Anal Bioanal Chem
DOI 10.1007/s00216-011-4707-y

ORIGINAL PAPER

Chiral analyses of dextromethorphan/levomethorphan and their metabolites in rat and human samples using LC-MS/MS

Ruri Kikura-Hanajiri · Maiko Kawamura · Atsuko Miyajima · Momoko Sunouchi · Yukihiro Goda

Received: 30 September 2010 / Revised: 7 January 2011 / Accepted: 22 January 2011

Abstract In order to develop an analytical method for the discrimination of dextromethorphan (an antitussive medicine) from its enantiomer, levomethorphan (a narcotic) in biological samples, chiral analyses of these drugs and their *O*-demethyl and/or *N*-demethyl metabolites in rat plasma, urine, and hair were carried out using LC-MS/MS. After the i.p. administration of dextromethorphan or levomethorphan to pigmented hairy male DA rats (5 mg/kg/day, 10 days), the parent compounds and their three metabolites in plasma, urine and hair were determined using LC-MS/MS. Complete chiral separation was achieved in 12 min on a Chiral CD-Ph column in 0.1% formic acid–acetonitrile by a linear gradient program. Most of the metabolites were detected as being the corresponding *O*-demethyl and *N*, *O*-didemethyl metabolites in the rat plasma and urine after the hydrolysis of *O*-glucuronides, although obvious differences in the amounts of these metabolites were found between the dextro and levo forms. No racemation was observed through *O*- and/or *N*-demethylation. In the rat hair samples collected 4 weeks after the first administration, those differences were more clearly detected and the concentrations of the parent compounds, their *O*-demethyl, *N*-demethyl, and *N*, *O*-didemethyl metabolites were 63.4, 2.7, 25.1, and 0.7 ng/mg for the dextro forms and 24.5, 24.6, 2.6, and 0.5 ng/mg for the levo forms, respectively. In order to fully investigate the differences of their metabolic properties between dextromethorphan and levomethorphan, DA rat and human liver microsomes were studied. The results suggested that there might be an enantioselective metabolism of levomethorphan, especially with regard to the *O*-demethylation, not only in DA rat but human liver microsomes as well. The proposed chiral analyses might be applied to human samples and could be useful for discriminating dextromethorphan use from levomethorphan use in the field of forensic toxicology, although further studies should be carried out using authentic human samples.

Keywords Levomethorphan · Dextromethorphan · Chiral analysis · Biological samples · LC-MS/MS · Enantioselective metabolism

Published in the special issue *Forensic Toxicology* with Guest Editors Frank T. Peters, Hans H. Maurer, and Frank Musshoff.

R. Kikura-Hanajiri (✉) · M. Kawamura · Y. Goda
Division of Pharmacognosy, Phytochemistry and Narcotics,
National Institute of Health Sciences,
1-18-1, Kamiyoga, Setagaya,
Tokyo 158-8501, Japan
e-mail: kikura@nihs.go.jp

A. Miyajima
Division of Medical Devices,
National Institute of Health Sciences,
1-18-1, Kamiyoga, Setagaya,
Tokyo 158-8501, Japan

M. Sunouchi
Division of Pharmacology, National Institute of Health Sciences,
1-18-1, Kamiyoga, Setagaya,
Tokyo 158-8501, Japan

Published online: 11 February 2011

Introduction

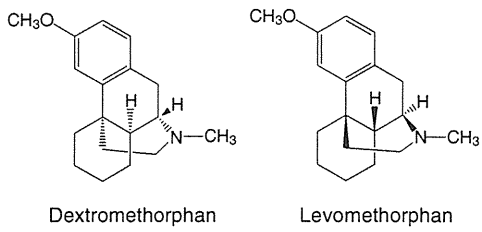
Dextromethorphan is widely used all over the world as an over-the-counter antitussive medicine. It produces little or no central nervous system depression at therapeutic doses, but it has dissociative effects similar to ketamine and phencyclidine in large doses as an *N*-methyl-*D*-aspartate receptor antagonist [1–6]. To obtain its hallucinogenic effect, young people abuse this drug by large doses and many fatalities from overdoses have been reported [7, 8]. In

contrast, its enantiomer, levomethorphan, is a potent narcotic analgesic [9] (Fig. 1), and an *O*-demethyl compound of levomethorphan, levorphanol, is known to have stronger opioid pharmacological effects [9, 10]. Levomethorphan is strictly controlled in the world as a narcotic and is never used for therapeutic purposes.

In humans, as shown in Fig. 2, it has been reported that dextromethorphan is primarily metabolized to dextrorphan via *O*-demethylation by cytochrome P450 2D6 (CYP2D6), which is polymorphically expressed in humans, who can be classified as poor, intermediate, and extensive metabolizers [11, 12]. Dextromethorphan is *N*-demethylated via an additional route to 3-methoxymorphinan (3-MEM), primarily mediated by CYP3A4 in human liver microsomes [11, 13]. Dextrorphan and 3-MEM are then demethylated to 3-hydroxymorphinan (3-HM) via CYP3A4 and CYP2D6, respectively. Moreover, dextrorphan and 3-HM are glucuronized to their *O*-glucuronides and these are mainly excreted into human urine [14, 15].

A variety of analytical methods for the determination of dextromethorphan and its metabolites in biological samples have been reported using capillary electrophoresis (CE) [16, 17], HPLC [18–22], GC-MS [23–26], and LC-MS (/MS) [15, 27–30]. However, there is little information regarding the metabolic properties of levomethorphan. Although a chiral separation method of dextromethorphan and levomethorphan using HPLC with fluorometric detection [22] or using CE [17] has been reported, there has been no report that describes a simultaneous determination of dextromethorphan, levomethorphan, and their metabolites in biological samples after administration of these drugs. Considering the possibility of the adulteration or substitution of dextromethorphan with levomethorphan due to its chemical similarities for illegal purposes, it is necessary to establish the enantiometric separation of dextromethorphan and levomethorphan in biological samples.

In order to develop an analytical method for the discrimination of dextromethorphan from levomethorphan in biological samples, we first investigated chiral analyses of these drugs and their *O*-demethyl and/or *N*-demethyl metabolites in plasma, urine, and hair samples of rats administered with each enantiomer, using LC-MS/MS. In addition, detailed metabolic properties of these drugs were investigated using rat and human liver microsomes.



Dextromethorphan Levomethorphan

Fig. 1 Chemical structures of dextromethorphan and its enantiomer, levomethorphan

Experimental

Materials

Levomethorphan was obtained from Cerilliant (Round Rock, TX, USA). Dextromethorphan hydrobromide, dextrorphan tartrate, (+)-3-HM hydrobromide, (+)-3-MEM hydrochloride, and levallorphan tartrate (used as internal standard, IS) were from Sigma-Aldrich (St. Louis, MO, USA). Levorphanol tartrate was given by Professor T. Nagano (Graduate School of Pharmaceutical Sciences, The University of Tokyo, Japan). A β-glucuronidase solution (EC 3.2.1.31, 103,000 units/mL, Source: *Helix pomatia*) was purchased from Wako Chemicals (Osaka, Japan). A solid-phase extraction column (OASIS HLB, 3 cc, 60 mg) was obtained from Waters (Milford, MA, USA), and the membrane filter (Ultrafree-MC, 0.45 μm) was from Millipore Corporation (Bedford, MA, USA).

Liver microsomes from individual male dark agouti (DA) rats (n=4, 6 weeks old, around 125 g mean weight) were prepared by ultracentrifugation as described [31, 32]. The concentrations of microsomal protein were estimated using a bicinchoninic acid (BCA) protein assay kit (Pierce, Rockford, IL). Human liver microsomes (50-donor pool, 20 mg/mL) were purchased from BD Biosciences (Woburn, MA, USA). Nicotinamide adenine dinucleotide phosphate (NADP) and glucose 6-phosphate (G-6-P) were obtained from Oriental Yeast Co., Ltd. (Tokyo, Japan) and G-6-P dehydrogenase (G-6-PDH) was from Roche Diagnostics (Indianapolis, IN, USA). All other chemicals and solvents were of an analytical reagent grade or HPLC grade (Wako Chemicals, Osaka, Japan).

Instrumentation

The UPLC analysis was performed using a Waters Acquity Ultra-Performance™ liquid chromatography system (Waters, Milford, MA, USA). The separations were achieved using a Chiral CD-Ph column (150×2.1 mm i.d., 5 μm) from Shiseido (Tokyo, Japan). The column temperature was maintained at 30 °C, and the following gradient system was used with a mobile phase A (0.1% formic acid) and mobile phase B (0.1% formic acid/acetonitrile) delivered at 0.25 mL/min: 80% A/20% B (2 min hold)—70% A/30% B (15 min). The mobile phase was used as a wash solvent to avoid any carry-over from previous injections. The autosampler was maintained at 4 °C, and the injection volume was 2 μL. Quantitation was achieved by MS/MS detection in a positive ion mode using a Quattro Premier XE mass

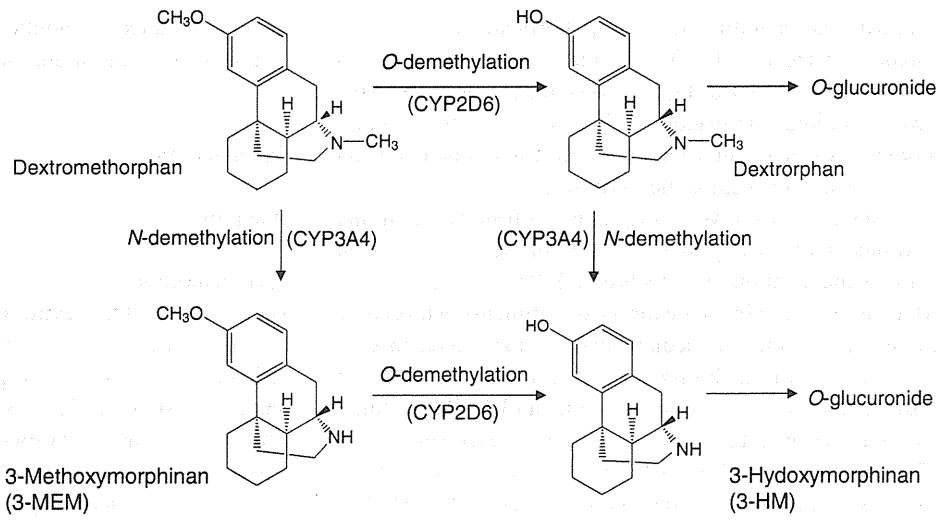


Fig. 2 Main metabolic pathway of dextromethorphan in humans

spectrometer (Waters, Milford, MA, USA) equipped with an electrospray ionization (ESI) interface. Quantification was performed using multiple reaction monitoring of the transitions of precursor ions to product ions with each cone voltage and collision energy as shown in Table 1. The optimal MS parameters obtained were as follows: capillary 3.0 kV, source temperature 120 °C, and desolvation temperature 350 °C. Nitrogen was used as the desolvation and cone gas, with a flow rate of 800 and 50 L/h, respectively. Argon was used as the collision gas, with a flow rate of 0.25 mL/min. All data collected in the centroid mode were processed using MassLynx™ NT4.1 software with a QuanLynx™ program (Waters, Milford, MA, USA).

Since the standard compounds of (−)-3-MEM and (−)-3-HM were not available, these peaks were confirmed by comparison of their retention times and mass fragmentations with those of the standard compounds of the dextro forms ((+)-3-MEM and (+)-3-HM) using an ODS column. The analyses were performed using an Acquity HSS T3 column (100×2.1 mm i.d., 1.8 μm) from Waters (Milford, MA, USA). The column temperature was maintained at 40 °C, and the following gradient system was used with a mobile phase A (1% formic acid) and mobile phase B (1% formic acid/acetonitrile) delivered at 0.3 mL/min: 90% A/10% B (0 min)–70% A/30% B (8 min). The MS parameters were the same as for the analyses using the chiral column described above.

Animal experiments

The animal experimental model was designed as shown in our previous reports [33, 34]. All experiments were carried out with the approval of the Committee for Animal Care and Use of the National Institute of Health Sciences, Japan. Dextromethorphan hydrobromide (dissolved in an isotonic sodium chloride solution, 2.5 mg/mL, rat 1–3) or levomethorphan (dissolved in a mixed solution of 5% Emulphor™ EL-620/5% ethanol/90% isotonic sodium chloride solution, 2.5 mg/mL, rat 4–6) was administered to male DA pigmented rats, which were 5 weeks old and around 90 g mean weight (Japan SLC, Shizuoka, Japan). The drugs were given once daily at 5 mg/kg by intraperitoneal injection for ten successive days. Blood samples were collected 5, 15, 30, 60, 120, and 360 min after the first administration from the orbital vein plexus. Plasma samples were prepared by centrifugation at 10,000×*g* for 3 min and stored at −20 °C until analysis. The area under the plasma concentration time curve (AUC) was calculated by the conventional method. Urine samples were collected 0–24, 24–48, and 48–72 h after the last administration and stored at −20 °C. Each animal had been shaved on the back just before the first drug administration. The new growing hair samples were collected 28 days after the first administration.

Extraction of parent compounds and their metabolites from rat plasma and urine samples

For the quantitative analysis of *O*-demethyl and *N*,*O*-didemethyl metabolites in the rat plasma and urine samples, the analytes were measured as free compounds after the hydrolysis of *O*-glucuronides. The optimal condition of the hydrolysis was evaluated, with the peak of putative *O*-glucuronide at nearly 2 min (*m/z* 434→258) on the MRM chromatogram disappearing from rat plasma and urine samples after the hydrolysis.

A 25-μL plasma sample with 50 μL of added 10 mM ammonium formate buffer (pH 5.0) was reacted with 20 μL of the β-glucuronidase solution at 37 °C for 20 h. To precipitate the proteins in the plasma, 40 μL of the IS methanol solution and 100 μL of methanol were poured into each tube, and the mixtures were then vigorously mixed. The

Table 1 Analytical conditions of LC-MS/MS using the Chiral CD-Ph column

Compounds	Retention time min	Precursor ions *m/z*	Cone voltage V	Product ions *m/z*	Collision energy eV
Dextromethorphan	10.6	272	40	171	45
Dextrorphan	6.1	258	45	157	40
(+)-3-MEM	8.1	258	40	170	35
(+)-3-HM	3.9	244	30	156	35
Levomethorphan	11.3	272	40	171	45
Levorphanol	5.5	258	45	157	40
(−)-3-MEM	9.8	258	40	170	35
(−)-3-HM	4.5	244	30	156	35
Levallorphan (IS)	7.5	284	40	157	40

mixed solution was centrifuged at 1,200×*g* for 3 min and filtered prior to the injection for the LC-MS/MS analysis.

To a 50-μL urine sample (20 μL for 0–24 h samples) was added 100 μL of the β-glucuronidase solution, 1 mL of 10 mM ammonium formate buffer (pH 5.0) and 50 μL of the IS aqueous solution, respectively. The mixed solution was incubated at 37 °C with gentle shaking. After an OASIS HLB column was pre-activated with 2 mL of methanol and distilled water, the reaction mixture was applied to the column. Following the wash of the column with 2 mL of distilled water, 1 mL of methanol was passed through the column to elute the target drugs. A 2-μL of the solution was automatically injected into the UPLC-MS/MS.

Extraction of parent compounds and their metabolites from rat hair samples

Hair samples were washed three times with 0.1% sodium dodecyl sulfate under ultrasonication, followed by washing three times with water under the same condition. After the sample was dried under a nitrogen stream at room temperature, approximately 10 mg of finely cut hair was precisely weighed and extracted with 1 mL of methanol/5 M hydrochloric acid mixed solution (20:1) containing 50 μL of the IS methanol solution for 1 h under ultrasonication. Following overnight storage at room temperature, the hair was filtered off, the filtrate was evaporated with a nitrogen stream, and the residue was dissolved in 1 mL of distilled water. The solution was treated with an OASIS HLB column and analyzed as described above.

Linearity, precision, and recovery of the analytical method for the rat samples

An individual standard solution of 1.0 mg/mL of each drug, dextromethorphan, levomethorphan, dextrorphan, 3-hydroxymorphinan, 3-methoxymorphinan, and levorphanol, was prepared in methanol and stored at 4 °C. The IS solutions of 1 μg/mL of levallorphan in methanol for the analysis of hair samples and those of 1 μg/mL of levallorphan in distilled water for plasma and urine samples were also prepared.

The drug concentrations in the samples were calculated using the peak–area ratios of the ions monitored for the target compounds versus IS. The calibration curves for the determination were constructed by analyzing extracted drug-free control samples spiked with the standard solution, as described above. The calibration samples containing 0, 1, 2, 4, 20, 40, 200, and 400 ng/mL of the target drugs for the rat plasma, 0, 5, 10, 50, 100, 500, 1,000 2,500, 5,000, and 10,000 ng/mL for the urine samples and 0, 0.1, 0.5, 1.0, 5.0, 10, 25, and 50 ng/mg for the hair samples were prepared just before analysis. The limit of quantitation (LOQ) of each drug was chosen to be the concentration of the lowest calibration standard with an acceptable limit of variance, while the limit of detection (LOD) was defined as concentrations in a sample matrix resulting in peak areas with signal-to-noise ratios (*S/N*) of 3.

The precision of the method was evaluated by five consecutive analyses of the plasma and urine samples that were spiked with the standard solutions containing 2, 20, and 200 ng/mL for the rat plasma samples and 5, 500, and 5,000 ng/mL for the urine samples, respectively. For the hair analyses, the control samples spiked with the standard solutions each containing 0.1, 5, and 50 ng/mg of the targeted drugs were evaluated. The recoveries of the four analytes from the rat samples were determined using each sample spiked with the analytes at a concentration of 80 ng/mL for the plasma, 500 ng/mL for the urine, and 10 ng/mg for the hair, respectively. To determine the recoveries, the responses of the analytes in the standard solutions and in the extracts from the rat control samples were compared. For the quantitative analysis of (−)-MEM and (−)-HM, the calibration curves of (+)-MEM and (+)-HM were used.

Demethylation of dextromethorphan/levomethorphan in rat and human liver microsomes

For the *in vitro* experiments with rat and human liver microsomes, the reaction mixture consisted of 0.1 M potassium phosphate buffer (pH 7.4) with an NADPH generating system (1.3 mM NADP, 3.3 mM G-6-P, 0.4 U/mL G-6-PDH, 3.3 mM $MgCl_2$), 50 μM substrate (dextromethorphan or levomethorphan), and 0.5 mg protein/mL microsomes (rat or human liver microsomes) in a final volume of 200 μL. Dextromethorphan and levomethorphan were dissolved in methanol, and the final concentration of the organic solvent was 0.1%. The incubation was started by adding the microsomal fraction and then continued for 0, 5, 10, or 20 min. The reaction was terminated by adding an equal volume of a mixed organic solution of 50% acetonitrile and 50% methanol, including 10 μM levallorphan (IS), and vigorous shaking. At the same time, a reaction mixture without the microsomal fraction was also incubated as an enzyme-free control. The mixture was centrifuged at 3,500×*g* for 3 min at 4 °C, and the supernatant was filtered prior to the injection for the LC-MS/MS analysis. The *in vitro* experiments for kinetic analyses were also performed as described above, except that 2, 5, 10, 50, 100, and 150 μM of substrates were incubated with the rat and human liver microsomes for 10 min. Each experiment was performed in duplicate and kinetic parameters were calculated with Eadie–Hofstee plots.

The results of the *in vitro* experiments were each evaluated by three consecutive analyses. The amounts of dextromethorphan/levomethorphan and their metabolites were calculated on the basis of calibration curves made by spiking known amounts of these compounds into the reaction mixture without the microsomal fraction.

Results

Chiral separation of dextromethorphan/levomethorphan and their metabolites

Complete chiral separation of dextromethorphan, levomethorphan, and their metabolites was achieved in 12 min on a Chiral CD-Ph column in 0.1% formic acid–acetonitrile by a linear gradient program. The retention time of each compound was as follows: the parent compounds (dextro/levo forms, 10.6/11.3 min) and their metabolites of *O*-demethyl (6.1/5.5 min), *N*-demethyl (8.1/9.8 min), and *O*, *N*-didemethyl (3.9/4.5 min) as shown in Table 1. Figure 3 shows LC-MS/MS total ion current chromatograms (MRM mode) of the extract from plasma (30 min after the first administration), urine (0–24 h after the last administration), and hair (collected 4 weeks after the first administration) of rats administered with dextromethorphan or levomethorphan. Under the chromatographic conditions used, there was no interference with any of the compounds or the internal standard by any extractable endogenous materials in the rat samples. The peaks 7 (9.8 min, *m/z* 258→170) and 8 (4.5 min, *m/z* 244→156) on the chromatograms shown in Fig. 3 were identified as those of (−)-3-MEM and (−)-3-HM when the mass fragmentations of these peaks were considered, although the standard compounds of these two metabolites were not available. These peaks were also confirmed by comparison of their retention times and mass fragmentations with those of the standard compounds of the dextro forms ((+)-3-MEM and (+)-3-HM) using an ODS column.

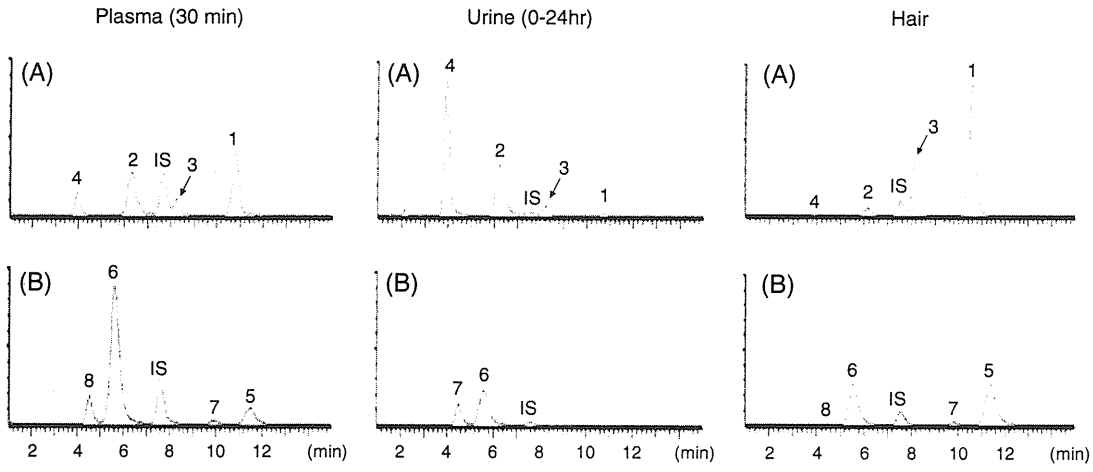


Fig. 3 LC-MS/MS total ion current (TIC) chromatograms (MRM mode) of the extracts from plasma, urine, and hair of rats administered with (**A**) dextromethorphan and (**B**) levomethorphan using a chiral column. *1* Dextromethorphan, *2* Dextrorphan, *3* (+)-3-MEM, *4* (+)-3-HM, *5* Levomethorphan, *6* Levorphanol, *7* (−)-3-MEM, *8* (−)-3-HM

Table 2 Validation of results of the LC-MS/MS analyses of dextromethorphan/levomethorphan and their metabolites in rat plasma, urine and hair samples (n=5)

Samples		Compounds	LOD (S/N>3)	LOQ (S/N>10)	Linear ranges	Recoveries (%)	Precision (%) (n=5)			Accuracy (%) (n=5)		
				(ng/mL)		80 ng/mL	2.0 ng/mL	20 ng/mL	200 ng/mL	2.0 ng/mL	20 ng/mL	200 ng/mL
Plasma (50 μL)	Dextro	Dextromethorphan	0.8	1.0	1.0-400	106.1	22.1	9.3	1.5	-19.2	5.5	-0.2
		Dextrorphan	0.4	0.8		81.7	10.2	3.8	1.5	10.2	2.2	-3.6
		(+)-3-MEM	0.8	1.0		110.5	15.0	3.2	2.5	23.5	2.1	2.6
		(+)-3-HM	0.8	1.0		92.5	15.7	6.1	1.8	13.7	-8.3	2.9
	Levo	Levomethorphan	0.8	1.0	1.0-400	100.8	8.6	4.9	2.5	21.6	-4.4	-5.7
		Levorphanol	0.8	1.0		90.7	15.9	4.1	2.3	-10.6	-5.6	-3.6
				(ng/mL)		500 ng/mL	5.0 ng/mL	500 ng/mL	5000 ng/mL	5.0 ng/mL	500 ng/mL	5000 ng/mL
Urine (100 μL)	Dextro	Dextromethorphan	1.0	2.5	5.0-10000	90.2	9.7	0.8	2.6	-4.8	-5.2	-4.8
		Dextrorphan	1.0	2.5		106.1	23.6	4.6	3.2	-17.9	11.1	-3.3
		(+)-3-MEM	2.5	5.0		102.5	19.7	6.1	4.2	10.4	-5.8	2.7
		(+)-3-HM	2.5	5.0		91.3	24.6	5.1	2.6	1.6	-9.9	1.5
							10 ng/mL			10 ng/mL		
	Levo	Levomethorphan	1.0	5.0	10-10000	94.6	10.9	9.5	2.6	-4.3	-17.0	-2.2
		Levorphanol	1.0	5.0		93.1	4.8	4.5	4.6	18.6	-8.0	6.8
				(ng/mg)		10 ng/mg	0.1 ng/mg	5.0 ng/mg	50 ng/mg	0.1 ng/mg	5.0 ng/mg	50 ng/mg
Hair (10 mg)	Dextro	Dextromethorphan	0.025	0.05	0.1-50	84.2	11.5	4.5	2.8	4.6	18.8	-6.6
		Dextrorphan	0.025	0.05		99.8	6.4	2.6	2.7	3.7	15.4	-3.5
		(+)-3-MEM	0.025	0.05		83.8	18.6	3.9	1.5	4.7	0.6	-2.2
		(+)-3-HM	0.025	0.1		91.4	11.2	6.2	2.8	4.6	18.8	-6.6
	Levo	Levomethorphan	0.025	0.1	0.1-50	98.1	9.9	9.8	5.5	0.1	-2.3	-5.1
		Levorphanol	0.025	0.05		112.2	8.8	2.8	4.2	11.9	1.3	-3.8

Linearity and precision of the analytical method for the rat urine, plasma, and hair samples

The calibration curves were linear over the concentration range 1.0–400 ng/mL for rat plasma, 5.0–10,000 ng/mL (compounds of dextro forms) and 10.0–10,000 ng/mL (compounds of levo forms) for rat urine, and 0.1–50 ng/mg for rat hair, with good correlation coefficients of $r^2 \geq 0.996$, respectively. The LOD of each drug was 0.4 or 0.8 ng/mL for the plasma, 1.0 or 2.5 ng/mL for the urine, and 25 pg/mg for the hair samples. The recoveries and the precision and accuracy data from the analytical procedures for the rat samples (n=5), spiked with a standard solution of the targeted compounds, are shown in Table 2.

Determination of dextromethorphan/levomethorphan and their metabolites in DA rat plasma, urine, and hair samples

It has been reported that a female DA rat lacks the CYP2D1 enzyme, which is known to be related to *O*-demethylation of dextromethorphan in the SD rat; it is therefore used as a model animal for the poor metabolizer phenotype of dextromethorphan [35–37]. As such, the metabolic data from female DA rats may not reflect the "normal" situation. On the other hand, pigmented hairy rats appear to be suitable for the investigation of analytical methods of basic drugs in hair samples, compared with albino rats (SD or Wistar rats) because pigmentation (the melanin contents) is one of the most important factors regarding the incorporation of basic drugs into hair, as described before [38]. Therefore, thus far, we have studied the analytical properties of various drugs in hair samples using the pigmented hairy male DA rats, avoiding female DA rats.

After the i.p. administration of dextromethorphan or levomethorphan to pigmented hairy male DA rats, the parent compounds and their three metabolites in the plasma, urine, and hair were determined using LC-MS/MS. The optical purities of the resulting metabolites were unchanged in any rat biological sample, and no racemation was observed through *O*- and/or *N*-demethylation (Fig. 3). In the rat plasma ($AUCs_{0-360\ min}$) and urine samples (total excretions for 0–72 h) after the hydrolysis of *O*-glucuronides, most metabolites were detected as being the corresponding *O*-demethyl and *N*, *O*-didemethyl compounds, as shown in Table 3. However, obvious differences in the amounts of these metabolites were found between the dextro and levo forms. After administration of dextromethorphan, dextrorphan and (+)-3-HM were the major metabolites in the plasma (59.4 and 64.3 mg/L·min) and urine (106.1 and 226.9 μg/mL). However, *O*-demethyl metabolites (levorphanol) were mainly detected in the plasma (197.1 mg/L·min) and urine (210.5 μg/mL) after administration of levomethorphan (Table 3).

In the hair samples, the differences in the amounts of the metabolites are more clearly detected. After administration of dextromethorphan, the parent compound and the *N*-demethyl metabolite ((+)-3-MEM) were mainly detected at 63.4 and 25.1 ng/mg, respectively, although the *O*-demethyl metabolite of dextromethorphan (dextrorphan) was detected at only 2.70 ng/mg, which was nearly one tenth of the level

Table 3 Rat plasma $AUCs_{0-360min}$, total excretion into rat urine, and concentrations in rat hair of dextromethorphan/levomethorphan and their metabolites

Administrations	Targeted compounds	Plasma $AUC_{0-360min}$ (mg/L·min)	Urine Total excretion 0–72 h (μg)	Hair Concentration (ng/mg)
Dextromethorphan (rat 1–3)	Dextromethorphan	23.8±17.6	2.13±1.05	63.4±4.6
	Dextrorphan	59.4±16.3	106.1±15.3	2.70±0.04
	(+)-3-MEM	3.10±2.15	6.95±0.68	25.1±1.9
	(+)-3-HM	64.3±31.3	226.9±51.3	0.70±0.11
Levomethorphan (rat 4–6)	Levomethorphan	6.90±5.12	0.59±0.61	24.5±5.3
	Levorphanol	197.1±48.2	210.5±36.2	24.6±2.4
	(−)-3-MEM	1.47±0.64	0.13±0.06	2.57±0.71
	(−)-3-HM	51.5±9.6	39.0±5.9	0.49±0.09

of levorphanol. In contrast, after the administration of levomethorphan, the parent compound and the *O*-demethyl metabolite (levorphanol) were mainly detected at 24.5 and 24.6 ng/mg, respectively, with a small amount of the *N*-demethyl metabolite ((−)-3-MEM). The *N*, *O*-didemethyl metabolites (3-HM) were hardly detected in either sample (Table 3). The ratios of the parent compounds, their *O*-demethyl, *N*-demethyl, and *N*, *O*-didemethyl metabolites in the hair samples were 100:4:40:1 for the dextro forms and 100:100:11:2 for the levo forms, respectively.

The rat plasma AUCs, total excretions into rat urine and concentrations in rat hair of dextromethorphan or levomethorphan, and their metabolites are summarized in Fig. 4. The metabolic ratios of dextromethorphan/levomethorphan, *O*-demethyl, *N*-demethyl, and *N*, *O*-didemethyl metabolites in rat plasma ($AUC_{0-360\ min}$) and hair (collected 4 weeks after the first administration) were 1:3:0.1:3 and 1:0.04:0.4:0.01 for the dextro forms and 1:29:0.2:7 and 1:1:0.1:0.02 for the levo forms, respectively. It is of interest that the concentrations of dextromethorphan and levome-

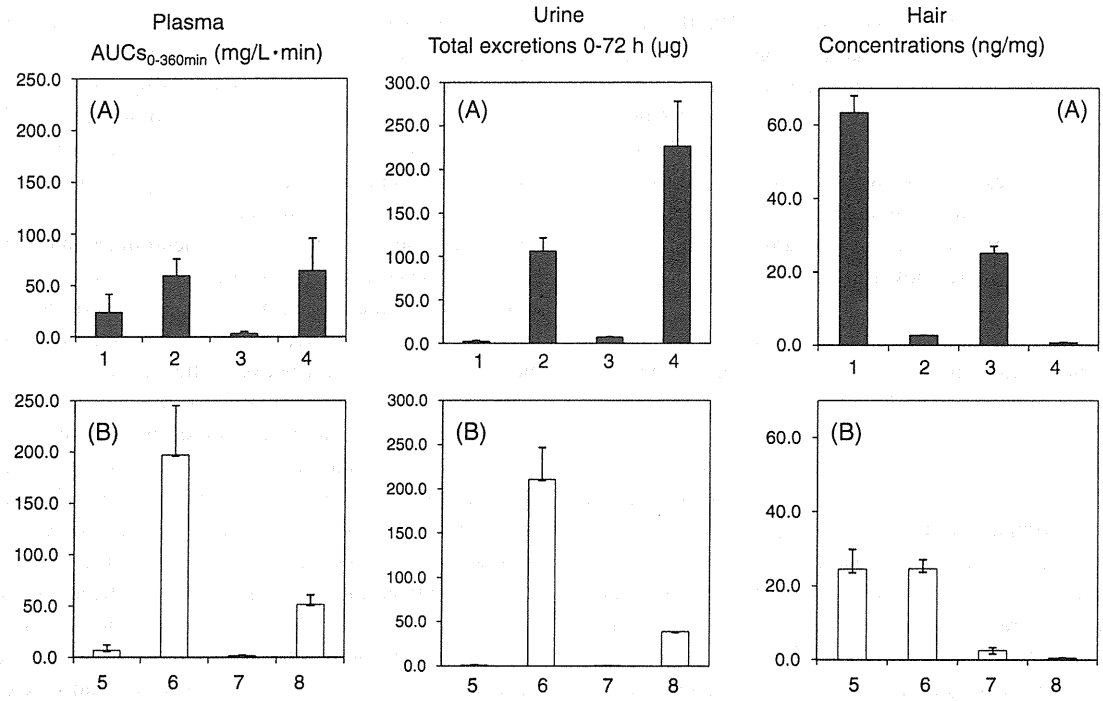


Fig. 4 Rat plasma $AUCs_{0-360min}$, total excretions into rat urine, and concentrations in rat hair of parent compounds and their metabolites after administration of (**A**) dextromethorphan and (**B**) levomethorphan. *1* Dextromethorphan, *2* Dextrorphan, *3* (+)-3-MEM, *4* (+)-3-HM, *5* Levomethorphan, *6* Levorphanol, *7* (−)-3-MEM, *8* (−)-3-HM

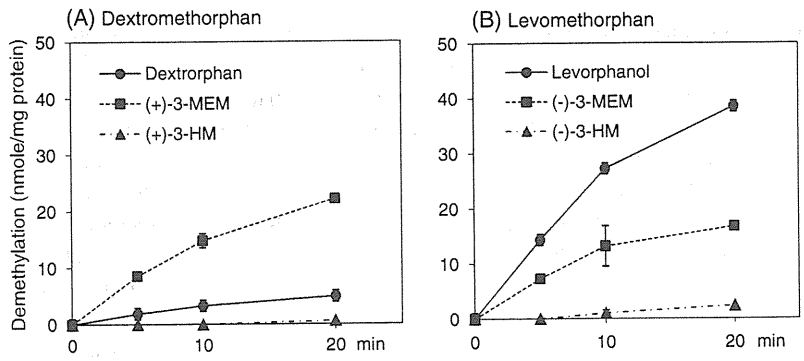


Fig. 5 Demethylation of (**A**) dextromethorphan and (**B**) levomethorphan in DA rat liver microsomes

thorphan in the rat hair were obviously high compared with those in the plasma, while those of their *O*-demethyl and *N*, *O*-didemethyl metabolites in the hair (which mostly existed as very hydrophilic metabolites, *O*-glucuronides in the plasma) were extremely low considering their high plasma AUCs.

Demethylation of dextromethorphan/levomethorphan in DA rat liver microsomes

In order to fully investigate the differences of the metabolic properties between dextromethorphan and levomethorphan, DA rat liver microsomes were studied. Figure 5 shows the *O*- and/or *N*-demethylation of dextromethorphan/levomethorphan in the rat liver microsomes.

The optical purities of the resulting metabolites were unchanged in the liver microsomes, and no racemation was observed through *O*- and/or *N*-demethylation. After 20-min incubation, 4.8% of dextromethorphan and 45% of levomethorphan were transformed to each *O*-demethyl metabolite, and 22% and 19% of the parent compounds were transformed to each *N*-demethyl metabolite. The *N*-demethylation was preferred over *O*-demethylation for dextromethorphan. In contrast, *O*-demethylation was preferred over *N*-demethylation for levomethorphan and the *O*-demethylation of levomethorphan was performed at levels 9.4 times that of dextromethorphan after 20-min incubation. The *N*-demethylation of levomethorphan was almost the same as that of dextromethorphan. Table 4 shows kinetic parameters for *O*-demethylation of dextromethorphan and levomethorphan by the DA rat microsomes. The V_{max} value for levomethorphan (3.8±0.3 nmol/min/mg protein) was 5.9 times higher than that of dextromethorphan (0.65±0.03 nmol/min/mg protein). The K_m values for levomethorphan and dextromethorphan were 22.1±5.0 and 44.1±4.0 μM, respectively. These results suggest that there might be an enantioselective *O*-demethylation of levomethorphan in the DA rat liver microsomes. This enantioselective metabolism might be the cause of the different amounts of the metabolites observed in the rat plasma, urine, and hair after administration of dextromethorphan and levomethorphan.

Demethylation of dextromethorphan/levomethorphan in pooled human liver microsomes

In order to investigate whether the enantioselective metabolism could be observed in humans as well as in DA rats, the pooled human liver microsomes were examined. Figure 6 shows the *O*- and/or *N*-demethylation of dextromethorphan/levomethorphan in the human liver microsomes.

The optical purities of the resulting metabolites were unchanged also in the human liver microsomes, and no racemation was observed through *O*- and/or *N*-demethylation. After 20-min incubation, 3.3% of dextromethorphan and 11% of levomethorphan were transformed to each *O*-demethyl metabolite and 2.5% and 7.1% of the parent compounds were transformed to each *N*-demethyl metabolite. The total amounts of the three metabolites from levomethorphan were higher than those from dextromethorphan in human (3.1 times) microsomes. Kinetic parameters for *O*-demethylation of dextromethorphan and

Table 4 Kinetic parameters for *O*-demethylation of dextromethorphan/levomethorphan by DA rat and human liver microsomes

	DA rat liver microsomes		Human liver microsomes	
	Dextromethorphan	Levomethorphan	Dextromethorphan	Levomethorphan
V_{max} (nmol/min/mg protein)	0.65±0.03	3.8±0.3[a]	0.26±0.03	0.58±0.02[a]
K_m (μM)	44.1±4.0	22.1±5.0[a]	4.5±0.9	8.9±1.7[a]

[a] Significantly different from dextromethorphan ($p<0.01$)

levomethorphan in the human liver microsomes are listed in Table 4. The V_{max} value for levomethorphan (0.58± 0.02 nmol/min/mg protein) was 2.2 times higher than that of dextromethorphan (0.26±0.03 nmol/min/mg protein). The K_m values for levomethorphan and dextromethorphan were 8.9±1.7 and 4.5±0.8 μM, respectively. There could also be an enantioselective metabolism of levomethorphan in human liver microsomes.

Discussion

In this study, we first investigated the analytical methods of dextromethorphan/levomethorphan and their metabolites in biological samples using DA male rats. As a result, chiral separation of dextromethorphan, levomethorphan, and their metabolites in biological samples was achieved in 12 min on a Chiral CD-Ph column. The optical purities of the resulting metabolites were unchanged in all rat biological samples, and no racemation was observed through *O*- and/or *N*-demethylation. The proposed chiral analyses might be applied to human samples and could provide useful information for discriminating dextromethorphan use from levomethorphan use, considering the possibility of the adulteration or substitution of dextromethorphan with levomethorphan for illegal purposes. However, for application to forensic toxicological purposes, further studies should be carried out using authentic human samples.

The concentrations of dextromethorphan and levomethorphan in the rat hair were obviously high compared with those of metabolites in the plasma and urine samples in this study. In our previous study [38], we determined the melanin affinity and lipophilicity of 20 abused drugs and these values were compared with the ratio of drug concentration in hair to plasma AUC as an index of the incorporation tendency into hair. As a result, the combination of melanin affinity (basicity) and lipophilicity showed a high correlation with the incorporation tendency into hair. Parent compounds can be detected relatively easily in hair in comparison with their hydrophilic metabolites. Actually, it has been reported that cocaine is detected in hair at a much higher concentration than its metabolite, benzoylecgonine, although cocaine is rapidly hydrolyzed to benzoylecgonine and disappears from plasma [39]. Considering those reports, the physico-chemical properties of dextromethorphan/levomethorphan and their metabolites could be significantly related to their concentrations in the hair samples. Additionally, the drug concentrations in the rat hair (collected 4 weeks after the first administration) reflected the total amounts of drugs in the plasma of rats administered with dextromethorphan/levomethorphan for ten successive days, and the differences might become more distinct. The detection of the parent compounds from hair samples would provide useful information regarding the monitoring of their use over a long period.

In the DA rat samples, obvious differences in the ratios of the metabolites were found between the dextro and levo forms. These differences were most clearly detected in the hair samples. The concentrations of the parent compounds, their *O*-demethyl, *N*-demethyl, and *N*, *O*-didemethyl metabolites were 63.4, 2.7, 25.1, and 0.7 ng/mg for the dextro forms and 24.5, 24.6, 2.6, and 0.5 ng/mg for the levo forms, respectively. In order to investigate the differences of their metabolic properties between dextromethorphan and levomethorphan, DA rat and human liver microsomes were studied. As a result, we have shown the enantioselective metabolism of levomethorphan, not only in DA rats but also in human liver microsomes, especially with regards to the *O*-demethylation. Because it is well-known that CYP2D6 (mainly related to *O*-demethylation of dextromethorphan) is polymorphically expressed in humans, it may be difficult to discuss the enantioselective metabolism in humans who can be classified as poor, intermediate and extensive metabolizers of dextromethorphan. In future studies, the metabolic properties of these drugs using CYP2D6 enzymes (having a variety of phenotypes) should be examined to clarify the effects of their genotypes on the enantioselective *O*-demethylation of levomethorphan observed in this study.

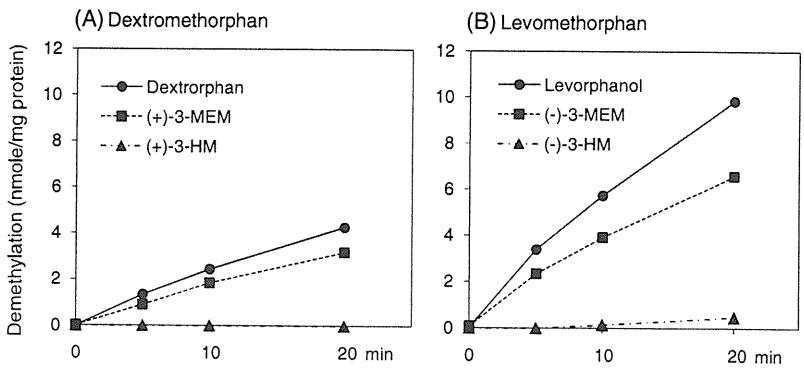


Fig. 6 Demethylation of (**A**) dextromethorphan and (**B**) levomethorphan in human liver microsomes

Conclusions

In this present study, we have established procedures for chiral analyses of dextromethorphan, levomethorphan, and their *O*-demethyl and/or *N*-demethyl metabolites in rat plasma, urine, and hair using LC-MS/MS. These analytical methods might be applied to human samples and could be useful for discriminating dextromethorphan use from levomethorphan use although further studies should be carried out using authentic human samples for forensic toxicological purposes. In addition, we have found the enantioselective metabolism of levomethorphan, not only in DA rats but also in human liver microsomes, especially with regards to the *O*-demethylation. This is the first report describing the differences in metabolic properties between dextromethorphan and levomethorphan in rats and humans.

Acknowledgments Part of this work was supported by a Health and Labor Sciences Research Grant from the Ministry of Health, Labor and Welfare in Japan.

References

1. Mutschler J, Koopmann A, Grosshans M, Hermann D, Mann K, Kiefer F (2010) Dtsch Arztebl Int 107:537–540
2. Chyka PA, Erdman AR, Manoguerra AS, Christianson G, Booze LL, Nelson LS, Woolf AD, Cobaugh DJ, Caravati EM, Scharman EJ, Troutman WG (2007) Clin Toxicol 45:662–677
3. Banken JA, Foster H (2008) Ann NY Acad Sci 1139:402–411
4. Shin EJ, Lee PH, Kim HJ, Nabeshima T, Kim HC (2008) J Pharmacol Sci 106:22–27
5. Bryner JK, Wang UK, Hui JW, Bedodo M, MacDougall C, Anderson IB (2006) Arch Pediatr Adolesc Med 160:1217–1222
6. Miller SC (2005) Addict Biol 10:325–327
7. Logan BK, Goldfogel G, Hamilton R, Kuhlman J (2009) J Anal Toxicol 33:99–103
8. Rammer L, Holmgren P, Sandler H (1988) Forensic Sci Int 37:233–236
9. Woods JW, Carney J (1978) NIDA Res Monogr 18:54–66
10. Trescot AM, Datta S, Lee M, Hansen H (2008) Pain Physician 11: S133–S153
11. Jacqz-Aigrain E, Funck-Brentano C, Cresteil T (1993) Pharmacogenetics 3:197–204
12. Schmid B, Bircher J, Preisig R, Küpfer A (1985) Clin Pharmacol Ther 38:618–624
13. Gorski JC, Jones DR, Wrighton SA, Hall SD (1994) Biochem Pharmacol 48:173–182
14. Köppel C, Tenczer J, Arndt I, Ibe K (1987) Arzneimittelforschung 37:1304–1306
15. Lutz U, Bittner N, Lutz RW, Lutz WK (2008) J Chromatogr B 871:349–356
16. Kristensen HT (1998) J Pharm Biomed Anal 18:827–838
17. Aumatell A, Wells RJ (1993) J Chromatogr Sci 31:502–508
18. Lin SY, Chen CH, Ho HO, Chen HH, Sheu MT (2007) J Chromatogr B 859:141–146
19. Bendriss EK, Markoglou N, Wainer IW (2001) J Chromatogr B 754:209–215
20. Afshar M, Rouini MR, Amini M (2004) J Chromatogr B 802:317–322
21. Hendrickson HP, Gurley BJ, Wessinger WD (2003) J Chromatogr B. B 788:261–268
22. Kim SC, Chung H, Lee SK, Park YH, Yoo YC, Yun YP (2006) Forensic Sci Int 161:185–188
23. Spanakis M, Vizirianakis IS, Mironidou-Tzouveleki M, Niopas I (2009) Biomed Chromatogr 23:1131–1137
24. Rodrigues WC, Wang G, Moore C, Agrawal A, Vincent MJ, Soares JR (2008) J Anal Toxicol 32:220–226
25. Kim EM, Lee JS, Park MJ, Choi SK, Lim MA, Chung HS (2006) Forensic Sci Int 161:198–201
26. Bagheri H, Es-haghi A, Rouini MR (2005) J Chromatogr B 818:147–157
27. Eichhold TH, McCauley-Myers DL, Khambe DA, Thompson GA, Hoke SH 2nd (2007) J Pharm Biomed Anal 43:586–600
28. Kuhlenbeck DL, Eichold TH, Hoke SH 2nd, Baker TR, Mensen R, Wehmeyer KR (2005) Eur J Mass Spectrom 11:199–208
29. Lutz U, Völkel W, Lutz RW, Lutz WK (2004) J Chromatogr B 813:217–225
30. Vengurlekar SS, Heitkamp J, McCush F, Velagaleti PR, Brisson JH, Bramer SL (2002) J Pharm Biomed Anal 30:113–124
31. Sunouchi M, Fukuhara M, Ohno Y, Takanaka A (1988) J Toxicol Sci 13:193–204
32. Ozawa S, Ohta K, Miyajima A, Kurebayashi H, Sunouchi M, Shimizu M, Murayama N, Matsumoto Y, Fukuoka M, Ohno Y (2000) Xenobiotica 10:1005–1017
33. Kikura-Hanajiri R, Kawamura M, Saisho K, Kodama Y, Goda Y (2007) J Chromatogr B 855:121–126
34. Kikura-Hanajiri R, Kawamura M, Miyajima A, Sunouchi M, Goda Y (2010) Forensic Sci Int 198:62–69
35. Bochner F, Somogyi AA, Chen ZR (1994) Xenobiotica 24:543–552
36. Kerry NL, Somogyi AA, Mikus G, Bochner F (1993) Biochem Pharmacol 45:833–839
37. Zysset T, Zeugin T, Küpfer A (1988) Biochem Pharmacol 37:3155–3160
38. Nakahara Y, Takahashi K, Kikura R (1995) Biol Pharm Bull 18:1223–1227
39. Nakahara Y, Ochiai T, Kikura R (1992) Arch Toxicol 66:446–449

Industrial Health 2011, 49, 559–565

Original Article

Expert Review for GHS Classification of Chemicals on Health Effects

Takeshi MORITA[1]* and Kaoru MORIKAWA[1]

[1]Division of Safety Information on Drug, Food and Chemicals, National Institute of Health Sciences, 1–18–1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

Received December 1, 2010 and accepted March 29, 2011
Published online in J-STAGE August 1, 2011

Abstract: Intoxication as a result of chemical accidents is a major issue in industrial health. The Globally Harmonized System of Classification and Labelling of Chemicals (GHS) provides a framework for hazard communication on chemicals using labelling or safety data sheets. The GHS will be expected to reduce the number of chemical accidents by communicating the hazards posed and prompting safety measures to be taken. One of the issues which may be a barrier to effective implementation of the GHS results from discrepancies in GHS classifications of chemicals across countries/regions. The main reasons are the differences in information sources used and in the expertise of people making the classification (Classifiers). The GHS requests expert judgment in a weight of evidence (WOE) approach in the application of the criteria of classification. A WOE approach is an assessment method that considers all available information bearing on the determination of toxicity. The quality and consistency of the data, study design, mechanism or mode of action, dose-effect relationships and biological relevance should be taken into account. Therefore, expert review should be necessary to classify chemicals accurately. However, the GHS does not provide any information on the required level of expertise of the Classifiers, definition of who qualifies as an expert, evaluation methods of WOE or data quality, and the timing of expert judgment and the need for updating/re-classification as new information becomes available. In this paper, key methods and issues in expert reviews are discussed. Examples of expert reviews and recommendations for harmonized classification are also presented.

Key words: GHS, Expert review, Weight of evidence, Data quality, Classification, Health effects

Introduction

Intoxication caused by chemicals, including organic solvents, is one of the major issues in industrial health. More than 50 selected case examples per year are reported by the Japanese Ministry of Health, Labour and Welfare to illustrate the need for prevention of chemical accidents, and the occurrence factors including lack of recognition of hazards, insufficient education of safety and health, and non-use of personal protective equipment, etc[1]. Implementation of the Globally Harmonized System of Classification and Labelling of Chemicals (GHS) will make improvements to these situations. The GHS is a scheme recommended by the United Nations issued in 2003, which aims to enhance the protection of human health and the environment by providing an internationally comprehensible system for hazard communication[2]. The classification and labeling of chemicals are key elements of industrial health to reduce the number of chemical accidents. Many efforts for implementation of the GHS are being made at national and international levels, since 2003. The efforts in Japan include issue of regulations (e.g., Revised Industrial Safety and Health Law), provisions of information for industries (e.g., Guidance on Consumer Product Risk Assessment for GHS Labeling, GHS Classification Guidance for Enterprises, or Support

*To whom correspondence should be addressed.
E-mail: morita-tk@nihs.go.jp

Tools for GHS Classification), education for GHS audiences (*e.g.*, holding of workshops or seminars, provision of education tools), or provision of GHS classification results of chemicals[3].

One of the issues on implementation of the GHS is discrepancies in GHS classifications of chemicals across countries/regions. As the GHS is a globally harmonized system, uniform GHS classification criteria are applied to each health hazard. However, different results of classification can be obtained for the same chemicals[4, 5]. The main reason is the difference in datasets (*i.e.*, information sources) used. The other reasons are adaptation of a building block approach[6] and the differences of expertise/experience of Classifiers[7]. The GHS is designed as a self-classification system, and it requires expert judgment in a weight of evidence (WOE) approach in the application of the criteria. The hazard classification process under the GHS is highly technical in nature, and it requires a certain background and level of expertise to perform it accurately. If a Classifier lacks understanding of the GHS classification criteria, the effort should be repeated by an expert or reviewed carefully before finalization of the results[7]. Therefore, an expert review process is important for GHS classification. However, the GHS does not provide any information on the necessary expertise of Classifiers, definition of the required expertise, evaluation methods for the WOE approach or data quality, and the timing of expert judgment. The authors have been involved in the review system for GHS classification projects for the Organization for Economic Co-operation and Development (OECD)[5] or in Japan[8]. Based on these experiences, key methods, examples, and recommendations on the application of expert review are presented in this paper.

Expert Review

Expert review for GHS classification in health effect is defined as evaluation process based on scientific evidence, expertise, experience, knowledge, and judgment in a WOE approach. Main targets of the evaluations are information sources, data quality, and WOE of the data. The experts should be people who have scientific knowledge, experimental skill and expertise in toxicology or industrial hygiene. They should understand well the classification criteria in the GHS and the regulatory sciences including test protocols. They should also recognize that the classification will be conducted based on hazard identification, not on risk assessment for humans.

Evaluation of information sources/datasets

One of the major factors of the different classifications for individual chemicals was the different sources used[4, 5]. Therefore, evaluation of information sources is an important factor for reliable classification. Experts know where to find the information necessary for classification and, more importantly, how to correctly interpret these data. Several types of information sources are available. These include review documents, peer-reviewed papers, industry based reports, abstracts, or databanks, etc. The most reliable source is international or national review documents in terms of the quality, availability and suitability of information that has to be used in decision making. Peer-reviewed papers and industry based reports have high quality and suitability, but low availability. Databanks have high availability, but low quality. Abstracts should not be used for classification without any supportive information. The age of the data differs among these sources. Newer information will be available from more recent documents, and this information could result in changed assessment of chemicals. Classification based on old or limited information will possess lower reliability. The evaluation of test results in each information source should be checked with multiple sources of information, if available. Original peer-reviewed papers are the best source for assessing difficult and comprehensive test results: these should be included in information collection, if possible.

Evaluation of data quality

Even for chemicals with testing data, inherent differences among test protocols and the interpretation of test data may confound hazard evaluation[9]. The determination of the quality of test data is a critical point for the classification. Therefore the evaluation of data quality has to be done by an expert. The evaluation of data quality includes assessment of three basic elements, *i.e.*, reliability, relevance and adequacy. Definitions of these terms are shown in Table 1[10].

In order to evaluate the reliability of the data, the following are examples of key points in an expert review[11]:

- Were the data obtained from the test using a standardized method (accordance with recent OECD test guideline or internationally recognized methods)?
- Was the test conducted in compliance with the principles of Good Laboratory Practice (GLP) or equivalent standards?
- Was purity or the physicochemical properties of the test chemical suitable for the test?
- Were the findings clear and plausible?
- Was the reporting information sufficient to make a judgment?

For regulatory purposes, a GLP study, in accordance

Table 1. Three basic elements of the evaluation of data quality

Element	Explanation
Reliability	Evaluating the inherent quality of a test report or publication relating to preferably standardised methodology and the way the experimental procedure and results are described to give evidence of the clarity and plausibility of the findings.
Relevance	Covering the extent to which data and tests are appropriate for a particular hazard identification or risk characterisation.
Adequacy	Defining the usefulness of data for hazard/risk assessment purposes. Where there is more than one study for each endpoint, the greatest weight is attached to the studies that are the most relevant and reliable.

Table 2. A scoring system to assess the reliability of toxicological data

Reliability of data	Explanation
Reliable without restrictions	Data generated according to generally valid and/or internationally accepted testing guidelines (preferably performed according to GLP) or in which the test parameters documented are based on a specific (national) testing guideline or in which all parameters described are closely related/comparable to a guideline method.
Reliable with restrictions	Data (mostly not performed according to GLP), in which the test parameters documented do not totally comply with the specific testing guideline, but are sufficient to accept the data or in which investigations are described which cannot be subsumed under a testing guideline, but which are nevertheless well documented and scientifically acceptable.
Not reliable	Data in which there were interferences between the measuring system and the test substance or in which organisms/test systems were used which are not relevant in relation to the exposure (e.g. unphysiological pathways of application) or which were carried out or generated according to a method which is not acceptable, the documentation of which is not sufficient for assessment and which is not convincing for an expert judgment.
Not assignable	Data which do not give sufficient experimental details and which are only listed in short abstracts or secondary literature (books, reviews, etc.).

with standardized methods, has a high level of reliability in toxicology. On the other hand, a research oriented study may be of low reliability. A scoring system to assess the reliability of toxicological data is shown in Table 2[10, 11].

Examples of key points of evaluation of the relevance of the data are as follows[11]:

- Was the study design suitable? It should include vehicle, animal species, route of administration, doses or concentration used, parameters examined, etc.
- Were there dose-effect relationships?
- Was the effect of statistical and biological significant?
- What test system was used (*e.g.*, *in vitro*, *in vivo*, or human)?

The level of relevance of toxicological findings will be higher usually in the following order: i) human data (meta-analysis, randomized controlled trial, case control study, cross-sectional study, and case report); ii) animal data (primate, rodent, other mammals, non-mammals); iii) *in vitro* data (mammalian cells, microorganisms, biochemical reactions).

Examples of key points of the evaluation of the data adequacy are as follows[11]:

- Recognition of the strengths and weaknesses of the test method (*e.g.*, sensitivity, specificity, accuracy).
- What was the key study?
- Was the finding supported by other data?
- What kind of mechanisms or mode of action was involved?

The above key points are important for *in vitro* data, *in silico* data, or human data. The level of adequacy of any toxicological findings will be higher usually in the following order: i) similar findings in more than single study; ii) the findings obtained with a validated test method; iii) the finding is supported by the other data; iv) single study; v) *in silico* data.

Evaluation of WOE among the data

Generally, three objectives of the WOE approach are suggested for regulatory decision-making: i) provision of a "clear and transparent framework" for evaluation of the evidence in risk determination; ii) offer of a consistent and standardized approach to evaluating toxic substances submitted to regulatory agencies; and iii) help of identification of the discretionary assumptions in risk determinations from experts[12–14]. The GHS defines WOE as follows[2]: "All available information bearing on the determination of toxicity is considered together, including the results of valid *in vitro* tests, relevant animal data, and human experience such as epidemiological and clinical studies and well-documented case reports and observations. Both positive and negative test results are assembled together in the weight of evidence determination. However, a single positive study performed according to good scientific principles and with statistically and biologically significant positive

results may justify classification." When multiple data for one endpoint exist, the WOE approach must be applied by experts. Toxicology experts must consider all available data (both positive and negative), weigh it with respect to validity, and finally reach a conclusion. In a WOE approach, quality and consistency of the data, study design, mechanism or mode of action, dose-effect relationships, reproducibility, biological relevance, strength of the evidence, and purity of the test substance should be taken into account. It is noticed that any discrepancy in classification will be based on the different weighting evidence used from expert to expert. Harmonization of expert judgment is not easy, and is not static[7].

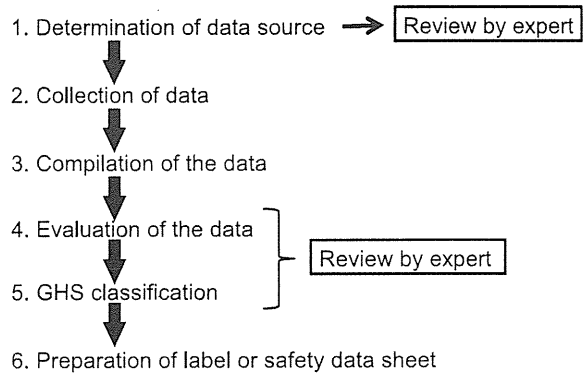


Fig. 1. Timing of expert review.

Issues in Expert Review

Expert review should promote and reflect the objective consideration of the full weight of evidence from alternative information sources, taking into account quality of data (*i.e.*, reliability, relevance and adequacy)[9]. The issues in expert review include: i) a single opinion from an expert might be low on transparency and high on subjectivity; ii) consistency of the judgment is unclear on between experts.

Transparency and objectivity

A WOE approach is one of key elements of expert review. Therefore, issues in applying it are also ones that might confound an expert review. The term WOE does neither constitute a scientifically well-defined term nor an agreed formalized concept characterized by defined tools and procedures. It is not clear which methods may be used, how they may be applied to the scientific evidence, what the results might be and how these may be used to make decisions in a specific hazard identification[14, 15]. The issues of a WOE approach in GHS classification include: i) application of WOE depends on the expertise of experts; ii) there are no canonical frameworks for weighting scientific evidence; iii) a process methodology is low on transparency and high on subjectivity; iv) WOE is usually applied in the case where there is no conclusive single study in demonstrating a cause-effect relationship; v) WOE looks like a 'seat-of-the pants' qualitative assessment. Without an explanation of how evidence is "weighed" or "weighted", the WOE approach may be to be a "black box" of scientific judgment[13]. To keep transparency of expert reviews, the review should be objective and taken into consideration of evidence based toxicology[16].

Consistency

Consistency of the results of expert review on similar subjects is a fundamental principal for hazard classification of chemicals. The outcome from an expert should be consistent in between experts and in chronological order if no new data or scientific evidence is available.

Timing of expert review

The practical application of an expert review (*i.e.*, how and when that review should be applied) is not mentioned in the GHS. This can result in ineffective hazard classification, hazard communication and chemical management[9]. Based on the experiences from the GHS classification projects, suitable timing of expert review is proposed (Fig. 1). General steps of GHS classification are as follows: Data collectors gather available data with a certain list. Classifiers read, select, compile and evaluate the information (*i.e.*, data), and then classify the chemical based on GHS criteria. Information gathering is the first step of hazard classification of chemicals. Different information source sets result in different classification results. It is important for GHS classification to establish useful and effective information source set. Experts should review suitability of the information sources at the beginning of GHS classification work, and should provide sufficient list of data sources. Examples of the sources are provided by Japan Ministry of Economy, Trade and Industry (METI)[17] or European Chemicals Agency[18]. Experts will review the necessity of additional source(s). Next timing of expert review is at time of evaluation of the data and following assignment of classification by Classifiers. Experts should review the relevance of the classification assigned by classifiers based on data quality of key studies and total weight of evidence of the findings[7].

Examples of Expert Reviews Where Re-classification Was Needed

Followings are examples of expert review in Japanese

GHS classification projects. Details of some examples are available from the web site of METI[19]. Examples for germ cell mutagenicity are given in a previous article[7].

Antimony trioxide [1309-64-4]

The original classification for this compound was Category 2B in regard to eye irritation, based on a mild irritation seen in rabbits[20]. An expert pointed out that draft EU Risk Assessment Report evaluates this compound as non-irritant based on the result from a new GLP study. The draft is not available yet, but the original test report is available through the OECD[21]. As the result is now non-irritating in the rabbit, "Not classified" was re-assigned in the review. This case suggests the importance of data collection.

4,4'-Thiobis(6-tert-butyl-m-cresol) (TBBC) [96-69-5]

The original classification was Category 1 in skin sensitization based on two patients with positive patch tests to TBBC who developed contact dermatitis to TBBC-containing latex gloves[22]. An expert questioned the reliability and relevancy of this information. The American Conference of Governmental Industrial Hygienists summarized that sufficient data were not available to recommend sensitization notation[22]. Therefore, "Classification not possible" was assigned by expert review. The point of debate was the evaluation of data quality.

p-Dichlorobenzene [106-46-7]

The original classification was Category 2 for germ cell mutagenicity based on a negative result from a dominant lethal test and a positive result in a micronucleus test[23]. However, a reviewing expert noticed that both positive and negative results existed for micronucleus tests of this compound. The positive result was not confirmed by additional tests including tests using a similar protocol to the first test. A positive result in a kidney micronucleus test was considered of low reliability and relevancy. Another 5 or more micronucleus tests showed negative results. Based on WOE, "Not classified" was assigned by the expert. The reasons for the changed classification were that multiple negative results had more weight than a single positive result and also an evaluation of data quality for the original positive result.

Styrene [100-42-5]

The original classification was Category 2 for carcinogenicity based on the classification in Group 2B by evaluation of International Agency for Research on Cancer[24]. A reviewing expert suggested that a recent analysis revealed that lymphatic and haematopoietic neoplasms seen in humans exposed to styrene are likely to be due to concomitant exposure to butadiene[25]. Mouse specific mode of action (MOA) exists in the induction of mouse lung tumor[26]. Opinions on the interpretation of the cancer data were different among experts. Finally, after much discussion resulted in a "Not classified" Category instead of Category 2. Thus the change in Classification resulted from the recent re-evaluation and analysis of the MOA.

Ethylene glycol [107-21-1]

The original classification was Category 1B for reproductive toxicity based on reduced skeletal ossification and malformations of the skeleton, which were observed without maternal toxicity[27]. A reviewing expert pointed out that the above effects were seen in rats at 1,500 mg/kg (over the limit dose of 1,000 mg/kg) or mice at 500 mg/kg. The expert introduced a recent evaluation document[28] which mentions that ethylene glycol is not directly responsible for developmental toxicity, but that this toxicity is due to the accumulation of glycolic acid (a metabolic breakdown of ethylene glycol). The saturation level of this compound is lower in humans than that in rodents. There is negligible concern (at current human exposure levels) for reproductive toxicity in humans. Therefore, "Not classified" was assigned by an expert review. The relevant points for the change in classification were the effective dose for toxicity shown in animal experiments which was not relevant to human exposure and the findings of different metabolism in human compared to rats for this compound.

Hydroquinone [123-31-9]

The original classification was Category 1B for reproductive toxicity based on an increase in foetal resorptions[29]. A reviewing expert pointed out that the above finding was based on old (1955–1964) studies. Recent evaluations generated negative results in rat and rabbit developmental tests and a rat two generation fertility test[30, 31]. These tests were conducted in accordance with recent guidelines, giving more reliability. Therefore, "Not classified" was assigned by the expert review for reproductive toxicity.

With respect to specific target organ toxicity (single exposure), the original classification was Category 1 (central nerve system and kidney) based on the appearance of tremor, vomiting and cyanosis which was observed in exposed humans and the observation of kidney damage in rats[29]. However, a reviewing expert pointed out that the human findings were based on exposure to mixtures containing hydroquinone plus other substances. Symptoms observed after exposure to

hydroquinone alone were transient central nerve system effects.

The rat is unique in susceptibility to kidney effects following hydroquinone exposure[31]. Based on the new evaluation document, Category 3 (narcotic effect) was assigned by an expert review. The relevant points for the change in classification were insufficient information gathering in the first instance, careful review of the documents and use of recent evaluation using more reliable data.

Ferric Chloride [7705-08-0]

The original classification was Category 1 for aspiration hazards based on the following finding[32]; a woman presented with vomiting after ingestion of 200 ml ferric chloride solution (pH 1.0). Three hours after her ingestion she presented with drowsy consciousness, tachycardia and protracted vomiting. Aspiration pneumonia was also noted. A reviewing expert noticed that the aspiration pneumonia was observed after vomiting of a corrosive solution, which does not necessarily indicate an aspiration hazard. In addition, the findings did not fit the GHS criteria for aspiration hazard. Therefore, "Classification not possible" was assigned by the expert review. The relevant points were recognitions of the definition in the GHS text and the effect by ingestion of a corrosive solution.

Recommendations

Consideration of different information sources can result in different GHS classification results. Judgments of data quality and weight given to findings will vary among experts. To minimize these variations, the following approaches will be needed for harmonized classification: i) Development of an internationally-constructed and maintained information database for GHS classification; ii) Provision of rationales of selection of (key) studies and a classification derived from them for maintaining transparency; iii) Discussion on how to apply expert judgment and how to assess the quality of data from limited studies; iv) Establishment of a GHS classification data bank which collects GHS classification results including related information; v) International review system of classification on specific chemicals; and vi) Consultation system for companies/Institutes without experts. These will be help the harmonization and transparency of GHS classifications. Duplication of classification will be also avoided.

Conclusions

It is clear that suitable classification depends on the correct interpretation of the data, the application of the weight of evidence approach and basing judgments only on high quality data. Toxicologists or industrial hygienists, as experts, play an important role in assigning supportable classifications. They should consider data quality, and should review critically several authoritative documents including original articles to support the classification of chemicals.

Acknowledgements

The authors are grateful to Dr. David Tweats (Swansea University, UK) for his review of manuscript. This work was supported in part by the Health and Labor Sciences Research Grants (H20-Labour-General-011 and H21-Chemistry-General-004).

References

1) Ministry of Health, Labour and Welfare (MHLW). Examples of chemical accidents (2005–2009). http://www.mhlw.go.jp/bunya/roudoukijun/anzeneisei10/. Accessed October 20, 2010 (in Japanese).
2) United Nations (2003) Globally Harmonized System of Classification and Labelling of Chemicals (GHS), New York and Geneva. Also available from: http://www.unece.org/trans/danger/publi/ghs/ghs_rev00/00files_e.html. Accessed October 20, 2010.
3) Jonai H (2008) Implementation of the GHS in Japan. Ind Health **46**, 443–7.
4) United Nations Institute for Training and Research (UNITAR). Implementation of GHS, Development of lists of classification, UN/SCEGHS/15/INF.32, Fifteenth session, Geneva, 9-11 July 2008. http://www.unece.org/trans/doc/2008/ac10c4/UN-SCEGHS-15-inf32e.pdf. Accessed October 20, 2010.
5) Organization for Economic Co-operation and Development (OECD). Outcome of the analysis of classification of selected chemicals listed in Annex III of the Rotterdam Convention, Nineteen session, Geneva, 30 June-2 July 2010. http://www.unece.org/trans/doc/2010/ac10c4/UN-SCEGHS-19-INF03.pdf. Accessed October 20, 2010.
6) United Nations Department of Economic and Social Affairs, Division for Sustainable Development. National Information, CSD-18-19 National Reports, Japan, Chemicals. http://www.un.org/esa/dsd/dsd_aofw_ni/ni_pdfs/NationalReports/japan/Chemicals.pdf. Accessed October 20, 2010.
7) Morita T, Hayashi M, Nakajima M, Tanaka N, Tweats DJ, Morikawa K, Sofuni T (2009) Practical issues on the application of the GHS classification criteria for germ cell mutagens. Regul Toxicol Pharmacol **55**, 52–68.
8) National Institute of Technology and Evaluation (NITE). Information about the status of the implementation

of GHS in Japan. http://www.safe.nite.go.jp/english/ghs_index.html#results. Accessed October 20, 2010.
9) Hamilton JD, Daggett DA, Pittinger CA (2006) The role of professional judgment in chemical hazard assessment and communication. Regul Toxicol Pharmacol **46**, 84–92.
10) European Chemical Agency (ECHA). Guidance on information requirements and chemical safety assessment, Chapter R.4: Evaluation of available information, May 2008. http://guidance.echa.europa.eu/docs/guidance_document/information_requirements_r4_en.pdf?vers=20_08_08. Accessed October 20, 2010.
11) Klimisch H, Andreae M, Tillmann U (1997) A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regul Toxicol Pharmacol **25**, 1–5.
12) Walker VR (1996) Risk characterization and the weight of evidence: adapting gatekeeping concepts from the courts. Risk Analysis **14**, 793–9.
13) Krimsky S (2005) The weight of scientific evidence in policy and law. Am J Public Health **95**, S129–36.
14) Weed DL (2005) Weight of evidence: a review of concept and methods. Risk Analysis **25**, 1545–57.
15) Vermeire TG, Aldenberg T, Dang Z, Janer G, de Knecht JA, van Loveren H, Peijnenburg WJGM, Piersma AH, Traas TP, Verschoor AJ, van Zijverden M, Hakkert B (2007) Selected integrated strategies (ITS) for the risk assessment of chemicals. RIVM-report 601050001/2007. http://www.rivm.nl/bibliotheek/rapporten/601050001.html. Accessed October 20, 2010.
16) Hoffmann S, Hartung T (2006) Toward an evidence-based toxicology. Human Experi Toxicol **25**, 497–513.
17) Ministry of Economy, Trade and Industry (METI). GHS Classification Guidance for the Japanese Government or Enterprises, March 2009. http://www.meti.go.jp/policy/chemical_management/int/ghs_tool_01GHSmanual.html. Accessed October 20, 2010.
18) European Chemicals Agency (ECHA). Guidance on information requirements and chemical safety assessment, Chapter R.3: Information gathering, May 2008. http://guidance.echa.europa.eu/docs/guidance_document/information_requirements_r3_en.pdf?vers=20_08_08. Accessed October 20, 2010.
19) Mitsubishi Chemical Techno-Research (2008) Report of investigation on GHS classification of PRTR chemicals, March 2008. http://www.meti.go.jp/policy/chemical_management/03kanri/H19%20GHS_report(re-classification).pdf. Accessed October 20, 2010 (in Japanese).
20) Chemicals Evaluation and Research Institute (CERI). Chemical hazard data sheet, 2001-7, Antimony trioxide. http://qsar.cerij.or.jp/SHEET/F2001_07.pdf. Accessed October 20, 2010 (in Japanese).
21) LPT Laboratory of Pharmacology and Toxicology KG (2005) Acute eye irritation/corrosion test of antimony trioxide in rabbits, LPT Report No.19227/05, October 12, 2005.
22) American Conference of Governmental Industrial Hygienists (2001) Documentation of the threshold limit values for chemical substances, 4,4'-Thiobis(6-tert-butyl-m-cresol). ACGIH, Cincinatti.
23) European Chemicals Bureau (2004) EU Risk Assessment Report, 31, p-Dichlorobenzene. http://ecb.jrc.ec.europa.eu/DOCUMENTS/Existing-Chemicals/RISK_ASSESSMENT/REPORT/14dichlorobenzenereport001.pdf. Accessed October 20, 2010.
24) International Agency for Research on Cancer (2002) IARC Monographs on the evaluation of the carcinogenic risks to humans, 82, Some Traditional Herbal Medicines, Some Mycotoxins, Naphthalene and Styrene. IARC, Lyon, France.
25) European Chemicals Bureau (2007) EU Draft Risk Assessment Report, Styrene, R034_0711_hh November 2007. http://ecb.jrc.it/DOCUMENTS/Existing-Chemicals/RISK_ASSESSMENT/DRAFT/R034_0001_0711_env_hh.pdf. Accessed October 20, 2010.
26) International Program on Chemical Safety (IPCS). IPCS Conceptual Framework for Evaluating a Mode of Action For Styrene Induced Mouse Lung Tumour, ECBI/19/06 Add. 2. Also available from: http://www.docstoc.com/docs/27315138/Mode-of-Action---styrene. Accessed October 20, 2010.
27) International Program on Chemical Safety (2002) Concise International Chemical Assessment Document, 45, Ethylene glycol. http://www.who.int/ipcs/publications/cicad/en/cicad45.pdf. Accessed October 20, 2010.
28) US National Toxicology Program (2004) NTP-CERHR Monograph on the Potential Human Reproductive and Developmental Effects of Ethylene Glycol, NIH Publication No. 04-4481, January 2004. http://cerhr.niehs.nih.gov/evals/egpg/ethylene/EG_Monograph.pdf. Accessed October 20, 2010.
29) International Program on Chemical Safety (1994) Environmental Health Criteria, 157, Hydroquinone. http://www.inchem.org/documents/ehc/ehc/ehc157.htm. Accessed October 20, 2010.
30) Bingham E, Cohrssen B, Powell CH (Eds.) (2001) Patty's Toxicology, 5th Ed., Vol. 4, Hydroquinone, 407–23, John Wiley and Sons, New York.
31) Organization for Economic Co-operation and Development (OECD). Screening Information Data Set (SIDS) Initial Assessment Profile, Hydroquinone. http://www.chem.unep.ch/irptc/sids/OECDSIDS/123319.pdf. Accessed October 20, 2010.
32) Hazardous Substance Data Bank. Ferric chloride, HSDB Number 449, Last revision data, May 2005. http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB. Accessed October 20, 2010.

Mutation Research 741 (2012) 32–56

Contents lists available at SciVerse ScienceDirect

Mutation Research/Genetic Toxicology and Environmental Mutagenesis

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

Community address: www.elsevier.com/locate/mutres

Effect of reducing the top concentration used in the *in vitro* chromosomal aberration test in CHL cells on the evaluation of industrial chemical genotoxicity

Takeshi Morita[a,*], Masamitsu Honma[b], Kaoru Morikawa[a]

[a] National Institute of Health Sciences, Division of Safety Information on Drug, Food and Chemicals, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan
[b] National Institute of Health Sciences, Division of Mutagenesis and Genetics, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

ARTICLE INFO

Article history:
Received 3 June 2011
Received in revised form
22 September 2011
Accepted 6 October 2011
Available online 17 October 2011

Keywords:
In vitro tests, Chromosomal aberration test, CHL cells, Test concentration limit, Industrial chemicals

ABSTRACT

A current concern with *in vitro* mammalian cell genotoxicity testing is the high frequency of false or misleading positive results caused in part by the past use of excessively high test concentrations. A dataset of 249 industrial chemicals used in Japan and tested for genotoxicity was analyzed. Of these, 116 (46.6%) were positive in the *in vitro* chromosomal aberration (CA) test, including 6 that were positive only at test concentrations >10 mM. There were 59 CA-positive chemicals at test concentrations ≤1 mM. At >1 mM, 51 chemicals were CA-positive, including 13 Ames-positive chemicals, which were therefore not "missed" by the test battery. Thus, 38 potentially positive chemicals would not have been detected in the test battery if the top test concentration was limited to 1 mM in CA test. Analysis of the relevance of CA results on the 38 missed chemicals was conducted based on a weight of evidence approach, including evaluations of effects of extreme culture conditions (low pH, high toxicity, or precipitation), *in silico* structural alert analysis, *in vivo* genotoxicity and carcinogenicity test data (where available), mode of action, or information from closely related chemicals. After an exhaustive review, there were four chemicals with some concern for human health risk assessment, nine with minimal concern, and the remaining 25 with negligible concern. We apply different top concentrations to the 38 missed chemicals to identify the most accurate approach for predicting the genotoxicity of industrial chemicals. Of these 2 mM or 1 mg/mL, whichever is higher, was the most effective in detecting these chemicals, *i.e.*, relatively higher (8/13) or lower (17/25) detection among 13 chemicals with some or minimal concern, or 25 with negligible concern, respectively. Lower top concentration limits, 1 mM or 0.5 mg/mL, whichever is higher, are not as effective (2/13) for detecting these chemicals with concern. Therefore, we conclude 2 mM or 1 mg/mL, whichever is higher, would be an appropriate top concentration limit for testing industrial chemicals for chromosome damage.

1. Introduction

The top concentrations used in *in vitro* mammalian cell genotoxicity tests are currently being re-examined with the goal of reducing the frequency of false or misleading positive results [1–6]. In the standard test battery, the mouse lymphoma assay and the *in vitro* micronucleus (MN) or chromosomal aberration (CA) test have low specificity for predicting carcinogenicity (e.g., <45%) [7–9], and the false positive results they generate lead to additional testing, often requiring the unnecessary use of animals [4]. The current top concentration limit specified in regulatory guidelines is 10 mM or 5 mg/mL, whichever is lower, when not limited by solubility or cytotoxicity [10,11]. The possible reasons for misleading or non-relevant results caused by testing at very high concentrations *in vitro* are: (1) un-physiological culture conditions including low pH, high osmolality and/or precipitation; (2) excessive cellular metabolic turn over, activation and defense/stress processes; and (3) results obtained at high concentrations that could not be reached *in vivo* and therefore not confirmed in *in vivo* genotoxicity or carcinogenicity tests. Therefore, more biologically relevant experimental conditions are needed. One way to lessen false positive results is to reduce the top test concentration [1]. The proposed revised International Conference on Harmonisation (ICH) test guidelines for pharmaceuticals recommends as the top concentration 1 mM or 0.5 mg/mL, whichever is lower [12], and almost same conclusion, *i.e.*,1 mM or 0.5 mg/mL, whichever is higher, was reached following an analysis of 384 genotoxic rodent carcinogens [3] and the retesting of selected chemicals [4]. A consensus for reducing the top concentration for testing in mammalian genotoxicity tests *in vitro* from 10 mM (but no agreement to what concentration should be) was reached in 2009 at the 5th International Workshop on Genotoxicity Testing, where the results from our preliminary analysis of 249 chemicals were presented [5]. Earlier in 2006, participants in a European Centre for the Validation of Alternative Methods workshop proposed that the published and industry data should

* Corresponding author. Tel.: +81 3 3700 1141; fax: +81 3 3700 1483.
E-mail address: morita-tk@nihs.go.jp (T. Morita).

1383-5718/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.mrgentox.2011.10.005

be reviewed to determine whether the top test concentration should be lowered [1]. Therefore, our analysis will provide useful information for this debate.

As a member of the Organization for Economic Co-operation and Development (OECD) since 1991, Japan has been performing safety tests of high production volume (HPV) chemicals. The results of the 249 chemicals analyzed in this paper as part of that effort were published as hard copies [13–26], while recent results are available only online (http://dra4.nihs.go.jp/mhlw_data/jsp/SearchPageENG.jsp). The dataset includes results of the *in vitro* CA test and the Ames test, both conducted in accordance with OECD or Japanese test guideline and according to Good Laboratory Practice (GLP) regulations. In this paper, we report our analysis of "missed" chemicals—that is, those that were positive in the *in vitro* CA test at the 10 mM top concentration but not at 1 mM and were negative in the Ames test. We discuss the significance and relevance of the induction of CAs by these chemicals based on the weight of evidence approach including *in silico* analysis, review of literature on *in vivo* genotoxicity and carcinogenicity, or effects in structurally related chemicals. The levels of concern for human health risk assessment are defined for each chemical in this paper. Finally, we apply different top concentrations to missed chemicals to identify the most accurate approach for predicting the genotoxicity of industrial chemicals.

2. Materials and methods

2.1. Source of published in vitro CA test results

We obtained *in vitro* CA test data from reports published from 1994–2006 by the Ministry of Health, Labor, and Welfare of Japan [13–26].

2.2. In vitro CA test protocol

In vitro CA tests were conducted in Chinese hamster lung (CHL) cells in accordance with OECD test guideline no. 473 [11] or the Japanese test guideline for new chemicals [27] under GLP conditions. The outline of the protocol under the former Japanese test guideline were as follows: Treatment length was 6 h (with 18 h recovery), 24 h (continuous without recovery time), or 48 h (continuous without recovery time) without S9 mix, a rat liver homogenate microsomal fraction with co-factors for metabolic activation, and 6 h (with 18 h recovery) with S9 mix. The top concentration was 5 mg/mL (or equivalent of 10 mM) when no cytotoxicity was observed. In the presence of cytotoxicity, the top concentration selected was one that caused 50% or greater inhibition of cell growth compared to the negative control. Methods for measuring cytotoxicity, as relative cell growth, were an estimation of monolayer confluence using a monocellater or other equipment, or survival cell counts. Structural aberrations and polyploidy were evaluated independently in 100 or 200 metaphases per concentration. In some cases, 800 metaphases per concentration were analyzed for polyploidy. Pre-1997 studies did not follow the current OECD guideline that was published that year. Major differences from the current OECD guideline are use of long exposure time (48-h continuous treatment), use of concentrations which shows much greater than 50% relative cell growth as the top concentration when cytotoxicity was observed. More preferably 5 mg/mL was used, rather than 10 mM as the top concentration when cytotoxicity was not observed. In general, there was no consideration of physiological culture conditions (pH, osmolality, or precipitation), no concurrent cytotoxicity measurement, no ensuring that at least 200 metaphases were analyzed per concentration, and no statisticalanalyses were carried out [27,28].

2.3. Reporting of results

Each experiment was classified as (a) positive (+): ≥10% cells with CAs; (b) equivocal (?): ≥5–10% cells with CAs, or (c) negative (−): less than 5% cells with CAs. Then only chemicals showing at least one positive or equivocal experiment were considered positive or equivocal compound, respectively, in the original reports [13–26]. Statistical significance, reproducibility, culture conditions, or concentration relationship of the response were taken into consideration in some cases. Basically, calls of the classification of chemicals were based on the original calls. However, there are some exceptions to this analysis. If a chemical was classified equivocal in the original "call" and showed a reproducibility and/or concentration-related response with statistical significance, the chemical was considered positive. For example, two chemicals (Identifications (IDs) 61, 82) assigned equivocal in the original call were considered positive, because the effect was reproducible or there was a CA-induction of equal to or more than 10%. Three chemicals (IDs 81, 83, 84) assigned negative in the original call were also considered positive, because CA-induction of them was equal to or more than 10%; the original reports judged the CA-induction was due to low pH (see Section 3.1). Chemicals that show negative and/or equivocal (without reproducibility or concentration-related response) results were also considered as negative compounds.

Note that the percentages of cells with CAs refer to structural aberrations and do not include polyploidy. The percentages of polyploidy cells are presented in two chemicals (IDs 95, 96) in this analysis.

2.4. Analysis of the data

The different steps of the analysis (weight of evidence approach) used in this paper for the 249 chemicals on which *in vitro* CA tests were conducted, are shown in Fig. 1: (1) analysis of *in vitro* CA data excluding 48 h results; (2) classification of chemicals into positive and negative compounds; (3) for positive chemicals exclusion of those with the lowest effective concentration (LEC) of ≤1 mM or >10 mM; (4) for chemicals with LEC of >1–10 mM, review of Ames test data; if positive, chemicals would not be missed; (5) if Ames-negative (*i.e.*, possible "missed" chemicals in the test battery), further evaluation of the relevance of CA results, including evaluation of effects of extreme culture conditions (low pH, high toxicity and precipitation), *in silico* analyses using Deductive Estimation of Risk from Existing Knowledge (DEREK) for Windows and/or the Optimized Approach Based on Structural Indices Set (OASIS) tissue metabolic simulator (TIMES), and review of the literature to see if more *in vivo* genotoxicity and carcinogenicity data including on structurally related compounds can be found; and (6) conclusion on level of concern for human health risk assessment on missed chemicals.

2.5. In silico structure alert analysis

We used DEREK for Windows (version 12) for structure alerts for mutagenicity, clastogenicity, and carcinogenicity [29] and TIMES (version 2.26.3) for structural alerts for clastogenicity [30]. TIMES can predict CAs induced by metabolically activated chemicals that do not elicit activity in the parent form, in addition to alerting for mutagenic structures [30–33]. Basically, DEREK was applied to all "missed" chemicals. TIMES was also applied missed chemicals with the exception of chemicals which were considered positive due to possible effects of extreme culture conditions.

2.6. Literature search

For the literature search, we used *PubMed* and *TOXNET* and searched for "CAS number", "carcinogenicity", "genotoxicity", "mutagenicity", "micronucleus" and "chromosomal aberration". Data from structurally related chemicals were also taken into account. We also searched Screening Information Data Set (SIDS) documents in the United Nations Environment Programme (UNEP, http://www.chem.unep.ch/irptc/sids/OECDSIDS/sidspub.html) or the OECD chemical database (http://webnet.oecd.org/hpv/ui/Search.aspx) [34–53].

2.7. Level of concern

The level of concern for human health risk assessment on "missed" chemicals was defined based on the analysis by weight of evidence approach. General criteria are as follows: (1) negligible concern, negative result(s) in *in vivo* genotoxicity or carcinogenicity test, clear evidence(s) of irrelevancy (e.g., extreme culture condition) for CA-induction, and/or mode of action of non-DNA target; (2) minimal concern, some evidence(s) of irrelevancy of CA-induction or of increasing level of negligible concern, or negative result(s) in *in vivo* genotoxicity tests with some limitations; (3) some concern, positive result(s) in *in vivo* genotoxicity or carcinogenicity test, or no supporting evidence(s) for reducing the level of concern.

2.8. Application of different top concentrations to the "missed" chemicals

Several top concentration limits were applied to the "missed" chemicals in order to investigate their effectiveness for predicting the genotoxicity of industrial chemicals, which include 1 mM or 0.5 mg/mL, whichever is higher; 2 mM or 1 mg/mL, whichever is higher; 4 mM or 2 mg/mL, whichever is lower; and 10 mM or 2 mg/mL, whichever is lower. These top concentrations except for 2 mM or 1 mg/mL, whichever is higher, are under discussion by an OECD expert group (unpublished). The number of chemicals detected among missed chemicals with minimal or some concern, or with negligible concern was calculated.

3. Results

3.1. Analysis of in vitro CA and Ames test data

At first, we regarded two chemicals, benzyltrimethylammonium chloride (ID61, Table 2) and glycerol triacetate (ID82, Table 4), as positive in this analysis; their original "call" were equivocal [16,18], but the effects observed by benzyltrimethylammonium chloride or glycerol triacetate were reproducible or induced more than 10%

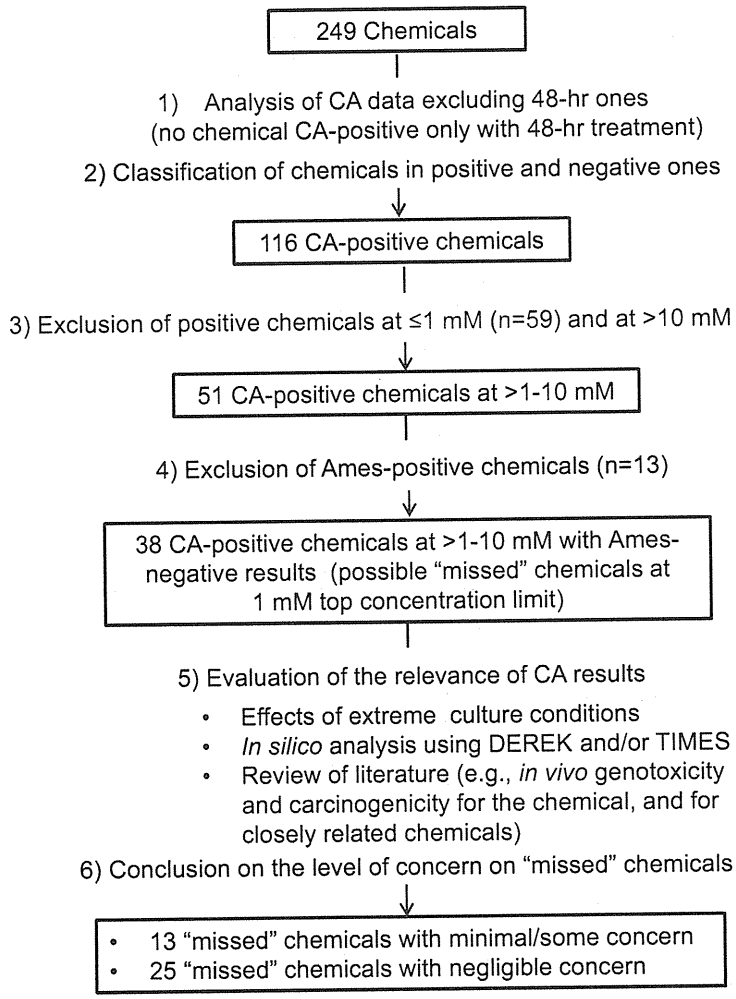


Fig. 1. Flow chart of the different steps of the analysis (weight of evidence approach).

cells with CAs, respectively. CA-induction by the latter chemical was considered to be due to low pH, so equivocal was assigned in the original report [18]. In addition, we regarded 3 chemicals, 2-amino-5-methylbenzenesulfonic acid (ID81), 4-hydroxybenzoic acid (ID83), and methyl acetoacetate (ID84) (Table 4), as positive in this analysis; their original "call" were negative, but the frequencies of CAs observed by the 3 chemicals were more than 10%. Clastogenicity induced by them was considered to be due to low pH, and neutralization of culture medium gave negative results. Thus a negative call was assigned for these chemicals in the original reports [16–18,54].

Among the 249 chemicals tested that had been subjected to the *in vitro* CA test, 116 (46.6%) were positive (Tables 1–4), and 133 (53.4%) were negative (Table 5). No chemicals were CA-positive only with a 48-h treatment. Almost all chemicals were also Ames tested, although we took some Ames data from other literature sources including US National Toxicology Program (NTP, http://ntp-apps.niehs.nih.gov/ntp_tox/) [55–57]. The Ames data are included in Tables 1–5.

Based on the steps of the analysis (Fig. 1), the 249 chemicals in the *in vitro* CA test were divided into five groups; (1) 59 chemicals positive at ≤1 mM (Table 1); (2) six chemicals positive at >10 mM which are considered negative following the criteria given in the current guidelines (Table 2); (3) 13 chemicals positive at >1–10 mM with Ames-positive results which were not "missed" by the test battery (Table 3); (4) 38 chemicals positive at >1–10 mM with Ames-negative results which would be missed in top concentration of 1 mM (possible "missed" chemicals, Table 4); and (5) 133 chemicals negative in the *in vitro* CA test (Table 5).

3.2. Evaluation of the relevance of in vitro CA results

Thirty eight chemicals were chemicals that would be missed if 1 mM was employed as the top concentration limit (*i.e., in vitro* CA negative at >1–10 mM and Ames-negative). The relevance of *in vitro* CA results was evaluated based on a weight of evidence approach including analysis of effects of extreme culture conditions (low pH, high toxicity, and precipitation), *in silico* structural alert analysis using DEREK and/or TIMES, and review of literature for *in vivo* genotoxicity and carcinogenicity tests and the genotoxicity/carcinogenicity of closely related chemicals (Table 4).

On measuring cytotoxicity, there were cases where even though there was apparently, for example 50% relative cell growth, measured by cell counts or confluence, there were insufficient mitotic cells to score.

3.2.1. Possible effects of extreme culture conditions (15 chemicals)

3.2.1.1. Low pH (seven chemicals). A low pH effect was defined as responsible for CA induction when the medium pH was 6.0 or below at the beginning of or just after treatment. Morita et al. reported that initial pH 6.2 or below in 6-h treatment with S9 mix, and initial pH 5.5 or below in 24-h treatment without S9 mix, were clastogenic