Annex VII

Analyses Using Multiple SI Decision Criteria

This page intentionally left blank

1.0 Introduction

This annex provides analyses associated with using two decision criteria for classifying substances using the results from the murine local lymph node assay modified by Daicel Chemical Industries, Ltd., based on ATP content (LLNA: DA); one criterion to classify substances as sensitizers and another criterion to classify substances as nonsensitizers. The data used for the analyses in this annex are the LLNA: DA results for the 44 substances (32 traditional murine local lymph node assay [LLNA] sensitizers and 12 traditional LLNA nonsensitizers) that were reviewed by the independent peer review panel at the public meeting on April 28-29, 2009. **Section 2** of this annex discusses the accuracy produced by using the two decision criteria and includes an evaluation of discordant, or indeterminate, substances that produced stimulation index (SI) values between the sensitizer and nonsensitizer SI criteria. **Section 3** provides the reproducibility analysis using the decision criterion for sensitizers (**Sections 3.1** and **3.2**) and for tests yielding SI values in three categories: sensitizer, nonsensitizer, and indeterminate (i.e., in the range of uncertainty) (**Section 3.3**). The two SI values determined to be optimal were based on four animals per dose group, and resulted in nine substances that could not be definitively classified because they produced SI values in the range of uncertainty. **Section 4** describes the impact of sample size on the range of the uncertainty between the sensitizer and nonsensitizer criteria. **Section 5** evaluates a number of physicochemical characteristics and other parameters to distinguish between traditional LLNA sensitizers and nonsensitizers in the LLNA: DA, when using multiple SI decision criteria, for potential use in providing additional information for classifying substances that produce SI values in the range of uncertainty.

2.0 Accuracy Analysis Using Multiple Stimulation Index Decision Criteria

As detailed in Section 6.5 of the background review document (BRD), the accuracy of the LLNA: DA when using various single alternative decision criteria was evaluated using the traditional LLNA as the reference test. Compared to the traditional LLNA (SI ≥ 3.0), the optimum performance (accuracy of 93% [41/44] and sensitivity of 100% [32/32]) was achieved using the decision criterion of SI ≥ 1.8 (Table C-8 of the BRD). Although the SI ≥ 1.8 produced a false positive rate of 25% (3/12) it yielded a false negative rate of 0% (0/32) (Table C-8 of the BRD). Increasing the SI decision criterion to SI ≥ 2.5 decreased the false positive rate to 0% (0/12) but increased the false negative rate to 13% (4/32). The 0% false positive rate using SI ≥ 2.5 and the 0% false negative rate using SI ≥ 1.8 prompted an evaluation using two SI decision criteria for determining LLNA: DA results: one criterion to classify substances as sensitizers (SI ≥ 2.5) and one criterion to classify substances as nonsensitizers (SI ≤ 1.8). The evaluation of this accuracy analysis is described below.

It should be noted that this analysis was based on the same strategy for combining results as that described in Section 6.5 of the BRD for the substances tested multiple times (i.e., the sensitizer/nonsensitizer outcome for each substance using the most prevalent outcome). **Section 3.0** details the reproducibility of substances tested multiple times and indicates that, there were no instances of false positive results for nonsensitizers (0% [0/80] of the substances classified as traditional LLNA nonsensitizers had an SI ≥ 2.5 in the LLNA: DA). See **Section 3.0** for more details regarding these results.

2.1 Indeterminate Results Using Multiple Stimulation Index Decision Criteria

While optimum false positive and false negative rates can be achieved for the 44 substances evaluated in the LLNA: DA accuracy analyses using these two different decision criteria, a range of SI values (i.e., between 1.8 and 2.5) exists for which the correct classification is not definitive (i.e., there is a chance for false positive or false negative results for substances that produce SI values in this range). Chemical class, physical form, molecular weight, peptide reactivity (see Annex II of the BRD for physicochemical properties), traditional LLNA EC3 range (estimated concentration needed to

produce a stimulation index of 3) (Table C-2 of the BRD), or potential for skin irritation (Annex III of the BRD) were examined to identify commonalities among the substances that produced SI values between 1.8 and 2.5 in an attempt to identify similar characteristics among these substances that could be used to correctly classify such substances. **Section 5.0** of this annex provides a comprehensive evaluation of a number of physicochemical characteristics and other parameters using the LLNA: DA database to distinguish between traditional LLNA sensitizers and nonsensitizers.

Of the nine substances that produced SI values between 1.8 and 2.5 (**Table C-VII-1)**, four are nonsensitizers (chlorobenzene, hexane, isopropanol, salicylic acid) and five are sensitizers (3-aminophenol, cobalt chloride, 2-mercaptobenzothiazole, methyl methacrylate, and nickel [II] sulfate hexahydrate) based on traditional LLNA results. Among the four traditional LLNA nonsensitizers, six chemical classes are represented; one substance is classified as both a carboxylic acid and phenol (salicylic acid), one substance is both a halogenated and a cyclic hydrocarbon (chlorobenzene), one substance is an acyclic hydrocarbon (hexane), and one substance is an alcohol (isopropanol). Other characteristics of the nonsensitizers (based on traditional LLNA data) include:

- Three substances are liquids (chlorobenzene, hexane, and isopropanol) and one substance is a solid (salicylic acid).
- Molecular weights range from 60 g/mol for isopropanol, 86 g/mol for hexane, 113 g/mol for chlorobenzene, to 138 g/mol for salicylic acid.
- All four substances are soluble in water.
- The peptide reactivity for chlorobenzene, hexane, and isopropanol is minimal; peptide reactivity information for salicylic acid is not available.
- Hexane and salicylic acid are considered irritants based on data in either mice or humans and isopropanol is considered negative based on data in rabbits; irritancy data for chlorobenzene are not available but irritancy potential is assumed as low based on clinical literature (**Table C-VII-1**).
- Among the five traditional LLNA sensitizers, five chemical classes are represented; one substance is a carboxylic acid (methyl methacrylate), two substances are metals (nickel [II] sulfate hexahydrate and cobalt chloride), one substance is both an amine and a phenol (3-aminophenol), and one substance is a heterocyclic compound (2-mercaptobenzothiazole).

Other characteristics of the substances that are classified as sensitizers by the traditional LLNA include:

- Four substances are solids (3-aminophenol, cobalt chloride, 2-mercaptobenzothiazole, and nickel [II] sulfate hexahydrate) and one substance is a liquid (methyl methacrylate).
- Molecular weights range from 100 g/mol for methyl methacrylate, 109 g/mol for 3-aminophenol, 130 g/mol for cobalt chloride, 155 g/mol for nickel (II) sulfate hexahydrate to 167 g/mol for 2-mercaptobenzothiazole.
- 2-Mercaptobenzothiazole is insoluble in water; the other four substances are soluble in water.
- The peptide reactivity for 2-mercaptobenzothiazole is high and that for 3-aminophenol is minimal; peptide reactivity data for the three other substances are not available.
- The EC3 values for the five substances identified as sensitizers by the traditional LLNA are: 0.6% for cobalt chloride, 1.7% for 2-mercaptobenzothiazole, 3.2% for 3-aminophenol, 4.8% for nickel [II] sulfate hexahydrate, and 90% for methyl methacrylate.
- All five substances are considered nonirritants based on GP data (**Table C-VII-1**).

Table C-VII-1 Indeterminate Results for the LLNA: DA When Multiple Stimulation Index Decision Criteria are Used[1]

Substance Name[2]	Vehicle[3]	LLNA: DA[4]	Traditional LLNA[4]	Skin Irritant?
Chlorobenzene	AOO	2.44, 25% (1/1 tests)	- (1.7, 25%)[5]	No data. Low irritancy potential assumed based on clinical literature.
Hexane	AOO	2.31, 100% (1/1 tests)	- (2.2, 100%)	Irritant at 100% (humans)
Isopropanol	AOO	1.97, 50%[5] (1/11 tests)	- (1.7, 50%)[5]	Negative at 100% (rabbits)
Salicylic acid	AOO	2.00, 25% (1/1 tests)	- (2.4, 25%)	Irritant at 20% aq. (mice)
3-Aminophenol (3.2%)	AOO	2.38, 10%[6] (1/3 tests)	+ (5.7, 10%)	Nonirritant at 5% (GP)
Cobalt chloride (0.6%)	DMSO	2.0, 5% (1/8 tests)	+ (7.2, 5%)	Negative at ≤ 0.5% (GP)
2-Mercaptobenzothiazole (1.7%)	DMF	2.01, 50%[5] (1/1 tests)	+ (8.6, 10%)	Nonirritant at 10% (GP)
Methyl methacrylate (90%)	AOO	1.81, 100% (1/1 tests)	+ (3.6, 100%)	Nonirritant at 3 M (GP)
Nickel (II) sulfate hexahydrate (4.8%)	DMSO	2.13, 10% 2.17, 5%[7] (2/8 tests)	+ (3.1, 5%)	Nonirritant at 0.15% (GP); irritant at 10% (humans)

Abbreviations: AOO = acetone: olive oil (4:1); aq. = aqueous; DMF = *N,N*-dimethylformamide; DMSO = dimethyl sulfoxide; GP = guinea pig; LLNA = murine local lymph node assay; LLNA: DA = murine local lymph node assay modified by Daicel Chemical Industries, Ltd., based on ATP content.

"+" = Sensitizer.

"-" = Nonsensitizer.

[1] Data source(s) indicated in Annex III of the BRD.

[2] Numbers in parentheses are EC3 values (concentrations needed to produce a stimulation index [SI] of three) for substances that are sensitizers in the traditional LLNA (see Table C-2 of the BRD).

[3] Vehicle listed is that used in both the LLNA: DA and the traditional LLNA, unless otherwise noted.

[4] Numbers indicated are highest SI and maximum concentration tested; highest SI is at maximum concentration tested, unless otherwise noted.

[5] Highest SI occurred at concentration 10%.

[6] Highest SI occurred at concentration 3%.

[7] Highest SI occurred at concentration 2.5%.

3.0 Test Method Reliability

An assessment of test method reliability (intralaboratory repeatability and intra- and interlaboratory reproducibility) is an essential element of any evaluation of the performance of an alternative test method (ICCVAM 2003). Repeatability refers to the closeness of agreement between test results obtained within a single laboratory when the procedure is performed on the same substance under

identical conditions within a given time period (ICCVAM 1997, 2003). Intralaboratory reproducibility refers to the extent to which qualified personnel within the same laboratory can replicate results using a specific test protocol at different times. Interlaboratory reproducibility refers to the extent to which different laboratories can replicate results using the same protocol and test substances, and indicates the extent to which a test method can be transferred successfully among laboratories. With regard to the LLNA: DA test method, there are no known intralaboratory repeatability studies, which was also the situation with the traditional LLNA.

The LLNA: DA data were amenable to both intralaboratory and interlaboratory reproducibility analyses. The evaluation of multiple SI decision criteria in **Section 2.0** of this Annex evaluated SI ≥ 2.5 as the decision criterion for classifying substances as sensitizers when used with a decision criterion of SI ≤ 1.8 to identify nonsensitizers. Thus, this section provides an assessment of reproducibility for the decision criterion of SI ≥ 2.5 to identify sensitizers.

3.1 Intralaboratory Reproducibility

Idehara et al. (2008) evaluated intralaboratory reproducibility of EC3 values for the LLNA: DA using two substances (isoeugenol and eugenol) that were each tested in three different experiments (**Table C-VII-2**). The data indicate coefficients of variation (CVs) of 21% and 11% for isoeugenol and eugenol, respectively. The authors state that for both compounds the EC3 values appeared to be close and that for each test substance the SI values for the same concentration were fairly reproducible (Idehara et al. 2008). The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) also determined the intralaboratory reproducibility of EC2.5 values (estimated concentrations needed to produce an SI of 2.5) for the same set of data. The results for EC2.5 values indicate slightly larger intralaboratory variability compared to EC3 values with CVs of 33% and 13% for isoeugenol and eugenol, respectively.

Table C-VII-2 Intralaboratory Reproducibility of EC3 and EC2.5 Values Using the LLNA: DA[1]

Isoeugenol			
Concentration (%)	**Experiment 1[2]**	**Experiment 2[2]**	**Experiment 3[2]**
Vehicle (AOO)	1.00 ± 0.54	1.00 ± 0.54	1.00 ± 0.30
0.5	1.50 ± 0.54	-------	1.22 ± 0.13
1	2.28 ± 0.60	-------	2.77 ± 1.01
2.5	2.78 ± 0.17	3.11 ± 1.15	3.01 ± 0.98
5	3.39 ± 0.69	4.39 ± 1.25	-------
10	5.68 ± 1.19	6.77 ± 0.23	-------
EC3	3.40%	2.35%	2.46%
EC2.5	0.82%	1.37%	0.75%
Mean EC3: 2.74% ± 0.58% and 21% CV *Mean EC2.5: 1.46% ± 0.48% and 33% CV*			

continued

Table C-VII-2 Intralaboratory Reproducibility of EC3 and EC2.5 Values Using the LLNA: DA[1] (continued)

Eugenol			
Concentration (%)	**Experiment 1[2]**	**Experiment 2[2]**	**Experiment 3[2]**
Vehicle (AOO)	1.00 ± 0.17	1.00 ± 0.17	1.00 ± 0.09
5	2.92 ± 1.00	2.80 ± 1.08	3.24 ± 0.70
10	7.35 ± 2.62	4.47 ± 0.98	4.79 ± 0.94
25	10.92 ± 3.63	5.62 ± 3.20	7.07 ± 0.44
EC3	5.09%	5.59%	4.50%
EC2.5	4.33%	3.59%	2.87%
Mean EC3: 5.06% ± 0.55% and 11% CV *Mean EC2.5: 4.23% ± 0.57% and 13% CV*			

Abbreviations: AOO = acetone: olive oil (4:1); CV = coefficient of variation; EC2.5 = estimated concentration needed to produce a stimulation index of 2.5; EC3 = estimated concentration needed to produce a stimulation index of three; LLNA: DA = murine local lymph node assay modified by Daicel Chemical Industries, Ltd., based on ATP content.

[1] Based on results discussed in Idehara et al. 2008; the number per group was not specified.

[2] Mean stimulation index value ± standard deviation.

3.2 Interlaboratory Reproducibility

Furthermore, data were submitted to NICEATM (Annex IV of the BRD) from a two-phased interlaboratory validation study on the LLNA: DA test method (Omori et al. 2008). In the first phase of the interlaboratory validation study, a blinded test of 12 substances was conducted in 10 laboratories. Three substances (2,4-dinitrochlorobenzene, hexyl cinnamic aldehyde, and isopropanol) were tested in all 10 laboratories. The remaining nine substances were randomly assigned to subsets of three of the 10 laboratories **(Table C-VII-3)**. In each laboratory, each substance was tested one time at three different concentrations. The dose levels for each substance were predetermined (i.e., the participating laboratories did not determine their own dose levels for testing). Nine substances are sensitizers and three substances are nonsensitizers according to traditional LLNA results. Six substances are reference substances included in LLNA performance standards recommended by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM): cobalt chloride, 2,4-dinitrochlorobenzene, hexyl cinnamic aldehyde, isoeugenol, isopropanol, and methyl salicylate.

The second phase of the interlaboratory validation study was designed to evaluate the reliability of the LLNA: DA for testing metallic salts using dimethylsulfoxide (DMSO) as a vehicle since two metals dissolved in DMSO (cobalt chloride and nickel [II] sulfate hexahydrate) from the first phase of the interlaboratory validation study yielded inconsistent results. Five coded substances (two of the five substances were unique to the second phase of the interlaboratory validation study) were tested in seven laboratories **(Table C-VII-4)**. One substance (i.e. hexyl cinnamic aldehyde) was tested in all seven laboratories. The remaining four substances (cobalt chloride, nickel (II) sulfate hexahydrate, lactic acid, and potassium dichromate) were randomly assigned to subsets of four of the seven laboratories. Each laboratory tested the substance one time at three different dose levels. Again, the dose levels for each substance were predetermined. Of the two substances not previously tested in the first phase of the interlaboratory validation study (lactic acid and potassium dichromate), one is a nonsensitizer and the other is a sensitizer according to traditional LLNA results, respectively. In addition, lactic acid is an ICCVAM-recommended LLNA performance standards reference substance.

The LLNA: DA test results from the two-phased interlaboratory validation studies are amenable to interlaboratory reproducibility analyses for three endpoints: sensitizer (positive) or nonsensitizer (negative) classification, and EC2.5 values. Analyses of interlaboratory reproducibility were performed using a concordance analysis for the qualitative results (sensitizer vs. nonsensitizer) (**Section 3.2.1**) and a CV analysis for the quantitative results (EC2.5 values) (**Sections 3.2** and **3.3**).

Table C-VII-3 Substances and Allocation for the First Phase of the Interlaboratory Validation Study for the LLNA: DA

Substance Name[1]	Vehicle	Concentration Tested (%)			Laboratory									
					1	2	3	4	5	6	7	8	9	10
2,4-Dinitro-chlorobenzene (+)	AOO	0.03	0.10	0.30	X	X	X	X	X	X	X	X	X	X
Hexyl cinnamic aldehyde (+)	AOO	5	10	25	X	X	X	X	X	X	X	X	X	X
Isopropanol (-)	AOO	10	25	50	X	X	X	X	X	X	X	X	X	X
Abietic acid (+)	AOO	5	10	25		X				X	X			
3-Aminophenol (+)	AOO	1	3	10	X		X					X		
Dimethyl isophthalate (-)	AOO	5	10	25	X		X				X			
Isoeugenol (+)	AOO	1	3	10				X	X				X	
Methyl salicylate (-)	AOO	5	10	25			X				X			X
Formaldehyde (+)	ACE	0.5	1.5	5.0	X	X			X					
Glutaraldehyde (+)	ACE	0.05	0.15	0.50	X	X			X					
Cobalt chloride[2] (+)	DMSO	0.3	1.0	3.0				X		X		X		
Nickel (II) sulfate hexahydrate (+)	DMSO	1	3	10				X		X		X		

Abbreviations: ACE = acetone; AOO = acetone: olive oil (4:1); DMSO = dimethyl sulfoxide; LLNA: DA = murine local lymph node assay modified by Daicel Chemical Industries, Ltd., based on ATP content.

[1] (+) indicates sensitizers and (-) indicates nonsensitizers according to traditional murine local lymph node assay results.

[2] Different doses tested for cobalt chloride in the first phase (0.3%, 1%, and 3%) and in the second phase (1%, 3%, and 10%) of the interlaboratory validation study.

Table C-VII-4 Substances and Allocation for the Second Phase of the Interlaboratory Validation Study for the LLNA: DA

Substance Name[1]	Vehicle	Concentration Tested (%)			Laboratory						
					11	12	13	14	15	16	17
Hexyl cinnamic aldehyde (+)	AOO	5	10	25	X	X	X	X	X	X	X
Cobalt chloride[2] (+)	DMSO	1	3	5	X		X	X			X
Lactic acid (-)	DMSO	5	10	25	X		X		X	X	
Nickel (II) sulfate hexahydrate (+)	DMSO	1	3	10	X	X		X		X	
Potassium dichromate (+)	DMSO	0.1	0.3	1.0	X	X			X		X

Abbreviations: AOO = acetone: olive oil (4:1); DMSO = dimethyl sulfoxide; LLNA: DA = murine local lymph node assay modified by Daicel Chemical Industries, Ltd., based on ATP content.

[1] (+) indicates sensitizers and (-) indicates nonsensitizers according to traditional murine local lymph node assay results.

[2] Different doses tested for cobalt chloride in the first phase (0.3%, 1%, and 3%) and in the second phase (1%, 3%, and 10%) of the interlaboratory validation study.

3.2.1 Interlaboratory Reproducibility – Qualitative Results

The qualitative (positive/negative) interlaboratory concordance analysis for the 12 substances that were tested during the first phase of the LLNA: DA interlaboratory validation study is shown in **Table C-VII-5** for SI ≥ 2.5. In a qualitative comparison of LLNA: DA calls (i.e., sensitizer/nonsensitizer), ten substances tested in either three or 10 laboratories had consistent results leading to 100% (3/3 or 10/10) interlaboratory concordance for those substances. There were two discordant substances (3-aminophenol and nickel [II] sulfate hexahydrate) for which interlaboratory concordance was 67% (2/3). One of the three laboratories that tested 3-aminophenol reported SI ≥ 2.5 at the highest dose tested (SI = 2.83 at 10%) and two laboratories did not achieve SI ≥ 2.5 at any dose tested (Annex IV of the BRD). One of the three laboratories that tested nickel (II) sulfate hexahydrate reported a maximum SI = 1.52, while the other two laboratories produced an SI ≥ 2.5 at all three doses tested (Annex IV of the BRD). Notably, when analyzing the dose response curves for the three tests performed for nickel (II) sulfate in the first phase of the two-phased interlaboratory validation study, only one study demonstrated a sufficient dose response (i.e., a parallel increase in SI relative to increase in concentration). Since the evaluation of interlaboratory reproducibility for the traditional LLNA did not include an evaluation of qualitative results (ICCVAM 1999), there were no traditional LLNA concordance data for comparison with the LLNA: DA concordance data from the first phase of the interlaboratory validation study.