples with higher absorption results for the test compounds were classified as invalid (23 out of 30). In contrast, five excised human skin samples were classified as invalid despite of absorption data in reasonable ranges. Analog to TEWL, differentiation with TEER (limit: 2 kΩ) and TWF resulted in obvious higher absorption means for invalid skin samples than for valid skin samples as well as in significant overlapping of results.

In a second step linear regression analyses for the absorption values (AD, maxKp, dependent variable *y*) and integrity test results (independent variable *x*) were used to check whether integrity tests TEER, TEWL, TWF, ISTD and BLUE are able to display minor barrier differences between skin samples continuously. Besides human skin, rat skin was included in these analyses. Table 7 shows mean, min and max values of slopes and R^2 derived from analysis for the different experimental groups. One group covers experiments using

Table 4
Range of absorption (minimum–maximum) of the test compounds for valid and invalid (excised and reconstructed human) skin samples differentiated by TEER at two limit values. Mean values are shown in brackets. maxKp (maximal permeability constant) is given in ($*10^{-5}$ cm h^{-1}) and AD (absorbed dose) in percentages. *n* is the number of skin samples used for calculations.

Limit value classification	2 kΩ		1 kΩ	
	Valid	Invalid	Valid	Invalid
Testosterone				
n	22	8	30	0
maxKp	12–283 (78)	48–805 (296)	12–805 (136)	–
AD	6–81 (29)	18–84 (52)	6–84 (35)	–
Caffeine				
n	20	10	30	0
maxKp	25–994 (108)	54–1647 (901)	25–1647 (374)	–
AD	12–97 (36)	54–103 (94)	12–103 (55)	–
MCPA				
n	11	9	18	2
maxKp	4–405 (78)	18–1004 (461)	4–585 (183)	715–1004 (860)
AD	9–86 (28)	13–96 (82)	9–96 (48)	94–96 (95)
MCPA-EHE				
n	9	1	10	0
maxKp	0.4–2.0 (0.9)	0.4–0.4 (0.4)	0.4–2.0 (0.9)	
AD	1.0–9.6 (4.6)	8.9–8.9 (8.9)	1.0–9.6 (5.1)	–

Table 5
Range of absorption (minimum–maximum) of the test compounds for valid and invalid (excised and reconstructed human) skin samples differentiated by TEWL at two limit values. Mean values are shown in brackets. maxKp (maximal permeability constant) is given in ($*10^{-5}$ cm h^{-1}) and AD (absorbed dose) in percentages. *n* is the number of skin samples used for calculations.

Limit value classification	10 g m^{-2} h^{-1}		13 g m^{-2} h^{-1}	
	Valid	Invalid	Valid	Invalid
Testosterone				
n	19	11	19	11
maxKp	12–138 (47)	20–805 (290)	12–138 (47)	20–805 (290)
AD	6–44 (19)	32–84 (63)	6–44 (19)	32–84 (63)
Caffeine				
n	19	11	20	10
maxKp	25–143 (64)	54–1647 (909)	25–143 (63)	607–1647 (995)
AD	12–68 (32)	54–103 (95)	12–68 (33)	95–103 (99)
MCPA				
n	10	10	17	3
maxKp	4–365 (46)	15–1004 (455)	4–1004 (192)	442–715 (579)
AD	9–84 (21)	28–96 (84)	9–96 (45)	95–96 (95)
MCPA-EHE				
n	3	7	7	3
maxKp	0.6–1.4 (0.9)	0.4–2.0 (0.9)	0.4–2.0 (0.8)	0.9–1.0 (1.0)
AD	5.1–9.6 (6.7)	1.0–8.9 (4.4)	1.0–9.6 (5.2)	2.6–8.7 (4.9)

Table 6
Range of absorption (minimum–maximum) of the test compounds for valid and invalid excised human skin samples differentiated by TWF at two limit values. Mean values are shown in brackets. maxKp (maximal permeability constant) is given in ($*10^{-5}$ cm h^{-1}) and AD (absorbed dose) in percentages. *n* is the number of skin samples used for calculations.

Limit value classification	$2.5 * 10^{-3}$ cm h^{-1}		$4.5 * 10^{-3}$ cm h^{-1}	
	Valid	Invalid	Valid	Invalid
Testosterone				
n	8	2	9	1
maxKp	12–37 (26)	20–33 (27)	12–37 (27)	20–20 (20)
AD	11–24 (17)	10–44 (27)	10–24 (16)	44–44 (44)
Caffeine				
n	8	2	10	0
maxKp	30–143 (68)	77–130 (103)	30–143 (75)	–
AD	30–68 (46)	51–53 (52)	30–68 (47)	–
MCPA				
n	4	6	8	2
maxKp	4–11 (6)	8–18 (14)	4–18 (10)	14–15 (14)
AD	9–15 (12)	12–28 (17)	9–28 (15)	15–17 (16)
MCPA-EHE				
n	3	7	8	2
maxKp	0.7–1.4 (1.0)	0.4–2.0 (0.8)	0.4–2.0 (0.9)	0.4–0.9 (0.7)
AD	2.6–9.6 (5.8)	1.0–8.9 (4.8)	1.0–9.6 (4.8)	3.3–8.9 (6.1)

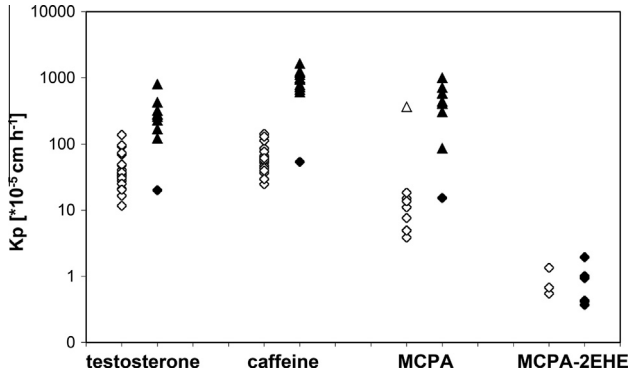

Fig. 1. Maximal permeability constants (maxKp) of four ^{14}C-labeled test compounds. Values are sorted by integrity class due to TEWL measurements (cut-off-value: 10 g m^{-2} h^{-1}). Valid human skin samples are shown as open symbols, invalid ones as filled symbols. Excised human skin is shown in diamonds, reconstructed human skin in triangles.

one defined combination of test compound (testosterone, caffeine, MCPA or MCPA-2EHE) and skin preparation (excised human skin, reconstructed human skin or excised rat skin). The correlations varied over a wide range for all five methods, four test compounds and three skin preparation types. The best correlations in average (R^2: 0.484) and maximal (R^2: 0.911) were achieved with the ISTD. Partially good correlations were observed for TEWL: the maximal R^2 of 0.790 was achieved with test compound ^{14}C-testosterone applied to reconstructed human skin. Even inverse correlations were occasionally obtained with TEWL, TEER, TWF and BLUE but not for ISTD. The dataset of the special investigation comprising all experiments with ^{14}C-MCPA applied to undamaged and gradually damaged rat skin covers a wide range of absorption (AD 6–100%) and absorption rates (Marzulli-Class: slow to very fast) (Marzulli and Brown, 1969). The results of individual skin preparations correlated well with the ISTD results (R^2 of 0.859 and 0.911, respectively). The plot for AD is shown in Fig. 2. In contrast, weak correlations exist between TEWL and AD (R^2: 0.598) and maxKp (R^2: 0.451) as well as TEER and AD (R^2: 0.386) and maxKp (R^2: 0.479). The quality of fit was not related

Table 7
Overview of correlations between penetration results (absorbed dose, AD and maximal permeability constant, maxKp) of test compound (TC) and outcome of integrity tests. Correlations were calculated for each homogeneous experiment – each TC (testosterone, caffeine, MCPA or MCPA-2EHE) and barrier system (excised human skin, reconstructed human skin or rat skin) combination – alone. Shown are the maximal (max) and minimal (min) slopes and correlation coefficients (R^2) of all experimental groups (n). Mean values for R^2 are only calculated from experiments with correct positive or negative correlation. Regarding theory the slope is expected to be positive for ISTD, TWF, TEWL, BLUE and negative for TEER.

	TEER		TEWL		TWF		^{3}H-ISTD				BLUE	
							AD		maxKp			
	Slope	R^2	Slope	R^2	Slope	R^2	Slope	R^2	Slope	R^2	Slope	R^2
maxKp (TC)												
n	8	7	9	7	4	2			6	6	4	3
Mean	−68.48	0.237	2.25	0.230	0.07	0.259			1.24	0.475	−0.04	0.243
Min	−359.2	0.017	−26.37	0.006	−0.005	0.212			0.05	0.139	−0.21	0.154
Max	2.93	0.691	27.62	0.757	0.27	0.307			4.10	0.911	0.02	0.391
AD (TC)												
n	8	3	9	7	4	3	6	6			4	1
Mean	0.39	0.360	1.33	0.377	0.02	0.257	0.73	0.484			0.06	0.433
Min	−8.70	0.036	−0.82	0.030	−0.005	0.064	0.13	0.142			−0.01	0.433
Max	7.64	0.658	4.59	0.790	0.05	0.628	1.13	0.859			0.26	0.433

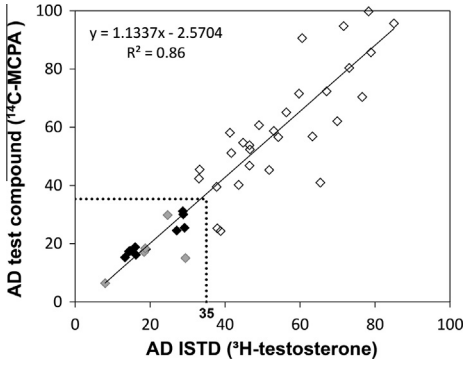

Fig. 2. Correlation of percent potentially absorbable dose (AD) of the test compound ^{14}C-MCPA and AD of the internal reference standard ^{3}H-testosterone with rat skin. The large absorption range was achieved by chemical and mechanical pretreatment of the rat skin samples as described in Table 2. Open diamonds represent damaged skin samples (experiments 4–9) and filled diamonds represent undamaged skin preparations (experiments 1 and 2: no pretreatment in black; experiment 3: pretreatment with water without mechanical washing in gray). The 35% line represents the provisional cut-off value.

to the skin preparations used, meaning that good, moderate and poor correlations were obtained with excised human skin, reconstructed human skin as well as excised rat skin.

Finally, to assess and compare the variabilities of the integrity tests (TEER, TEWL, TWF and BLUE), the overall, inter-donor and intra-donor or method variabilities were calculated. The results are given in Table 8. For instance, TEER resulted in CVs of 65%, 45% and 43%, respectively. Furthermore the method variability of the in vitro dermal absorption experiments (45% and 33% regarding AD and maxKp, respectively) and the ISTD (30% and 38% regarding AD and maxKp, respectively) are given.

3.2. Verification of ISTD approach

The independency of ^{3}H- and ^{14}C-analytics was proven by the quantification of ^{14}C-testosterone standards in presence of ^{3}H-testosterone at two dose levels in comparison to ^{14}C-testosterone standards without ^{3}H in the matrix (Fig. 3). The R^2 was 0.9991 and the slope 1.0077. No general influencing effects were apparent. This holds also true with ^{3}H-testosterone levels measured without the addition of the ^{14}C-labelled steroid and following the addition of this label at a high and low amount. Then the R^2 was 0.9998 and the slope 1.0008 (data not shown). In the very low Bq range (<200 Bq) of ^{3}H-testosterone the presence of ^{14}C increased the variability of ^{3}H-testosterone data.

To assess the co-absorption of test compound and internal reference standard, Table 9 lists absorption characteristics for three ^{14}C-labeled test compounds in absence and presence of a ^{3}H-labeled ISTD. Except for a significantly different lag time for ^{14}C-testosterone with and without ^{3}H-caffeine all endpoints of dermal absorption were close and the ISTD did not influence the absorption of the test compound.

Table 8
Variability of in vitro dermal absorption experiments and five integrity tests based on results with excised human skin from different donors.

	^{14}C-TC AD[a]	^{14}C-TC maxKp[a]	TEER	TEWL	TWF	^{3}H-ISTD AD[a]	^{3}H-ISTD maxKp[a]	BLUE[b]
Overall variability:[c] CV	–	–	64%	55%	61%	–	–	56%
(*n*)			(65)	(65)	(40)			(37)
Inter-donor variability:[d] CV	–	–	45%	48%	32%	–	–	44%
(*D*;*N*)			(11;16)	(11;16)	(5;10)			(5;10)
Intra-donor variability:[d] pooled CV	45%	33%	43%	57%	53%	30%	38%	43%
(*D*;*N*)	(11;20)	(11;20)	(11;16)	(11;16)	(5;10)	(9;18)	(9;18)	(5;10)

[a] Only the pooled variation coefficient (CV) was calculable for the test compound (TC) and ISTD due to the various TCs and ISTDs used.
[b] Results from three dermatomed skin samples with excessive high values for BLUE indicating damaged skin samples were excluded from the calculations.
[c] Overall variability was calculated with the mean and standard deviation (SD) for n replicates of humen skin samples and is given as CV in percentages and corresponds to the relative SD.
[d] Inter-donor variability was calculated with the mean values for each single donor (*D*: number of donors); intra-donor variability or pooled CV, corresponding to the method variability, was calculated with the CVs of each single donor weighted by the number of replicates; for both means and CVs were calculated separately for dermatomed and full-thickness skin, whereby both resulted in generally comparable values (data not shown) – and separately for each ISTD or TC for the respective calculations. The resulting number of independent values is given as *N*.

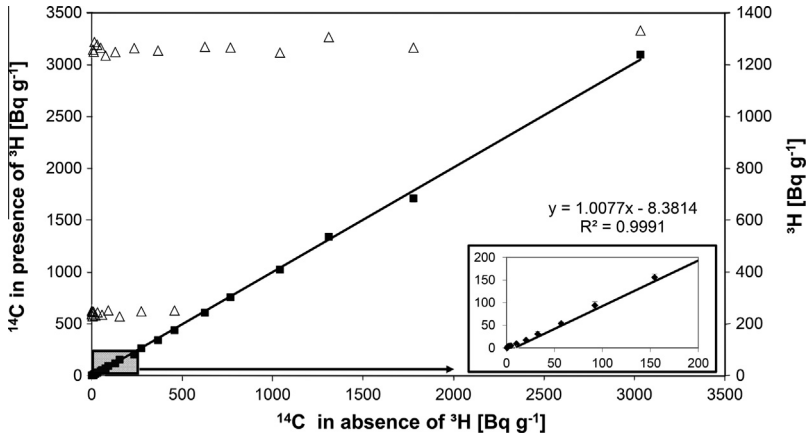

Fig. 3. Correlation of radioactive concentration of ^{14}C-testosterone in absence and presence of ^{3}H-testosterone as determined by LSC. ^{14}C-concentrations are shown in filled squares. Added high and low dose ^{3}H-concentrations are shown in open triangles. Insertion shows magnified low concentration rage (<200 Bq g^{-1}).

Table 9
Absorption results of 3 different test compounds (TC) in absence (–) and presence (+) of a tritium labeled internal reference standard (^{3}H-ISTD) through excised human skin. Statistics: students *t*-test.

^{14}C-TC	^{3}H-ISTD	Skin (%)	Receptor (%)	maxKp ($*10^{-5}$ cm h^{-1})	Lag time (h)
Caffeine	–	0.5 ± 0.2	19.6 ± 6.0	62.5 ± 13.8	3.0 ± 1.1
	+ Testosterone	0.6 ± 0.3	16.5 ± 3.7	40.9 ± 16.6	4.0 ± 0.9
Testosterone	–	1.7 ± 1.3	24.5 ± 10.6	69.3 ± 25.1	4.4 ± 1.5
	+ Caffeine	0.4 ± 0.2	17.3 ± 14.8	61.8 ± 47.0	2.4 ± 0.4*
MCPA	–	2.5 ± 2.4	9.2 ± 3.6	15.6 ± 14.6	1.1 ± 0.9
	+ Testosterone	4.3 ± 0.9	12.6 ± 1.2	21.4 ± 10.4	0.7 ± 0.4

* $p < 0.05$.

4. Discussion

TEER, TEWL and TWF are widely used skin integrity tests, each with a large historical dataset (Bronaugh et al., 1986; Davies et al., 2004; Diembeck et al., 1999; Elkeeb et al., 2010; Meidan and Roper, 2008). Nevertheless there are still discussions about the experimental performances, limit values and fields of application (Chilcott et al., 2002; Meidan and Roper, 2008; Netzlaff et al., 2006). Impairment of the skin barrier identified by these methods is expected to allow excessive penetration and permeation of the test compound and therefore yield invalid results. Usually cut-off values are used to distinguish impaired from intact skin preparations with no intermediate stages: a skin sample is either valid or invalid. This is helpful in case of a pre-test that rejects inappropriate samples for absorption testing. Compromising the barrier of a skin preparation is, however, a continuous phenomenon ranging from intact to increasingly more permeable barriers and can occur before the study started, during the test compound application (sometimes even caused by the compound, more often by the vehicle) or after the application (e.g. during washing of the skin). To obtain a complete picture of the barrier integrity, an advanced integrity test would detect the continuum of barrier impairments and barrier defects may correlate with the absorption of the test compound through the very skin preparation.

4.1. Evaluation of the three standard methods TEER, TEWL and TWF

To address the binary differentiation of human skin samples into valid and invalid, we compared the absorption results (AD and maxKp) of four test compounds (caffeine, testosterone, MCPA and MCPA-EHE) applied to excised or reconstructed human skin. The results were grouped by integrity test classification (valid/invalid) according to the three standard tests TEER, TEWL and TWF operated at two cut-off levels. Mean values for valid human skin samples sorted by TEWL or TWF were generally higher than means for invalid skin samples. The valid absorption results for ^{14}C-caffeine and ^{14}C-testosterone (Tables 5 and 6) were in good accordance with absorption studies for (^{14}C-) caffeine $56 \pm 36 * 10^{-5}$ cm h^{-1} (maxKp) and 30 ± 14% (AD) and (^{14}C-) testosterone $41 \pm 48 * 10^{-5}$ cm h^{-1} (maxKp) and 20 ± 15% (AD) through human skin (van de Sandt et al., 2004). 29 out of 30 reconstructed human skin samples were identified as invalid by TEWL measurements, which was in accordance to obviously higher absorption values in comparison to excised human skin samples. Generally higher absorption through reconstructed human epidermis and reconstructed human full-thickness skin in comparison to native human skin and pig skin was reported previously (Ackermann et al., 2010; Schäfer-Korting et al., 2008). The outlined observations confirm a meaningful differentiation of skin samples using integrity tests TEWL or TWF. However, some single skin samples with average permeability were identified as invalid and a few as valid which presented obvious too high maxKp and AD values. Deterioration of the skin during the experiment just due to time or caused by detergent and manipulation during washing procedure can be reasons for false valid classifications (Buist et al., 2005). Such effects can only be considered and evaluated by concurrent or post-experimental integrity tests. Interestingly the EFSA "Guidance on Dermal Absorption" recommends to avoid post experimental integrity tests (EFSA Panel on Plant Protection Products and their Residues, 2012). Prevention of inappropriate skin rejection due to compound related barrier damages could be reasons for this recommendation. However, diminished barrier function of single skin preparations after an experiment may provide valuable information, for instance, hints for an inappropriate over-prediction of dermal absorption. In contrast, if all skin preparations are affected, showing impaired barrier functionality, a destructive effect of the applied formulation is implied.

Skin samples could be falsely classified as 'invalid' if limit values are set to strict. To address this we also applied besides our standard TEWL limit of 10 g m^{-2} h^{-1} and the well-established TWF limit value of $2.5 * 10^{-3}$ cm h^{-1} (Bronaugh et al., 1986) for human skin, higher values of 13 g m^{-2} h^{-1} and $4.5 * 10^{-3}$ cm h^{-1} (Meidan and Roper, 2008). For TEWL it makes no significant difference: with both restrictions the valid mean for ^{14}C-caffeine and ^{14}C-testosterone was in accordance with reference values (van de Sandt et al., 2004); but inclusion of several high maxKp values and ADs for ^{14}C-MCPA – due to the less strict limit value – led to obviously higher mean values for skin that was classified as valid. To avoid inclusion of such apparent over-predicted values for mean calculations, the stricter limit value for TEWL or a combination of different integrity tests is advisable. Both limit values for TWF led to similar valid and invalid values. With both limits many skin samples were considered as invalid in contrast to absorption results in reasonable ranges and TEWL classifications. To avoid unnecessary rejection of skin samples by this sensitive method, the higher limit value is recommendable.

A large number of the reconstructed human skin samples showing increased absorption results were not identified as invalid with the standard TEER limit of 1 kΩ, but almost all with the stricter limit of 2 kΩ. It seems that the standard limit value of 1 kΩ, originally derived from experiments with native versus punched human skin samples, is unable to detect minor damages. Furthermore the 2 kΩ limit provides more reasonable mean values for valid samples as ^{14}C-caffeine and ^{14}C-testosterone absorption in accordance with previous data (van de Sandt et al., 2004). Rather homogeneous MCPA-2EHE absorption appears to indicate that no impaired skin sample was apparent (Fig. 1). However, some skin samples identified as invalid by TEWL, TWF and TEER (2 kΩ) (Tables 4–6) once more highlights the probability to discard integer skin samples and the usefulness of concurrent or post-experimental integrity tests. Furthermore, the applicability of TEWL, TWF and TEER as integrity tests in dermal absorption studies for highly lipophilic compounds could be questioned in general. Focusing on the permeation/loss of water or permeation of small electrolytes through the skin, these tests are suitable to identify changes in the polar pathway of the skin. Changes in the lipid pathway, which is relevant for highly lipophilic compounds like MCPA-2EHE, can be overlooked; meaning that these tests are not representative for the penetration of highly lipophilic compounds. The contribution of polar- and lipid-intercellular, intracellular and appendageal pathways through skin depend on the physico-chemical properties of the test compound (Flynn et al., 1974; Roberts and Cross, 2002).

Rougier et al. observed good correlations between TEWL and absorption of benzoic acid in vivo (Rougier et al., 1988). However, the comparability of water flow through skin tissue in vivo and in vitro is limited. Previous work about TEWL application in vitro indicates that only severe damages can be detected (Netzlaff et al., 2006). The same conclusion is drawn for the current work where no, poor or even inverse correlations were observed between TEER, TEWL or TWF and test compound absorption (Table 7). Yet, the stated general applicability for in vitro testing failed to reflect ^{14}C-mannitol (Lawrence, 1997) and ^{35}sulfur mustard absorption in vitro (Chilcott et al., 2002). A lack of correlation to highly lipophilic test compounds was reported, too (Levin and Maibach, 2005).

Taken together all three standard tests are able to sort out a substantial part of impaired human skin samples in general. Limit values of 2 kΩ, 10 g m^{-2} h^{-1} and $4.5 * 10^{-3}$ cm h^{-1} for TEER, TEWL and TWF, respectively, seem appropriate to judge between unwanted use of impaired skin and unnecessary rejection of skin samples. However, destruction of barrier function during the experiment does not become obvious by these tests and – shown by falsely classified skin – only a rough differentiation is possible. Furthermore, none of the named integrity tests seems universally applicable. Defined 'applicability domains' for each integrity test which limits their use to test compounds in specific physico-chemical spaces or to specific experimental conditions (in vitro and/or in vivo, human and/or rat skin, excised and/or reconstructed skin etc.) can help to choose the most indicative test for the relevant case. Moreover, future use of reconstructed human skin for testing of dermal absorption asks for the adjustment of the generated data to human skin based on a prediction model (Schäfer-Korting et al., 2008) which still needs to be set up. For this purpose, the cut-off values need to be adapted as well.

4.2. Evaluation of ISTD and BLUE in relation to standard methods

Because of the limitations of the standard integrity tests (TEER, TEWL and TWF), two other integrity parameters (ISTD, BLUE) were checked for their ability to correlate with absorption results and explain continuous differences of the skin barrier function. Extreme outliers were clearly identified with BLUE, but correlations to test compound absorption were poor and partly even inverse. Although a general applicability of BLUE cannot be ruled out, lack of advantage over established tests makes further investigations redundant. The opposite was true for ISTD. These results were positively and highly correlated with test compound results. The correlation over a wide absorption range of ^{14}C-MCPA (6–100%) to ^{3}H-testosterone as internal reference standard was 0.859 ($n = 45$). Comparison of results for normal and intentionally damaged rat skin samples suggests under these experimental conditions (rat skin, receptor fluid water) a provisional cut-off value of 35% AD ^{3}H-testosterone (Fig. 2). The good correlation and the possibility to monitor the skin over the whole experimental time make the ISTD a promising tool. Besides other aspects it could help to distinguish compound-specific wash-in effects from barrier-disruption related effects.

In contrast to the recommendation of the OECD-Guideline we decided against ^{3}H-sucrose as ISTD because of poor information about applicability and the set limit value of 5% absorption (Walters et al., 1997). Moreover, the very high hydrophilic compound sucrose is not representative for routinely tested lipophilic test compounds. In accordance with the above-mentioned 'applicability domain' for integrity tests, the ISTD should be selected on the basis of the physico-chemical properties of the test compound, to indicate representatively the barrier function in relation to the respective pathway through the skin. Another suggested reference compound for ISTD is phenol red. Yet a 100 times higher concentration of phenol red is needed to achieve the same analytical sensitivity as the ^{3}H-labeled reference compounds and high concentrations increase the risk to influence the test results (Dugard and Scott, 1986).

To get a first impression of the performance of different ISTDs, ^{3}H-caffeine and ^{3}H-mannitol were tested in parallel to ^{3}H-testosterone in human skin experiments. The combination ^{3}H-testosterone and ^{14}C-MCPA resulted in moderate and weak correlations (R^2 0.52 and 0.16 for AD and maxKp comparison, respectively). This is probably due to the divergent physico-chemical properties (logP 3.32 and −0.71 (at pH 7) and MW 288.4 and 200.6 g mol^{-1} for testosterone and MCPA, respectively), but also due to the narrow absorption range which was covered. In fact, once the absorption range was expanded, as done in the special investigation with damaged and undamaged rat skin, the correlation was improved (R^2 0.859 and 0.911 against AD and maxKp, respectively). Weak correlations were obtained with ^{3}H-mannitol as ISTD with ^{14}C-testosterone (R^2 0.34 and 0.14 for AD and maxKp comparison, respectively) and ^{14}C-caffeine (R^2 0.20 and 0.40 for AD and maxKp comparison, respectively). Also in this case, the distance of the logP

values for the very polar ISTD ^{3}H-mannitol and the rather lipophilic test compounds was probably too large. For the combination ^{14}C-testosterone and ^{3}H-caffeine, having closer logP values, the best correlations with human skin were obtained (R^2 0.62 and 0.81 for AD and maxKp comparison, respectively). However, the reverse case (^{3}H-testosterone and ^{14}C-caffeine) resulted in weaker correlations (0.59 for maxKp comparison) and even no correlation (R^2 0.04 for AD comparison) – probably due to a lower number of replicates ($n = 5$) and one obvious outlier. Summing up, an ISTD with close physico-chemical properties to the test compound is preferable; however, the results imply that also ISTDs with a certain distance to the test compound are applicable.

Finally, the suitability of the current ISTD approach was proven by the independence of ^{14}C-analytics by LSC in the presence of ^{3}H (Fig. 3) and at maximum a negligible influence of ^{14}C-presence on ^{3}H-analytics (Table 9) as well as the independence of absorption results from the presence of an internal reference standard in negligible concentrations.

4.3. Variability of integrity tests and donor influence

To assess the intrinsic variability of the integrity tests and the effect of the human donor on the results, the overall, the inter-donor and the intra-donor variabilities were calculated for TEER, TEWL, TWF and BLUE (Table 8). For TEER, CVs for the overall, the inter-donor and the intra-donor variability were 64%, 45% and 43%, respectively. This implies that the variability of the method, given as the intra-donor variability, is close to the inter-donor variability and therewith covering the donor effect. The same is true for the other integrity tests (TEWL, TWF and BLUE), for which the donor effect was also close to the method variability. Therewith, a clear separation of human donors based on the integrity test results is hardly possible. Additionally, means and overall variability of the different integrity tests were calculated for full-thickness and dermatomed human skin separately (data not shown). In general, the values were close within each integrity test. For instance, TWF results were $302 \pm 188 * 10^{-5}$ cm h^{-1} ($n = 20$, CV = 62%) and $248 \pm 146 * 10^{-5}$ cm h^{-1} ($n = 20$, CV = 59%) for dermatomed and full-thickness skin from the same human donors, respectively. This is in line with the previously reported comparability of absorption results through both skin preparation types (Guth, 2013). Furthermore, the donor effect was consistent over all methods with values ranging from 32% to 48%. These values were also in line with the general donor effect observed for dermal absorption experiments in vitro being ~43% (Southwell et al., 1984). The overall method variabilities determined in this study for four different test compounds are with CVs of 33% and 45% for maxKp and AD, respectively, in line with the reported variability ranging from 2% to 111% (Southwell et al., 1984; van de Sandt et al., 2004). The method variabilities obtained for all five integrity tests, including ISTD, are in the same range (30–57%).

4.4. Solitary versus continuous integrity tests

The ISTD is advantageous over the 'solitary' integrity tests conducted in advance or after an absorption experiment, since outliers or abnormalities observed for the kinetics of the test compound can be interpreted in parallel with the kinetics of the ISTD. For instance, an abrupt increase of absorption of the test compound after the washing procedure is classified as a wash-in effect due to mechanical disruption of the barrier if the ISTD shows a parallel effect, or it is classified as a substance-specific wash-in effect if the absorption of the ISTD is not affected. The latter case – washing increases the test compound absorption – can be relevant for regulatory purposes. In addition, formulation-related barrier impairment could be identified. If all skin samples are affected, resulting in higher absorption rates for ISTD in comparison to historical control data, the formulation influences the barrier function (irritation, disturbance or disruption). In contrast, if only few, e.g. two out of eight, skin samples are affected, it is a sign for per se impaired skin samples whose results need to be rejected. Furthermore, systematic errors could be evaluated. For instance, if higher skin temperatures or higher receptor flow rates are logged during an experiment, ISTD results in the historical range will argue against an effect of these variations on the test compound absorption or in other words will argue for a valid experiment. And finally a continuous test avoids any kind of pretreatment and elongation of the experiment which could alter the skin properties as outlined above (Buist et al., 2005). However, besides all these advantages, a continuous test is unsatisfactory as a stand-alone method. No preselection of skin samples is made, why impaired skin samples might be used. To avoid an insufficient number of valid skin samples for the entire study, we recommend a combined use of the binary standard test TEWL in advance of an experiment – which is able to identify the majority of defect skin samples without pretreatment of the skin samples – and the outlined continuous ISTD approach – to evaluate effects observed during the absorption experiment. Since good correlations were observed for all skin preparation types (excised human, reconstructed human and excised rat skin), the ISTD approach is probably transferable to diseased skin or reconstructed diseased skin (Kuechler et al., 2011; Oji et al., 2010) as well.

However, an obstacle for the routine application of the ISTD approach is the need of a broad, publicly available, historical dataset. In theory, this dataset should be a matrix of various ISTDs with different physico-chemical properties applied under several experimental conditions. Compounds with various logP values and MWs should be included, since these properties mainly determine their dermal absorption (Riviere, 2011). This would allow adjustment of the reference compound to the physico-chemical properties of the test compound in order to address the same pathway through the skin. However, to keep it practicable for routine application it is recommended to establish at least representatives for high, medium and low logP ranges. This would be in parallel to the suggested reference compounds stated in the guideline (caffeine, benzoic acid and testosterone) (OECD, 2004b) and cover different pathways through the skin. The ISTD with the logP value closest to the logP value of the test compound should be chosen for the experiment. That a certain distance is generally acceptable was shown in the current work. Since different conditions (like donor or receptor fluid) can influence the ISTD results (Kielhorn et al., 2006; Schäfer and Redelmeier, 1996a) there is also a need to generate data under the relevant scenarios. For the investigation of pesticides a breakdown to three scenarios is conceivable: aqueous donor and aqueous receptor, organic donor and organic/aqueous receptor, organic donor and aqueous receptor. To gain experience concerning the effect of formulation on the ISTD, additional experiments using ISTD in parallel to standard routine experiments without ISTD are feasible. If no data of intentionally damaged skin is available for setting a cut-off limit (as done in the current work, Fig. 2), routine data could be used to depict a frequency histogram and use the 95th percentile threshold as previously done for TWF (Fasano et al., 2002; Meidan and Roper, 2008).

5. Conclusion

In conclusion the standard integrity tests TEER, TEWL and TWF are useful to distinguish between impaired and intact human skin samples prior to a dermal absorption experiment, if limit values of 10 g m^{-2} h^{-1}, $4.5 * 10^{-3}$ cm h^{-1} and 2 kΩ, respectively, are applied. The application of one of these tests is recommended for routine experiments. Furthermore, adding an internal reference standard

to the test compound allows a continuous assessment of the barrier functionality over the entire experimental period. Combining both, an effective and non-invasive pre-test like TEWL and the concept of ISTD could improve the quality of dermal absorption experiments in the future. However, the routine application of ISTD is hampered by the need of a historical dataset which is required to define thresholds of integrity and develop a general protocol.

Conflict of Interest

Katharina Guth, Eric Fabian, Robert Landsiedel and Ben van Ravenzwaay are employees of BASF SE – a chemical company which may use the described models in the development of commerical products.

Transparency Document

The Transparency document associated with this article can be found in the online version.

Acknowledgements

We would like to thank Geoffrey Pigott for providing the test compounds MCPA and MCPA-2EHE as well as ingredients for the MCPA formulation.

References

Ackermann, K., Borgia, S.L., Korting, H.C., Mewes, K.R., Schäfer-Korting, M., 2010. The Phenion (R) full-thickness skin model for percutaneous absorption testing. Skin Pharmacol. Physiol. 23, 105–112.

Balaguer, A., Salvador, A., Chisvert, A., Melia, M., Herraez, M., Diez, O., 2006. A liquid chromatography-fluorimetric method for the in vitro estimation of the skin penetration of disodium phenyldibenzimidazole tetrasulfonate from sunscreen formulations through human skin. Anal. Bioanal. Chem. 385, 1225–1232.

Barry, B.W., 1987. Penetration enhancers – mode of action in humans. In: Shroot, B., Schäfer, H. (Eds.), Skin Pharmacokinetics, pp. 121–137.

Brain, K.R., James, V.J., Walters, K.A., 1995. Validation of the integrity of epidermal membranes during assessment of percutaneous penetration. Predict. Percut. Penet.

Britisch Crop Protection Council, 2011. The e-pesticide manual V5.2. British Crop Production Council, Hampshire, UK.

Bronaugh, R.L., Stewart, R.F., Simon, M., 1986. Methods for in vitro percutaneous-absorption studies. 7. Use of excised human-skin. J. Pharm. Sci. 75, 1094–1097.

Buist, H.E., van de Sandt, J.J.M., van Burgsteden, J.A., de Heer, C., 2005. Effects of single and repeated exposure to biocidal active substances on the barrier function of the skin in vitro. Regul. Toxicol. Pharmacol. 43, 76–84.

Chilcott, R.P., Dalton, C.H., Emmanuel, A.J., Allen, C.E., Bradley, S.T., 2002. Transepidermal water loss does not correlate with skin barrier function in vitro. J. Invest. Dermatol. 118, 871–875.

Davies, D.J., Ward, R.J., Heylings, J.R., 2004. Multi-species assessment of electrical resistance as a skin integrity marker for in vitro percutaneous absorption studies. Toxicol. In Vitro 18, 351–358.

Diembeck, W., Beck, H., Benech-Kieffer, F., Courtellemont, P., Dupuiss, J., Lovell, W., Paye, M., Spengler, J., Steiling, W., 1999. Test guidelines for in vitro assessment of dermal absorption and percutaneous penetration of cosmetic ingredients. Food Chem. Toxicol. 37, 191–205.

Dugard, P.H., Scott, R.C., 1986. A method of predicting percutaneous-absorption rates from vehicle to vehicle – an experimental assessment. Int. J. Pharm. 28, 219–227.

EFSA Panel on Plant Protection Products and their Residues, 2012. Guidance on Dermal Absorption. EFSA J., 10.

Elkeeb, R., Hui, X.Y., Chan, H.D., Tian, L.L., Maibach, H.I., 2010. Correlation of transepidermal water loss with skin barrier properties in vitro: comparison of three evaporimeters. Skin Res. Technol. 16, 9–15.

Fasano, W.J., Manning, L.A., Green, J.W., 2002. Rapid integrity assessment of rat and human epidermal membranes for in vitro dermal regulatory testing: correlation of electrical resistance with tritiated water permeability. Toxicol. In Vitro 16, 731–740.

Flynn, G.L., Yalkowsk, S.H., Roseman, T.J., 1974. Mass-transport phenomena and models – theoretical concepts. J. Pharm. Sci. 63, 479–510.

Guth, K., 2013. Models for percutaneous absorption – influential parameters and in silico predictions. Freie Universität, Berlin.

Imhof, R., De Jesus, M., Xiao, P., Ciortea, L.I., Berg, E., 2009. Closed-chamber transepidermal water loss measurement: microclimate, calibration and performance. Int. J. Cosmetic Sci. 31, 97–118.

Kielhorn, J., Melching-Kollmuss, S., Mangelsdorf, I., 2006. Dermal absorption, Environmental Health Criteria 235. World Health Organization.

Kuechler, S., Henkes, D., Eckl, K.M., Ackermann, K., Plendl, J., Korting, H.C., Hennies, H.C., Schäfer-Korting, M., 2011. Hallmarks of atopic skin mimicked in vitro by means of a skin disease model based on FLG knock-down. Atla-Altern. Lab. Animals 39, 471–480.

Lawrence, J.N., 1997. Electrical resistance and tritiated water permeability as indicators of barrier integrity of in vitro human skin. Toxicol. In Vitro 11, 241–249.

Leo, A., Hansch, C., Elkins, D., 1971. Partition coefficients and their uses. Chem. Rev. 71, 525+.

Levin, J., Maibach, H., 2005. The correlation between transepidermal water loss and percutaneous absorption: an overview. J. Control. Rel. 103, 291–299.

Marzulli, F.N., Brown, D.W.C.M.H.I., 1969. Techniques for studying skin penetration. Toxicol. Appl. Pharmacol. 3, 76–83.

Meidan, V.M., Roper, C.S., 2008. Inter- and intra-individual variability in human skin barrier function: a large scale retrospective study. Toxicol. In Vitro 22, 1062–1069.

Merck, 2006. Safety Data Sheet Caffeine.

Netzlaff, F., Kostka, K.H., Lehr, C.M., Schäfer, U.F., 2006. TEWL measurements as a routine method for evaluating the integrity of epidermis sheets in static Franz type diffusion cells in vitro. Limitations shown by transport data testing. Eur. J. Pharm. Biopharm. 63, 44–50.

OECD, 2004a. Guidance document for the conduct of skin absorption studies – OECD Series in Testing and assessment – Number 28. Organisation for Economic Co-operation and Development.

OECD, 2004b. OECD Guideline for the testing of chemicals. Skin absorption: in vitro method 428. Organisation for Economic Co-operation and Development.

Oji, V., Eckl, K.M., Aufenvenne, K., Natebus, M., Tarinski, T., Ackermann, K., Seller, N., Metze, D., Nurnberg, G., Folster-Holst, R., Schafer-Korting, M., Hausser, I., Traupe, H., Hennies, H.C., 2010. Loss of corneodesmosin leads to severe skin barrier defect, pruritus, and atopy: unraveling the peeling skin disease. Am. J. Human Genet. 87, 274–281.

Peck, K.D., Ghanem, A.H., Higuchi, W.I., 1995. The effect of temperature upon the permeation of polar and ionic solutes through human epidermal membrane. J. Pharm. Sci. 84, 975–982.

Pendlington, R.U., Sanders, D.J., Cooper, K.J., Howes, D., Lovell, W., 1997. The use of sucrose as a standard penetration in in vitro percutaneous penetration experiments. In: Perspectives in Percutaneous Penetration. Wiley-Blackwell, p. 55, vol. 5a.

Riviere, J.E., 2011. Absorption. In: Riviere, J.E. (Ed.), Comparative Pharmacokinetics. Wiley-Blackwell, pp. 39–71.

Roberts, M.S., Cross, S.E., 2002. Skin transport. In: Walters, K.A. (Ed.), Dermatological and Transdermal Formulations. Marcel Dekker Inc, New York, pp. 89–195.

Rougier, A., Lotte, C., Maibach, H.I., 1988. Invivo relationship between trans-epidermal water-loss and percutaneous-absorption in man. Clin. Res. 36, A690.

Schäfer, H., Redelmeier, T.E., 1996a. Factors effecting percutaneous absorption. In: Schäfer, H., Redelmeier, T.E., Skin Barrier, Principles of Percutaneous Absorption, 153–212.

Schäfer, H., Redelmeier, T.E., 1996b Prediction and Measurement of Percutaneous Absorption. In: Schäfer, H., Redelmeier, T.E., Skin Barrier, Principles of Percutaneous Absorption, 118–152.

Schäfer-Korting, M., Bock, U., Gamer, A., Haberland, A., Haltner-Ukomadu, E., Kaca, M., Kamp, H., Kietzmann, M., Korting, H.C., Krachter, H.U., Lehr, C.M., Liebsch, M., Mehling, A., Netzlaff, F., Niedorf, F., Rubbelke, M.K., Schäfer, U., Schmidt, E., Schreiber, S., Schroder, K.R., Spielmann, H., Vuia, A., 2006. Reconstructed human epidermis for skin absorption testing: results of the German prevalidation study. Atla-Altern. Lab. Animals 34, 283–294.

Schäfer-Korting, M., Bock, U., Diembeck, W., Duesing, H.J., Gamer, A., Haltner-Ukomadu, E., Hoffmann, C., Kaca, M., Kamp, H., Kersen, S., Kietzmann, M., Korting, H.C., Krachter, H.U., Lehr, C.M., Liebsch, M., Mehling, A., Mueller-Goymann, C., Netzlaff, F., Niedorf, F., Rubbelke, M.K., Schäfer, U., Schmidt, E., Schreiber, S., Spielmann, H., Vuia, A., Weimer, M., 2008. The use of reconstructed human epidermis for skin absorption testing: results of the validation study. Atla-Altern. Lab. Animals 36, 161–187.

Shah, D.K., Khandavilli, S., Panchagnula, R., 2008. Alteration of skin hydration and its barrier function by vehicle and permeation enhancers: a study using TGA, FTIR, TEWL and drug permeation as markers. Meth. Find. Exp. Clin. Pharmacol. 30, 499–512.

Sigma-Aldrich, 2007. Sicherheitsdatenblatt D-Mannitol.

Southwell, D., Barry, B.W., Woodford, R., 1984. Variations in permeability of human-skin within and between specimens. Int. J. Pharm. 18, 299–309.

van de Sandt, J.J.M., van Burgsteden, J.A., Cage, S., Carmichael, P.L., Dick, I., Kenyon, S., Korinth, G., Larese, F., Limasset, J.C., Maas, W.J.M., Montomoli, L., Nielsen, J.B., Payan, J.P., Robinson, E., Sartorelli, P., Schaller, K.H., Wilkinson, S.C., Williams, F.M., 2004. In vitro predictions of skin absorption of caffeine, testosterone, and benzoic acid: a multi-centre comparison study. Regul. Toxicol. Pharmacol. 39, 271–281.

Walters, K.A., Brain, K.R., Howes, D., James, V.J., Kraus, A.L., Teetsel, N.M., Toulon, M., Watkinson, A.C., Gettings, S.D., 1997. Percutaneous penetration of octyl salicylate from representative sunscreen formulations through human skin in vitro. Food Chem. Toxicol. 35, 1219–1225.

Yalkowsky, S.H., Valvani, S.C., Roseman, T.J., 1983. Solubility and partitioning. 6. Octanol solubility and octanol–water partition-coefficients. J. Pharm. Sci. 72, 866–870.