4.0 RECOMMENDED BASAL CYTOTOXICITY TESTS: BALB/C 3T3 AND NORMAL HUMAN KERATINOCYTE (NHK) NEUTRAL RED UPTAKE (NRU) TESTS

4.1 Validation Status of the 3T3 NRU Test

The BALB/c 3T3 NRU test is probably the cytotoxicity test that has been used most frequently in formal validation programs, all of which were aimed at evaluation of cytotoxicity in predicting eye irritancy. Large-scale studies to be mentioned here are Phases I, II, and III of the Cosmetic, Toiletry, and Fragrance Association (CTFA) validation program (Gettings et al., 1991, 1992, 1994a, 1994b); the German eye irritation validation study (Spielmann et al., 1991, 1993, 1996); the European Commission/British Home Office (EC/HO) eye irritation validation study (Balls et al., 1995); and the European Cosmetic Toiletry and Perfumery Association (COLIPA) eye irritation study (Brantom et al., 1997). The 3T3 NRU Phototoxicity Test is a modification of the BALB/c 3T3 NRU test and involves a shorter chemical exposure and the additional application of light. The 3T3 NRU Phototoxicity Test has been fully validated (Spielmann et al., 1998a,b) and has gained regulatory acceptance.

For the purpose of evaluating the NRU test, and specifically the BALB/c 3T3 NRU test, as a standard test for basal cytotoxicity, all results available from these studies regarding the reliability (reproducibility within and between laboratories and over time) should be used to avoid wasting resources in repeating the establishment of reliability. Section 4.2 contains an example of establishing reliability of the BALB/c 3T3 NRU test from one of these studies.

4.2 Reliability of the 3T3 NRU Test

To establish interlaboratory reproducibility in the first phase of the German eye irritation validation study (Spielmann et al., 1991), 32 chemicals were tested in 12 laboratories using two tests: the hen's egg test-chorioallantoic membrane (HET-CAM) and the BALB/c 3T3 NRU test. (NRU tests using 3T3 cells were done in accord with the SOP presented in Appendix C.) Five independent repeat tests were conducted per laboratory. Of these 32 chemicals, three compounds [n-hexane, aluminum hydroxide, and di-(2-ethyl-hexyl)phthalate] showed unacceptably high inter-laboratory variability. For the other 29 chemicals, interlaboratory variability was acceptable (Table 2). Interlaboratory reproducibility was assessed with a standard procedure recommended by ISO 5725 (a program for analysis and reporting of proficiency tests and method evaluation studies). ISO 5725 describes reproducibility as an estimate of the limit below which the absolute value of the difference between two results determined in two different laboratories can be expected to fall, with a probability of 95%. The value tabulated in the far right column in Table 2 represents the span of about four standard deviations.

Table 2. Interlaboratory reproducibility of the 3T3 NRU cytotoxicity test determined according to ISO 5725 in 12 laboratories for 29 chemicals[a]

Substance	CAS No.	NR_{50} [b] (mg/ml)	Interlaboratory reproducibility[c] (mg/ml)
Dimethylsulphoxide	67-68-5	44.06	18.36
Propylene glycol	57-55-6	36.27	25.40
Acetone	67-64-1	18.41	14.74
Ethanol	64-17-5	18.01	14.69
Acetonitrile	75-056-8	13.72	15.38
Sodium chloride	7647-14-5	7.74	3.66
Thiourea	62-56-6	6.41	5.49
2-Butoxyethanol	111-76-2	5.43	8.73
Nicotinamide	98-92-0	5.36	5.78
Glutamic acid	56-86-0	4.84	2.01
Lactic acid	598-82-3	4.16	1.56
Pyridine	110-86-1	3.71	4.78
Benzoic acid	65-85-0	3.09	1.67
Isobenzoic furano dione	85-44-9	2.47	0.63
Cyclohexanol	108-93-0	1.89	2.07
Toluene	108-88-3	1.72	3.96
Salicylic acid	69-72-7	1.63	2.04
Tin(II) chloride	7772-99-9	1.55	2.35
Nitrobenzene	98-95-3	1.39	1.33
Tetrachlorethene	127-18-4	1.08	2.35
Aniline	62-53-3	1.07	1.25
EDTA-Na salt	13235-36-4	0.95	0.50
Ascorbic acid	50-81-7	0.49	0.81
Phenol	108-95-2	0.35	0.74
Acrylamide	79-06-1	0.29	0.19
Copper (II) sulfate	7758-98-7	0.10	0.05
Sodium lauryl sulfate	151-21-3	0.093	0.09
2-Propane-1-ol	107-18-6	0.05	0.06
Benzalkonium chloride	8001-54-5	0.01	0.01

[a]From Spielmann et al., 1991.

[b]NR_{50}= mean concentration of test substance reducing the viability of cells to 50% of the viability of controls.

[c]ISO 5725 describes reproducibility as an estimate of the limit below which the absolute value of the difference between two results determined in two different laboratories can be expected to fall, with a probability of 95%.

The second phase of the German eye irritation validation study was a blind trial for database development and involved the testing of 150 chemicals (Spielmann et al., 1993, 1996). Each chemical was assigned at random to two of the 12 total laboratories, since reproducibility of the BALB/c 3T3 NRU test was not an issue at this stage of the study. The final publication (Spielmann et al., 1997) on this phase focused on predictivity and test strategies for identification of severe eye irritants. The data from this publication have been re-analyzed for the present guidance document in the following way: since each chemical was tested in two different laboratories, the IC_{50} values obtained in two laboratories were plotted against each other, as shown in Figure 5 for 147 of the 150 chemicals. (Three chemicals had to be excluded because they were not tested according to the SOP.) Note that "Lab 1" represents the total of all participating laboratories, as does "Lab 2". Thus, Figure 5 does not show the comparability of results between two given laboratories. Rather, it shows the comparability of data obtained under routine conditions between randomly selected laboratories performing the BALB/c 3T3 NRU test according to the same SOP.

Results of the correlation analysis shown in Figure 5 are quite promising, since the linear correlation line (black) deviates only slightly from the ideal line (gray line at 45° angle). The linear correlation coefficient of $r = 0.88$ ($R^2 = 0.775$) shows a sufficient comparability of the data. Outliers, where data of the two laboratories differed by more than 1 log, occurred for less than 10% of the chemicals. A predominant reason for these interlaboratory deviations, discussed in Spielmann et al. (1997), was that one laboratory had used an adequate solvent for a test chemical, while the other laboratory had tested the chemical in media at concentrations above the aqueous solubility of the chemical. Thus, concentrations reported by the second laboratory were nominal rather than actual. As a consequence of this experience, later validation studies (Spielmann et al., 1998a,b) emphasized guidance for the use of solvents.

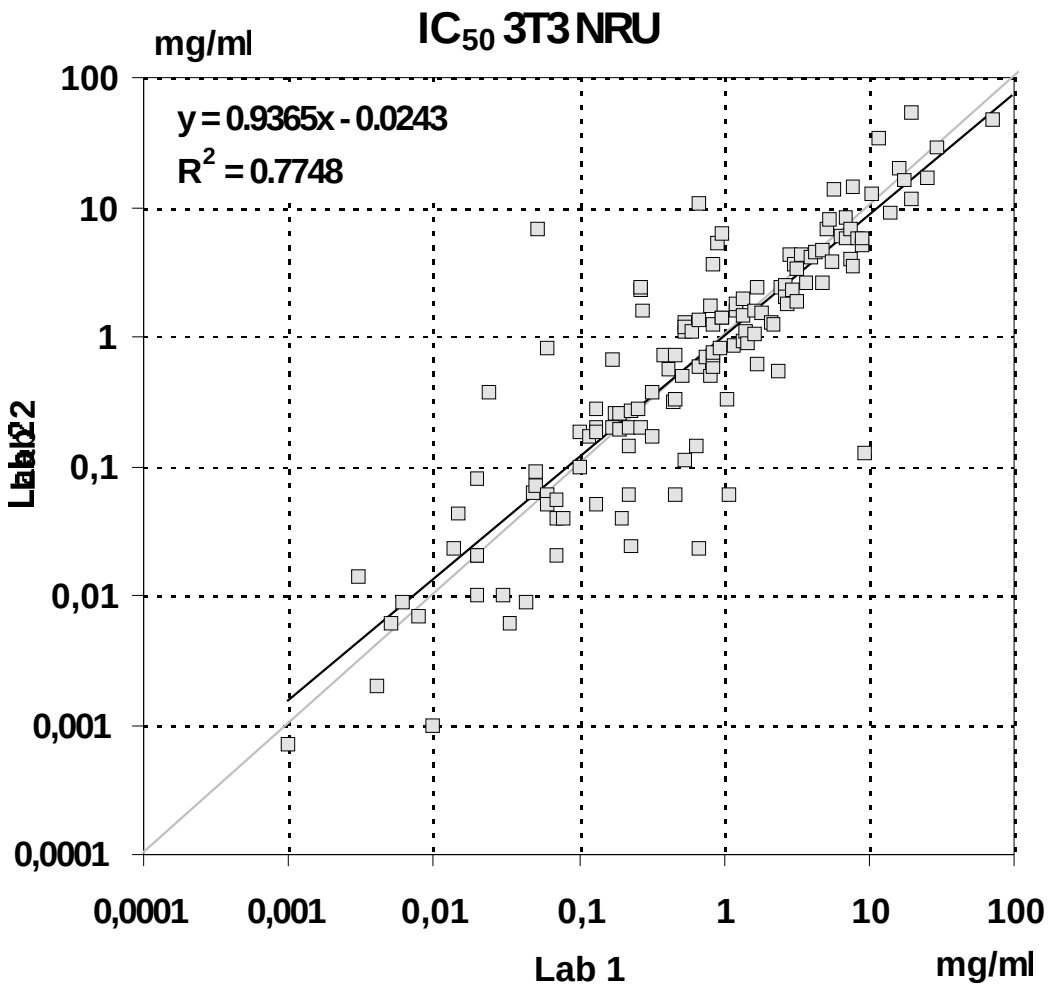

Figure 5. Interlaboratory comparability of the 3T3 NRU cytotoxicity test for 147 test chemicals in 2 different laboratories per chemical.
(Note: see text for explanation of the term "two laboratories per chemical".)

4.3 Validation Status of the NHK NRU Test

Although the NHK NRU test has been used less frequently in validation studies than has the BALB/c 3T3 NRU, the NHK NRU has been evaluated in several studies for its ability to predict eye irritation potential as reflected by Draize scores. It was used in Phases I, II, and III of the CTFA evaluation program (hydroalcoholic formulations, oil-and-water emulsions and surfactants and surfactant-containing formulations) (Gettings et al., 1991, 1994,1996); Phase III of the Soap and Detergent Manufacturers study using primarily neat surfactants and surfactant-containing formulations

(Bagley et al., 1994); as well as an independent study of surfactants and surfactant-containing formulations (Triglia et al., 1989). Many of these studies were subsequently reviewed by the Interagency Regulatory Alternatives Group, as part of a workshop review to evaluate the results of voluntary data submissions of *in vitro* methods to predict Draize scores (Harbell et al., 1997).

Gettings et al. (1996) evaluated the results of 34 different *in vitro* assays in testing 25 surfactant-based formulations for the prediction of Draize scores. The *in vitro* tests were ranked by discordance and separation index (i.e., the ability of the test to rank the toxicity of the 25 chemicals with the same relative rank as the Draize test). The NHK NRU test was not among the *in vitro* tests with the lowest discordance and highest separation index. Triglia et al. (1989), testing 12 surfactant-based formulations, suggested that sensitivity and specificity of the NHK NRU were sufficient for the test to be used as a screening tool as part of a battery of *in vitro* tests. Likewise, Harbell et al. (1997), in evaluating six data sets containing 9-45 surfactant or surfactant-containing materials, concluded that the NHK NRU had sufficient performance in predicting Draize scores that the assay could be used as a screen or adjunct over the range of toxicities found in personal care and household products.

4.4 Reliability of the NHK NRU Test

The reliability of the NHK NRU assay has been less well documented than that of the 3T3 NRU assay; however several reports have described the intralaboratory and interlaboratory variability of the test. Triglia et al. (1989) reported that 10 cytotoxicity trials in a single laboratory using the surfactant sodium lauryl sulfate (SLS) at five different concentrations produced coefficients of variation (CVs) <18% for all but the lowest concentration. (The average NRU_{50} [i.e., concentration reducing NRU to 50 % of control value] from one laboratory in these trials was 4.4 µg/ml; twelve years later the same laboratory has an average NRU_{50} for SLS of 4.4 +/- 0.97 µg/ml). Triglia et al. (1989) also reported interlaboratory variability for 12 compounds replicated in four laboratories. The interlaboratory CVs for the NRU_{50} means ranged from 19% - 60%. More recently, as part of the Interagency Regulatory Alternatives Group evaluation, Harbell et al. (1997) analyzed data from two laboratories that tested 22 materials in a blind fashion. NRU_{50} values for these materials showed an excellent interlaboratory correlation of 0.99. Dickson et al. (1993) also reported on variability for the NHK NRU assay and found that the NRU_{50} values for SLS tested in four different keratinocyte isolates were nearly identical at 66.7, 67.5, 70.9 and 73.4 µg/ml. Dickson et al. (1993) used a 24 h exposure rather than the 48 h exposure used for the other tests described in Sections 4.3 and 4.4.

5.0 CONCLUSION

This document provides guidance for using *in vitro* basal cytotoxicity assays to reduce the number of animals required for the conduct of *in vivo* lethality assays. The recommended approach takes advantage of the relationship between *in vitro* IC_{50}s and *in vivo* LD_{50}s derived from the RC for 347 chemicals (Halle and Spielmann, 1992; Halle, 1998). Detailed protocols for two recommended NRU assays, one using a rodent cell line, BALB/c 3T3 cells, and one using primary human cells, NHK, are included. Guidance is also provided for qualifying these tests, or any other *in vitro* cytotoxicity assay, for use with the RC regression to predict the starting dose for lethality assays.

ACKNOWLEDGEMENT

The authors wish to thank Dr. Elke Genschow with the Centre for the Documentation and Validation of Alternative Methods, Federal Institute for Health Protection of Consumers and Veterinary Medicine (Germany) for performing the correlation analysis for Figure 1 and the regression analyses for Figures 3 and 4.