2.0 TEST METHOD PROTOCOL COMPONENTS OF THE 3T3 AND NHK *IN VITRO* NRU TEST METHODS

The *Guidance Document* (ICCVAM 2001b) recommended that the following be incorporated into any *in vitro* cytotoxicity protocol used to predict rodent acute oral lethality:

- A cell line (or primary cells) that divides rapidly (e.g., with a doubling time of <24 hours)
- An initial seeding density that allows for exponential cell growth throughout the exposure period
- An exposure period that spans at least one cell cycle
- Appropriate positive control (PC) and vehicle control (VC) substances for which toxicity and lack of toxicity, respectively, has been well characterized by the performing laboratory
- Solvents that are used only at concentrations that do not cause significant toxicity to the cell system over the entire period of the assay
- A well-established, quantifiable cytotoxicity endpoint that has good interlaboratory reproducibility
- Tests that are compatible with at least 96-well plates and equipment (e.g., spectrophotometric microplate reader) that allow a quick and precise measurement of the endpoint of interest
- Use of a progression factor in the concentration-response experiment that yields graded effects between 0% and 100% cytotoxicity

Section 2.1 provides the basis for the selection of the *in vitro* 3T3 and NHK NRU test methods. **Section 2.2** provides descriptions of the NRU protocols applicable to this validation study. **Section 2.3** provides details for performing the 3T3 and NHK NRU test methods and explains the rationales for the various test method components, and **Section 2.4** describes any 3T3 and NHK NRU test method proprietary aspects. **Section 2.5** discusses the basis for the replicate and repeat tests conducted during validation of these two test methods. **Section 2.6** details the modifications and revisions made during the first two phases of the validation study which contributed to the development of the final protocol used in Phase III. **Section 2.7** describes the differences between the protocols used in this study and the protocols outlined in the *Guidance Document*. **Sections 2.8**, **2.9**, and **2.10** provide details on the solubility protocol evaluated during the validation study and used to identify the appropriate solvent for dissolving the reference substances.

The 3T3 and NHK NRU test method protocols were provided to the three laboratories that participated in the validation study (see **Section 5.6.3** for additional laboratory information). These were:

- The U.S. Army Edgewood Chemical Biological Center (ECBC)
- The FRAME Alternatives Laboratory (FAL)
- The Institute for *In Vitro* Sciences (IIVS)

A fourth laboratory (BioReliance Corporation, Rockville, MD) was used to procure and distribute the coded reference substances, and to perform solubility tests on the validation study reference substances prior to their distribution to the participating laboratories.

2.1 Basis for Selection of the *In Vitro* NRU Cytotoxicity Test Method

As stated in **Section 1**, in agreement with the recommendations of the Workshop 2000 participants (ICCVAM 2001a), ICCVAM made the following recommendations and forwarded them to U.S. Federal agencies along with the Workshop 2000 Report (ICCVAM 2001a) and *Guidance Document* (ICCVAM 2001b).

> "ICCVAM concurs with the Workshop recommendation that near-term validation studies should focus on two standard cytotoxicity assays: one using a human cell system and one using a rodent cell system. Since the murine BALB/c 3T3 cytotoxicity assay has been evaluated for only a limited number of chemical classes, there is merit in determining its usefulness with a broader array of chemical classes. Cell lines established from the rat rather than the mouse might also be considered, as most acute oral toxicity testing is conducted in this species. Human cell lines should also be considered since one of the aims of toxicity testing is to make predictions of potential toxicity in humans. Future validation studies should therefore compare rodent and human *in vitro* data with one another, with rodent *in vivo* data, and with human *in vivo* data. Correlations between *in vitro* and *in vivo* data might help in selecting cytotoxicity assays for further evaluation". (ICCVAM 2001a)

Based on this recommendation and the *Guidance Document* recommendation, NICEATM and ECVAM selected the 3T3 and NHK NRU basal cytotoxicity test methods for validation.

2.1.1 *Guidance Document* Rationale for Selection of *In Vitro* NRU Cytotoxicity Test Methods

The *Guidance Document* (ICCVAM 2001b) provided the basic approach for the use of *in vitro* NRU basal cytotoxicity test methods to predict starting doses for rodent acute oral toxicity assays using the RC millimole regression. The 3T3 and NHK NRU test method protocols used in the validation study were derived from those proposed in the *Guidance Document*.

2.1.2 *Guidance Document* Rationale for Selection of Cell Types

The Workshop 2000 participants (ICCVAM 2001a) concluded that there were no significant differences between the basal cytotoxicity results obtained using permanent mammalian cell lines, primary human cells (e.g., NHK cells), or the IC_{50x} approach of Halle and Spielmann (Halle 1998, 2003; Spielmann et al. 1999; Halle and Spielmann 1992). Further, the *Guidance Document* recommended that *in vitro* basal cytotoxicity test methods not use hepatocytes (or related metabolically competent cells) or other types of highly differentiated cells because they may not give the best prediction of acute lethality for the large variety of chemicals likely to be tested (ICCVAM 2001b). However, it was recognized that, ultimately, simple predictive systems (*in vitro* or *in silico*) would be needed for early identification of those substances likely to be metabolized to more toxic or less toxic species than the parent chemical as well as those that were likely to exhibit cell-specific toxicity (e.g., Fentem et al. 1993; Seibert et al. 1996; Curren et al. 1998; Ekwall et al. 1999).

Established rodent cell lines were recommended for validation because (ICCVAM 2001b):

- It was assumed that such cells would give the best prediction of rat and mouse acute oral lethality (i.e., like correlates with like).
- The use of a readily available, easy to culture, immortalized cell line for *in vitro* cytotoxicity testing would accelerate the development of a database that can be used to analyze the usefulness of this approach.

Human cells also offer potential advantages. As determined in the MEIC project, the *in vitro* test methods with the best predictivity for peak acute LC values in humans generally used human cell lines (Ekwall et al. 1998b). Thus, a long-term advantage of using human cells is that *in vitro* human cell cytotoxicity data can be added to human toxicity databases to facilitate the development of test methods that may better predict acute oral human lethality.

3T3, an immortalized mouse fibroblast cell line, and NHK, primary human cells, were selected as representative rodent and human cells, respectively, for the NICEATM/ECVAM validation study. Historical data for the 3T3 NRU test were available from a variety of studies, including controlled and blinded validation studies, indicating the reliability of this test method (Gettings et al. 1991, 1992, 1994a, 1994b; Spielmann et al. 1991, 1993, 1996; Balls et al. 1995; Brantom et al. 1997). NHK cells have also been used in validation studies for basal cytotoxicity test methods with good results (Willshaw et al. 1994; Sina et al. 1995; Gettings et al. 1996; Harbell et al. 1997).

2.2 Overview of the 3T3 and NHK NRU Test Methods

The *Guidance Document* (ICCVAM 2001b) includes a proposed 3T3 NRU test method protocol based on the 3T3 Cytotoxicity Test (INVITTOX Protocol No. 46; available from the FRAME-sponsored INVITTOX database [http://embryo.ib.amwaw.edu.pl/invittox/]), which in turn was based on the Borenfreund and Puerner (1985) protocol, as elaborated on in Spielmann et al. (1991, 1996). This protocol was updated based on experience obtained during the validation of the 3T3 NRU Phototoxicity Test (INVITTOX Protocol No. 78; also available at the FRAME INVITTOX database). The RC millimole regression for prediction of acute oral rat and mouse toxicity (Halle 1998, 2003; Spielmann et al. 1999) was included as the prediction model (ICCVAM 2001b; see **Section 1.1.2**).

The NHK NRU protocol provided in the *Guidance Document* was based on the protocol used by IIVS, which was based on a NRU protocol of Borenfreund and Puerner (1984) and a rat epidermal keratinocytes protocol (Heimann and Rice 1983). Formulations for the media and solutions, and general NHK cell culture techniques, correspond to Clonetics® products from the CAMBREX Corporation.

The protocol components for the 3T3 and NHK NRU test methods used in this validation study are similar (see **Figure 2-1**). The nature of the NRU response is described in **Section 1.3.1**. **Figure 2-1** provides an overview to the major steps for performance of the *in vitro* NRU test methods. The following procedures are common to both cell types:

- Preparation of substances and the PC
- Cell culture environmental conditions
- Determination of test substance solubility

- 96-well plate configuration for testing samples
- Range finder and definitive tests
- Microscopic evaluation of cell cultures for toxicity based on morphological alterations
- Procedures for measurement of NRU
- Data analysis procedures

The main protocol differences between the two cell lines are:

- The conditions of propagation of the cells in culture (e.g., time needed to reach appropriate confluence)
- The growth media components
- The volumes of substances applied to the 96-well plates
- The number of cell divisions undergone by each cell line during exposure to a test substance

2.2.1 The 3T3 NRU Test Method

2.2.1.1 *Initiating and Subculturing 3T3 Cells*

Each laboratory initially prepared a large pool of 3T3 cells (described further in **Section 2.3.1.1**), cryogenically preserved multiple ampules of these cells in liquid nitrogen, and periodically removed an ampule when needed. Although the NRU protocols used for each study phase provided cell culture density guidelines for subculturing the cells, each laboratory refined the final seeding density to achieve optimal growth.

Cryopreserved 3T3 cells were thawed, resuspended in a culture medium containing Dulbecco's Modification of Eagle's Medium (DMEM) supplemented with non-heat-inactivated 10% newborn calf serum (NCS), transferred into tissue culture flasks (25 or 75-80 cm^2), and incubated at 37 °C ±1 °C, 90% ±5% humidity, and 5.0% ±1% CO_2/air. When cells reached 50 to 80% confluence (as estimated from a visual inspection of cell density), they were removed from the flask by trypsinization. A single-cell suspension was added to new flasks for propagation and the cells were passaged/subcultured at least two times[1] before seeding into 96-well plates for testing. This study did not evaluate the potential effects that cell passage number may have on the performance of the 3T3 NRU test method.

[1] 3T3 cells were maintained in culture for approximately two months (approximately 18 passages) and used for the NRU test. The *Guidance Document* (ICCVAM 2001b) did not provide a rationale for using 18 passages as the limit, but it was probably recommended to maintain homogeneity of the 3T3 cell population (i.e., decrease the potential of the population to drift genetically). The more passages the cells undergo, the more likely their response to chemical stress may change.

Figure 2-1 Major Steps in the Performance of the NRU Test Methods

(1) Cells (3T3 or NHK) are seeded into 96-well plates to form a sub-confluent monolayer; plates are incubated at 37 ºC (24 hours for 3T3 cells; 48-72 hours for NHK cells)

⇓

(2) Culture medium is removed (3T3 cells only)

⇓

(3) Reference substances in the appropriate solvents are added to the cells; cells are exposed for 48 hours at 37 ºC over a range of eight (8) concentrations

⇓

(4) Cells are evaluated microscopically for toxicity based on morphological appearance

⇓

(5) Treatment medium is removed; cells are washed once with Dulbecco's Phosphate Buffered Saline (D-PBS); Neutral Red (NR) dye medium is added (3T3 cells: 25 µg/mL NR dye; NHK cells: 33 µg/mL NR dye); plates are incubated for 3 hours at 37 ºC

⇓

(6) NR medium is discarded; cells are washed once with D-PBS; NR desorbing fixative is added to the wells

⇓

(7) Plates are shaken for 20 minutes at room temperature

⇓

(8) NR absorption is measured at optical density (OD) 540 ±10 nm

⇓

(9) NRU is calculated as a percent of vehicle control values to define IC_{20}, IC_{50}, and IC_{80} concentrations (µg/mL)[2]

Abbreviations: NRU=Neutral red uptake; 3T3=BALB/c 3T3 fibroblasts; NHK=Normal human epidermal keratinocytes; IC_{20}, IC_{50}, IC_{80}=Substance concentration that reduces cell viability by 20, 50, and 80%, respectively.

[2] IC_{50} values are used for estimating the LD_{50} value of a reference substance. The IC_{20} and IC_{80} values were determined for possible use in estimating human lethal concentrations in blood.

2.2.1.2 *Preparation of Cells for 96-well Plate Assays*

After subculturing the cells, 100 µL of the cell suspension ($2.0 – 3.0x10^3$ cells/well) were placed in the appropriate wells and 100 µL of cell-free culture medium were dispensed into the 36 peripheral wells (blanks). The peripheral wells were in rows 1 and 8 and columns 1 and 12 (See **Figure 1** in **Appendix B1** or **B2**). Peripheral wells were used only for blanks because they may be subjected to more evaporation than interior wells. The *Guidance Document* authors (and the SMT and Study Directors) concluded that such conditions would ultimately affect cell growth in these wells. One plate was prepared for each reference substance. The cells were incubated for 24 ±2 hours at 37 ºC and checked visually to be sure that approximately a 50% confluent monolayer was present at the time of substance application.

2.2.1.3 *Reference Substance Application*

After the appropriate incubation period to achieve a half-confluent monolayer, the medium was removed and 50 µL of culture medium with 10% NCS were added to each well. Then, 50 µL treatment medium containing the appropriate substance concentrations were added for a final concentration of 5% NCS. The cells were then incubated at 37 ºC for 48 ±0.5 hours. At the end of the incubation period, the cells were microscopically evaluated for changes in morphology and their appearance was documented (as per Visual Observation Codes in the protocol) prior to measurement of NRU.

2.2.2 The NHK NRU Test Method

2.2.2.1 *Initiating and Subculturing NHK Cells*

Cryopreserved NHK cells (ampules of cryopreserved cells were obtained from CAMBREX Corporation and stored in liquid nitrogen until needed) were thawed, resuspended in serum-free keratinocyte complete growth medium (see **Section 2.3.1.4** for components of the medium), transferred into tissue culture flasks (25 cm^2 without fibronectin-collagen coating), and incubated at 37 °C ±1 °C, 90% ±5% humidity, and 5.0% ±1% CO_2/air. When the cells reached 50 to 80% confluence (as estimated from a visual inspection of cell density), they were removed from the flask by trypsinization and prepared for subculturing into the 96-well plates. Care was taken to prevent the keratinocyte cultures from becoming 100% confluent as this may lead to cell differentiation, which would alter the intrinsic sensitivity of these cells to cytotoxic substances. To minimize potential sources of experimental variability, the laboratories used the same lot of Clonetics® cells throughout the validation study, the same brand of growth medium and supplements (and concentrations of supplements), and cells were not used beyond their second passage. The protocols for each study phase provided cell culture density guidelines, but each laboratory refined the final seeding densities to achieve appropriate growth.

2.2.2.2 *Preparation of Cells for 96-well Plate Assays*

After subculturing the cells, 125 µL of the cell suspension ($2.0 – 2.5x10^3$ cells/well) were placed in the appropriate wells and 125 µL of cell-free culture medium were dispensed into the peripheral wells (blanks). One plate per reference substance was prepared. The cells were incubated at 37 ºC for 48-72 hours and checked to be sure that cultures were at 20 to 50% confluence at the start of exposure to the reference substance.

2.2.2.3 *Reference Substance Application*

To add the reference substances, 125 µL of culture medium containing the appropriate reference substance concentrations were added to the existing 125 µL of culture medium in the test wells. The cells were then incubated at 37 ºC for 48 ±0.5 hours. At the end of the exposure period, the cells were microscopically evaluated for changes in morphology and their appearance was documented (as per Visual Observation Codes in the protocol [see **Appendices B1** and **B2**]) prior to measurement of their NRU.

2.2.3 Measurement of NRU in the 3T3 and NHK Test Methods

The treatment medium was removed from the 96-well plates, the cells were rinsed with phosphate buffered saline (PBS), and 250 µL NR dye medium was added to the wells (25 µg NR/mL for 3T3 cells; 33 µg NR/mL for NHK cells). The plates were then incubated (37 °C ±1 °C, 90% ±5% humidity, and 5.0% ±1% CO_2/air) for three hours. After incubation, the NR medium was removed, the cells were rinsed with PBS, and 100 µL of the desorb solution were applied. The plates were shaken on a microtiter plate shaker for 20 to 45 minutes to extract NR from the cells and to form a homogeneous solution. The optical density (OD) of the resulting colored solution was measured (within 60 minutes of adding the desorb solution) at 540 nm ±10 nm (OD_{540}) in a spectrophotometric microtiter plate reader, using the blank wells as reference. Data from the plate reader were transferred to a Microsoft® EXCEL® (Microsoft Corporation, Redmond, WA, USA) spreadsheet template (hereafter know as EXCEL® template) designed by the SMT and the testing laboratories for statistical analyses.

2.3 Descriptions and Rationales of the 3T3 and NHK NRU Test Methods

The protocols used in Phases I, II, and III of the validation study (**Appendices B** and **C**) are modifications of the protocols reported in the *Guidance Document* (ICCVAM 2001b). The participating laboratories provided comments and recommendations during the development of these protocols. The following information is specific to the protocols used in this validation study.

2.3.1 Materials, Equipment, and Supplies

2.3.1.1 *3T3 Cells*

The CCL-163, 3T3 BALB/c mouse fibroblast, cell line, clone 31 from the American Type Culture Collection (ATCC), Manassas, VA, USA, was used. The 3T3 cells, an immortalized mouse fibroblast cell line, were procured from the ATCC by IIVS at passage 64. IIVS cultured the cells to expand their number and cryogenically preserved them as a pool at passage number 69. ECBC and FAL received frozen ampules of cells at passage number 69 from IIVS, propagated the cells, and cryopreserved multiple ampules of cells at a slightly higher passage number to establish their working cell banks for use throughout the study. Each laboratory determined the doubling time for the 3T3 cell line prior to NRU testing in Phase Ia as required by the protocol in **Appendix C1**. The following doubling times were reported: 18.6 hours by ECBC; 17 hours by FAL; and 17 hours by IIVS. No other doubling time measurements were made. The extent of cell confluence was monitored during the study to identify when the cultures were in exponential growth.

2.3.1.2 *NHK Cells*

A single lot of pooled donor, primary neonatal foreskin keratinocyte (NHK) cells (Clonetics® # CC-2507; lot # 1F0490N) from CAMBREX Bio Science Walkersville, Inc., Walkersville, MD, USA, was used throughout the validation study. Keratinocytes from other sources would be acceptable if they meet the growth requirements identified in the protocols. Each laboratory determined the doubling time for the NHK cells prior to testing in Phase Ia (as required by the protocol in **Appendix C2**). The following doubling times were reported: 21 hours by ECBC; 10 hours by FAL; and 15.8 hours by IIVS. No other doubling time measurements were made. The extent of cell confluence was monitored during the study to identify when the cultures were in exponential growth.

2.3.1.3 *Tissue Culture Materials and Supplies*

The 3T3 and NHK NRU test methods require general tissue culture materials and supplies (see **Appendices B1** and **B2** [protocols] for formulations, and concentrations of solutions and media). Both test methods used the same materials for solubility testing (**Section 2.8.1**). Freshney (2000) provides information on all aspects of cell culture, including materials, supplies, and equipment needed. The following materials were needed for both test methods:

- Trypsin (0.05%)
- PBS
- Hanks' Balanced Salt Solution (HBSS) without Ca^{2+} and Mg^{2+}
- NR dye
- Glacial acetic acid
- Dimethyl sulfoxide (DMSO) [analytical grade]
- Ethanol (ETOH) [100% non-denatured for test substance preparation]
- Distilled water

2.3.1.4 *Cell Culture Materials*

Laboratory items needed include the following:

- Sterile, disposable tissue culture plasticware (e.g., 25 cm^2,75-80 cm^2 flasks; multiwell/microtiter [96-well] plates; petri dishes) [Note: The laboratories in this study used tissue culture plasticware from various suppliers.]
- Cryogenic ampules
- Pipettes, pipettors, pipette tips
- Multichannel solution reservoirs
- Centrifuge tubes
- Microporous sterilization filters
- General plastic containers
- Glass tubes (for preparation of substance dilutions)

2.3.1.5 *Equipment*

Performance of the NRU tests requires a laboratory equipped with a designated cell culture area. Essential equipment for cell culture work and the NRU test methods include:

- Incubator (37 ºC ±1 ºC, 90% ±5% humidity, 5.0% ±1% CO_2/air)
- Laminar flow clean bench/cabinet (standard: "biological hazard")
- Water bath (37 ºC ±1 ºC)
- Inverted phase contrast microscope (with 10X to 40X objectives)

- Centrifuge (capable of 220 x g)
- Laboratory balance (capable of measuring to 10 mg)
- Spectrophotometer for reading 96-well plates (i.e., microtiter plate reader) equipped with 540 nm ±10 nm filter
- Shaker for microtiter plates
- Cell counter or hemocytometer
- Pipetting aid (e.g., vacuum pipettor unit)
- Pipettes, pipettors (multi-channel and single channel, multichannel repeater pipette)
- Waterbath sonicator
- Refrigerator
- Freezer (to at least -70 ºC)
- Cryostorage container (and liquid nitrogen supply)
- Magnetic stirrer
- Antistatic bar ionizer
- Personal computer
- Osmometer
- pH meter

2.3.1.6 *Culture Medium*

For 3T3 Cells

DMEM containing high glucose (4.5 gm/L) and supplemented with NCS, L-glutamine, penicillin, and streptomycin was used for the 3T3 cells. Heat-inactivated serum was not used in this study. Heat-inactivation of serum is often used to destroy heat-labile components such as complement factors, and microbial contaminants such as mycoplasma (Hyclone® 1996; Mediatech, Inc. 2006). However, some heat-labile complement factors can also be inactivated by the standard cell culture practice of warming serum-containing medium to 37 °C prior to use, and mycoplasma can be eliminated by filtering the medium (e.g., using 0.1 µm pore-size rated filters). Heating serum to 56 °C (heat-inactivation temperature) can destroy other heat-labile components such as growth factors, vitamins, amino acids, and hormones. Loss of these components can diminish the capacity of the serum to promote attachment of cells to culture vessel surfaces and to support cell growth. An additional confounding factor is that the procedure for heat-inactivation is highly precise, and deviation from the basic protocol can create additional issues such as protein denaturation and serum turbidity.

For NHK Cells

Although the contents of the NHK basal culture medium are proprietary, the formulation is based on a commercially available, non-proprietary basal medium (MCDB 153 medium formulation [Tsao et al. 1982]; e.g., MCDB 153 medium - SIGMA-ALDRICH product number #M 7403 http://www.sigmaaldrich.com/sigma/datasheet/m7403dat.pdf). The laboratories recommended this medium for use with the CAMBREX Clonetics® NHK cells because they all had access to this supplier. Other media are acceptable for NHK NRU testing if the performance standards prescribed in the media prequalification protocol are met (see **Appendix B4** and **Section 2.6.3.5**).

The serum-free culture medium used for NHK cells was Clonetics® keratinocyte basal medium (KBM®) supplemented with KBM® SingleQuots® (epidermal growth factor, insulin, hydrocortisone, antimicrobial agents, bovine pituitary extract [BPE]) and Calcium SingleQuots® (calcium) [all from CAMBREX Corporation] to make keratinocyte complete growth medium. Although the keratinocyte complete growth medium is a defined serum-free medium, it contains BPE collected from bovine pituitary glands. BPE contains growth factors and hormones, and is added to serum-free medium as a mitogenic supplement. Variability in the composition of the BPE could be a factor in cell growth kinetics. However, it is suggested that the undefined BPE components could be replaced with defined growth supplements, such as insulin, epidermal growth factor, and fibroblast growth factor, without adversely affecting the cellular proliferation rates and general physiology of human keratinocytes (Life Technologies, Inc. 1997).

2.3.2 Reference Substance Concentrations/Dose Selection

Each laboratory weighed and dissolved the reference substances on the same day as the start of the exposure period. The highest concentration of dissolved reference substance was identified using the solubility protocol and designated as the 2X stock solution. All reference substance dilutions for an assay were serially derived from this stock solution (see *Guidance Document* for serial dilution methods).

2.3.2.1 *Range Finder Test*

A range finder 3T3 or NHK NRU test was performed to determine the concentrations of a reference substance to be used for the definitive (concentration-response) test (see **Section 2.3.2.2**). The range finder test used eight concentrations of the reference substance prepared by diluting the stock solution using log intervals to cover a large concentration range (e.g., 1:10, 1:100, 1:1000, etc.; up to eight orders of magnitude). The highest concentrations applied to the cells were 10 mg/mL for substances dissolved in culture medium and 1 mg/mL in medium for substances dissolved in DMSO, unless precluded by solubility. ETOH was not used as a solvent for any of the substances in the validation study (see **Sections 2.8, 2.9,** and **2.10**).

If the range finder test did not produce cytotoxicity, then a second range finder test was conducted at higher concentrations (e.g., the highest concentration would be >10 mg/mL if in medium, >1 mg/mL if in DMSO) unless precluded by solubility. If the substance being tested was insoluble or poorly soluble, then more stringent solubility procedures were employed to increase the stock concentration (to the maximum concentration specified in **Appendices B1** and **B2**). If the range finder test produced a biphasic dose-response curve[3] for NR uptake, the concentrations selected for the definitive tests covered the response range that included the lowest concentration that reduced viability by 50% (see **Section 2.6.3.2**).

2.3.2.2 *Definitive Test*

The concentration-response determination is referred to as the definitive test because it is used to determine the IC_{50} value of the substance being tested. The concentration closest to the calculated IC_{50} value in the range finder test served as the midpoint of the eight concentrations tested in the definitive test. In the absence of other information (e.g.,

[3] A biphasic dose-response curve is a dose-response in which cytotoxicity increases (as dose increases), plateaus, and then increases again.

knowledge of the slope of the toxic response), the recommended dilution factor was 1.47 ($\sqrt[6]{10}$), which divides a log interval into six equidistant steps (e.g., 10, 14.7, 21.5, 31.6, 46.4, 68.1, 100). The *Guidance Document* considered a progression factor of 1.21 ($\sqrt[12]{10}$) to be the smallest factor practically achievable, and this was the lowest required concentration interval. The PC was tested similarly to the reference substances in the definitive test and the same recommended dilution factors were used (dilution factor at the discretion of the Study Director).

A definitive test was considered successful if it met all of the test acceptance criteria outlined in the NRU protocols. Definitive tests were repeated as per the protocols if the test failed to meet all of the test acceptance criteria. **Section 2.5** addresses the basis for replicate testing.

If minimal or no cytotoxicity was observed in the range finder test, the maximum concentration for the definitive test was determined as follows:

- For Substances Prepared in NHK or 3T3 Medium: A review of the RC chemicals used in this validation study showed that, among water-soluble chemicals, glycerol had the highest reported IC_{50} value (57 mg/mL). To capture this value, and that of other relatively non-toxic chemicals, the highest concentration of a substance applied to the cells in the definitive test was either 100 mg/mL (using 200 mg/mL 2X stock) or the maximum soluble dose if the substance was not soluble at that concentration.
- For Substances Prepared in DMSO: Based on the maximum concentration of DMSO that could be added to culture medium without causing cytotoxicity (i.e., 0.5%), the highest concentration of a substance that could be applied to the cells in the definitive test was 2.5 mg/mL. In the event that the reference substance was not soluble at this concentration, the highest soluble concentration was used.

2.3.3 NRU Endpoints Measured

2.3.3.1 *NRU and Measurement*

After cells were exposed to the reference substance or the controls (PC; VC) for 48 hours, they were washed and incubated with the NR dye at 37 ºC for an additional three hours. The dye was eluted from the cells using a desorb solution and the OD of the resulting solutions were measured using a spectrophotometric microtiter plate reader. Because NR is absorbed by healthy cells, the amount of dye eluted, as measured by the spectrophotometer, is proportional to NRU and thus to the number of live cells present at culture termination. The OD data from the spectrophotometer were recorded on the EXCEL® template. Relative cell viability for each reference substance and the PC was determined using six replicate wells (six wells [minimum of four scorable] in the 96-well plate) per concentration. Cells treated with the VC were considered to have 100% cell viability (i.e., the mean OD of the VC wells = 100% viability). Cell viability in other test wells was computed in reference to the mean VC OD value (i.e., [well OD/mean VC OD] x 100 = % viability).

2.3.3.2 *Determination of IC_{50}, IC_{20}, and IC_{80} Values*

IC_{50} values were determined from the concentration-response curve using a Hill function, which is a four parameter logistic mathematical model relating the concentration of a

substance to the response (typically following a sigmoidal shape). Modifications to the Hill function used in later phases of the study are described in **Section 2.6.3**.

Data from the EXCEL® template were transferred to a template designed by the SMT for GraphPad PRISM® 3.0, a commercially available statistical software (GraphPad Software, Inc., San Diego, CA, USA – hereafter known as PRISM® template). The PRISM® template used the Hill function to calculate the IC_{50}, IC_{20}, and IC_{80} concentrations, reported as μg/mL of reference substance in solution. IC_{20} and IC_{80} data were collected for potential use in designing a prediction model for estimating human lethal blood concentrations.

2.3.4 Duration of Reference Substance Exposure

The SMT and laboratory representatives reevaluated the reference substance exposure duration recommended in the *Guidance Document* (ICCVAM 2001b) before initiating the study. The *Guidance Document* recommended an exposure of 24 hours for the 3T3 cells and 48 hours for the NHK cells. However, Riddell et al. (1986) showed large differences in cytotoxicity for 3T3 cells in response to some chemicals, depending on whether the exposure duration was 24 or 72 hours. Although the toxicity induced by substances that damage, for example, cell membranes is likely to be observed in a relatively shorter time, the toxic effects of substances that interfere with cell functions/processes specifically relating to DNA replication (e.g., protein and nucleic acid synthesis) and cell division (e.g., mitotic spindle formation) are more pronounced after longer exposure periods. This occurs because cells are affected only at certain phases of the cell cycle.

IIVS conducted studies to evaluate the effect of exposure durations of 24, 48, and 72 hours and of 48 and 72 hours on the sensitivity of 3T3 cells and NHK, respectively, to six chemicals selected from the list in Riddell (1986). Because the closest fit to the RC millimole regression occurred when a 48-hour exposure duration was used, this exposure duration was selected for use with both cell types in the validation study (Curren et al. 2003) (see **Appendix E**).

2.3.5 Known Limits of Use

2.3.5.1 *Solubility/Precipitation/Volatility*

In vitro test methods cannot be used for substances that cannot be dissolved in media, DMSO, or ETOH at a sufficiently high concentration to induce cytotoxicity in excess of 50%. Also, chemicals that are unstable or exothermic in water cannot be adequately tested with these *in vitro* test methods (as well as *in vivo* methods).

Precipitation of a test substance in the dosing solution or in the culture medium after the substance to be tested has been added can affect the concentration-response and thus reduce the accuracy of the calculated IC_{50}. Some reference substances used in the validation study had precipitates in their medium/DMSO 2X concentrations prior to dilution for application to the test wells. Precipitates were also observed for some substances in a number of test wells after addition of the media/DMSO 1X solutions (see **Section 5.8** and **Table 5-11**) to the cultures and/or at the end of the exposure period.

Volatility was detected for a number of reference substances during the range finder tests by observance of cross contamination (i.e., high cytotoxicity) in VC wells. Plate sealers were

used during the definitive tests to control volatility (see **Section 2.6.3** – *Testing Volatile Reference Substances*), and could be used during the range finder tests if the Study Director suspected that the reference substance might be volatile. The use of plate sealers required additional laboratory training, and some volatile substances were difficult to test even with the use of plate sealers. Furthermore, some test substances (e.g., organic solvents) may react chemically with the plastic in the sealers.

2.3.5.2 *Biokinetic Determinations*

The Workshop 2000 report (ICCVAM 2001a) discussed the role of chemical biokinetics *in vivo* vis-a-vis acute toxicity, as illustrated in the following quote:

> "Results obtained from *in vitro* studies in general are often not directly applicable to the *in vivo* situation. One of the most obvious differences between the situation *in vitro* and *in vivo* is the absence of processes regarding absorption, distribution, metabolism and excretion (i.e., biokinetics) that govern the exposure of the target tissue in the intact organism. The concentrations to which *in vitro* systems are exposed may not correspond to the actual situation at the target tissue after *in vivo* exposure. In addition, the occurrence of metabolic activation and/or saturation of specific metabolic pathways or absorption and elimination mechanisms may also become relevant for the toxicity of a compound *in vivo*. This may lead to misinterpretation of *in vitro* data if such information is not taken into account. Therefore, predictive studies on biological activity of compounds require the integration of data on the mechanisms of action with data on biokinetic behavior."

The 3T3 and NHK NRU test methods do not account for biokinetics.

2.3.5.3 *Organ-Specific Toxicity*

The Workshop 2000 report also addressed concerns about the *in vitro* prediction of organ-specific toxicity, and identified the organ systems for which failure after acute exposure could lead to lethality (i.e., liver, central nervous system, kidney, heart, lung, and hematopoietic system) (ICCVAM 2001a). Each organ system was reviewed individually. Although the 3T3 and NHK NRU test methods do not assess organ-specific toxicity, they may be useful in a test method battery such as that proposed by the Workshop 2000 participants (see **Section 2.3.5.4**).

2.3.5.4 *The Role of Cytotoxicity Tests in an In Vitro Battery Approach for Possible Replacement of In Vivo Acute Toxicity Testing*

A five-step *in vitro* testing scheme was proposed for a test battery that may eventually be demonstrated to be an adequate replacement for rodent acute oral toxicity test methods for regulatory purposes (ICCVAM 2001a).

Step 1: Perform a physico-chemical characterization and biokinetic modeling.

Step 2: Evaluate basal cytotoxicity using, for example, the 3T3 or NHK NRU test methods.

Step 3: Evaluate the potential that metabolism will mediate the basal cytotoxicity effect.

Step 4: Assess the test substance's effect on energy metabolism.

Step 5: Assess the ability of the test substance to disrupt epithelial cell barrier function.

The Workshop 2000 participants suggested that implementation of the 5-step testing scheme would require the following:

- Identification of the most appropriate cell culture systems to use based on accuracy, reproducibility, cost, and availability
- Development of a standardized protocol for each test method used in each of the five steps, and validation of each test method using that protocol
- Development of prediction models for the relevant human toxic levels required by regulatory agencies
- Evaluation of the test battery using substances that are appropriate for all endpoints, and then test sufficient substances to develop a prediction model
- Validation of the entire testing scheme and the prediction model

2.3.6 Basis of the Response Assessed

Neutral red is a weakly cationic, water-soluble, supravital dye that stains living cells by readily diffusing through the cell membranes and concentrating in lysosomes. The intensity of the dye desorbed from the cells in a culture is directly proportional to the number of living cells. Cell death and/or growth inhibition decreases the amount of neutral red taken up by the culture (see **Section 1.3.1**).

2.3.7 Appropriate Positive, Vehicle, and Negative Controls

2.3.7.1 *Positive Control*

The *Guidance Document* recommended sodium lauryl sulfate (SLS; Chemical Abstracts Service Registry Number [CASRN] 151-21-3) as an appropriate PC for *in vitro* cytotoxicity test methods (ICCVAM 2001b), and historical data are available (e.g., Spielmann et al. 1991). A PC test plate was included with every 3T3 and NHK NRU test method assay and was treated the same as any reference substance assay plate.

The historical mean PC IC_{50}, standard deviation (SD), and acceptance limits, were determined separately for each laboratory (see **Table 5-3**), based on their individual historical databases (see **Figure 1-2**). The acceptable range for the PC IC_{50} was based on the statistical approach recommended in the *Guidance Document*. In Phase Ib, the IC_{50} limits accepted for the PC tests were within two SD of the historical mean PC IC_{50} value. In the Phase II studies, the IC_{50} limits for PC tests were within 2.5 SD of the historical mean value (i.e., from Phases Ia and Ib). In Phase III, the IC_{50} limits used for the PC were within 2.5 standard deviations of the mean PC IC_{50} from Phases I and II. The exception to this was the FAL NHK data, where only the Phase II data were used as the basis for establishing the acceptable PC range. The SLS data produced by FAL during Phase I was not used in subsequent historical database compilations because FAL used a modified cell culture protocol in Phase II (see **Section 2.6.2.6**).

2.3.7.2 *Vehicle Control*

The VC consisted of complete DMEM (see **Appendix B1**) for 3T3 cells and complete growth medium (Clonetics® KBM® with supplements [see **Appendix B2**]) for NHK cells when the reference substances were dissolved in culture medium. For reference substances

dissolved in DMSO, the VC consisted of medium with the same amount of DMSO (0.5% [v/v]) as was applied to the 96-well test plate.

2.3.7.3 *Negative Control*

Negative control cultures (i.e., those that were not exposed to the solvent) were not used in this validation study. Neither DMSO, at the concentration used, nor the culture medium affected the performance of the 3T3 and NHK NRU test methods.

2.3.8 Acceptable Ranges of Control Responses

The *Guidance Document* established an absolute value (i.e., uncorrected for blank absorbance) range of the OD_{540} for the VC to indicate whether the cells seeded in the 96-well plate had grown with a normal doubling time during the assay. A mean $OD_{540} \geq 0.3$ was recommended as the acceptable range of VC responses and was made a test acceptance criterion for both cell types at the start of the study. However, prior to Phase II, this was rescinded as a test acceptance criterion. The protocols for Phases II and III provide a range of OD values for use as guidance in future studies with these test methods (**Table 2-1**).

Table 2-1 Measured VC OD_{540} Values[1] and Targets

Laboratory	Phase Ia	Phase Ib	Phase II	Phase III
3T3 NRU Test Method				
Target Range[2]	0.3≤ OD ≤1.1	0.30≤ OD ≤0.80	0.103≤ OD ≤0.813	0.103≤ OD ≤0.813
ECBC	0.326 – 0.457	0.214 – 0.839	0.217 – 0.730	0.191 – 0.797
FAL	0.490 – 0.780	0.247 – 0.742	0.289 – 0.768	0.126 – 1.161
IIVS	0.336 – 0.538	0.319 – 0.598	0.307 – 0.578	0.256 – 0.544
NHK NRU Test Method				
Target Range[2]	0.3≤ OD ≤1.1	0.60≤ OD ≤1.70	0.35 ≤ OD ≤ 1.50	0.205 ≤ OD ≤ 1.645
ECBC	0.863 – 2.312	0.788 – 1.282	0.139 – 1.175	0.114 – 1.344
FAL	0.484 – 1.698	0.146 – 1.706	0.110 – 1.292	0.183 – 1.347
IIVS	0.550 – 1.883	0.487 – 1.001	0.201 – 0.841	0.430 – 0.834

Abbreviations: VC=Vehicle control; OD_{540}=Optical density at 540 nM; 3T3=BALB/c 3T3 fibroblasts; NRU=Neutral red uptake; NHK=Normal human epidermal keratinocytes; ECBC=Edgewood Chemical Biological Center; FAL=Fund for Replacement of Animals in Medical Experiments Alternatives Laboratory; IIVS=Institute for *In Vitro* Sciences.
[1]Lowest to highest OD values for tests that meet test acceptance criteria.
[2]Ranges used for all laboratories. Ranges for Phases Ia and Ib were test acceptance criteria. Ranges for Phases II and III were used as target ranges, rather than as test acceptance criteria.

In Phase III, 99.5% (914/919) of all 3T3 mean VC OD values and 97% (913/944) of all NHK mean VC OD values were within the target ranges. Most OD values outside the ranges were from range finding tests and were usually the result of volatile reference substances affecting the VC cells adjacent to the highest reference substance concentration wells.

The VC OD values had a tendency to be lower in Phases II and III as compared to Phases Ia and Ib. Protocol revisions made throughout Phases Ia, Ib, and II (as listed below) most likely contributed to the differences in the OD values. Possible explanations for changes in OD values for the 3T3 cells include:

- Some tests in Phases Ia and Ib exhibited NR crystals that caused higher OD readings.

- Cell seeding densities were revised from 2.5 x 10^3 cells/well to a range of 2.0 – 3.0 x 10^3 cells/well.

Possible explanations for changes in OD values for the NHK cells include:

- The minimum percent confluence of cells necessary before the reference substance could be applied was reduced from 30% to 20% confluence.
- Cell growth was reduced in some tests in the later study phases as a result of medium and supplement issues (e.g., certain lots of basal medium and medium supplements for NHK cells did not provide optimum growth conditions for the keratinocytes).

2.3.8.1 *Vehicle Controls as a Quality Control Tool*

To check for systematic cell seeding errors and reference substance volatility, VCs were placed both at the left side (row 2) and the right side (row 11) of the 96-well plate (see **Figure 1** in **Appendix B1**). Volatile reference substances generally affected the left side VC, which was next to the highest reference substance concentration in the 96-well plate. The test acceptance criterion for the VC was that the means for the left and the right set of VCs had to be within 15% of the mean of all VCs. This criterion, which was adopted from the protocols in the *Guidance Document* (ICCVAM 2001b), was used for reference substances and the PC in all phases of the validation study.

2.3.9 Nature of Experimental Data Collected

Each laboratory maintained a study workbook to document all aspects of the study and included the raw data for all steps of each assay (e.g., cell growth, test substance treatment, weighing and dilution of reference substances), as well as for all solubility studies.

2.3.9.1 *NRU OD Measurements*

At the conclusion of the NRU desorb step, the OD of the resulting colored solution in each well of the 96-well plates was measured at 540 ±10 nm in a spectrophotometric microtiter plate reader. Each laboratory followed its in-house Standard Operating Procedure (SOP) for use of the microplate readers. These SOPs included instructions for operation and calibration of the instruments. Critical specifications such as alignment, accuracy, reproducibility, and linearity were included as standard parameters for review and routine calibration. Raw OD data from the plate reader was electronically transferred to the EXCEL® template. The template converted the raw data from each treatment well (six wells/reference substance concentration) to derived data by subtracting the mean blank OD value (two blank wells/reference substance concentration) from each reference substance well OD. There were 12 VC wells and 20 associated blank wells. The corrected VC OD values were used to calculate the mean VC OD, which was then used to calculate relative viability (% of mean VC OD) in each test well for the reference substance or PC. The percent viability values were then transferred to the PRISM® template to calculate the IC_{20}, IC_{50}, and IC_{80} values.

2.3.9.2 *Information and Data Collected*

Originals of the raw data (i.e., the Study Workbook and computer printouts of absorbance readings from the plate reader) and copies of other raw data, such as instrument logs, were collected and archived under the direction of the Study Director according to Good Laboratory Practice (GLP)-compliant procedures.

The Study Director/technicians entered the following information into the EXCEL® template:

- Testing identification for: test facility, chemical code, study number, 96-well plate number, experiment number
- Reference substance preparation: solvent used, solvent concentration in dosing solutions, highest stock concentration, dilution factor, pH of 2X dosing solutions, medium clarity/color, presence/absence of precipitate in 2X solutions, PC concentration range
- Cell line/type: cell supplier, lot number, cryopreserved passage number, passage number in assay
- Cell culture conditions: medium, supplements, suppliers and lot numbers, serum concentrations
- Timeline: dates of cell seeding, dose application, OD_{540} determination
- Raw data: OD values from each well from the microtiter plate reader
- Test results: mean corrected OD_{540} value, Hill function R^2 value, logs of IC_{20}, IC_{50}, and IC_{80} (PRISM® template presents data as logs of the IC_x; EXCEL® converts values to μg/mL)
- Test acceptance criteria: acceptable number of values on each side of the IC_{50} (i.e., number of points >0 and ≤50% viability, and >50 and <100% viability), acceptable percent difference for the VCs, acceptable Hill function R^2 value (coefficient of determination) and calculated IC_{50} concentration for the PC
- Visual observations: protocol codes for cell culture conditions for all reference substance concentrations (i.e., relative level of cell cytotoxicity, cell morphology, presence of precipitate)

2.3.10 Data Storage Media

Raw and derived data from the NRU tests were saved in the EXCEL® template file format provided by the SMT. All EXCEL® and PRISM® files were copied and transferred to compact disks. NICEATM and the laboratories printed copies of all data sheets (stored at NICEATM and at the testing facilities), and included copies in the laboratories' final reports.

2.3.11 Measures of Variability

Each 96-well plate used in the NRU tests had three main measures of variability.

1) Each plate contained VCs on each end of the plate (columns 2 and 11) (see **Figure 1** in **Appendix B1** for plate map). The difference between the mean NRU OD for each VC column and mean of the pooled VC wells was used as a test acceptance criterion. The Study Director rejected the test if the difference was greater than 15%, which indicated cross-contamination from a volatile substance or possible cell seeding errors.
2) A mean relative viability was determined for each concentration of the substance tested along with the SD and coefficient of variation (%CV=SD/mean x 100).
3) Macros were included in the EXCEL® template to perform an outlier test (Dixon and Massey 1981) on the data for the six replicate wells for each concentration. Outliers (i.e., individual well values that exceeded the 99% confidence interval [CI] for the replicate wells) were highlighted and could be excluded from the resulting analysis to improve curve fit. The Study Director made the decision as to whether or not to remove outliers and provided a justification for the decision.

Other test-to-test measures of variability were considered in this study.

- Each set of assays for reference substances included a PC plate. If the SLS PC test did not meet test acceptance criteria, then the tests for the associated reference substances were rejected. The SMT recommended testing a manageable number of definitive test plates (e.g., 4 to 6) with each PC to limit the number of definitive NRU tests rejected for PC failure. In this validation study, 4.2% of all definitive tests performed were rejected because the PC failed (i.e., the PC IC_{50} was outside the acceptable confidence limits).
- SDs and CVs were determined for mean IC_{50} values from replicate tests. Replicate testing included three definitive tests for each reference substance, each performed on a different day.

2.3.12 Methods for Analyzing NRU Data

Relative cell viability for each reference substance concentration was calculated using the ODs of the six replicate values (minimum of four acceptable replicate wells) per test concentration. Relative cell viability was expressed as a percentage of the mean VC OD. Absolute OD data from the microtiter plate reader was transferred to the EXCEL® template for performance of these calculations. Where possible, the concentration range (eight concentrations) tested for each reference substance ranged from no effect to 100% toxicity.

The IC_{20}, IC_{50}, and IC_{80} values were determined from the concentration-response curve using the PRISM® template and applying a Hill function to the % viability data. The IC_{20} and IC_{80} values were calculated for potential use in the development of a human prediction model (reported elsewhere).

2.3.13 Decision Criteria for Classification of Reference Substances

The 3T3 and NHK NRU test methods will not be used to classify reference substances in hazard categories but rather to aid in setting the starting dose for sequential rodent acute oral toxicity test methods (i.e., the UDP and ATC) (see **Section 10** for an analysis of the estimated animal savings). The RC millimole regression procedure was used to predict a rodent LD_{50} value from an NRU IC_{50} value. **Section 6.3** addresses the accuracy of the 3T3 and NHK NRU test methods for predicting GHS hazard categories when used with IC_{50}-LD_{50} regressions, calculated using a subset of the RC data (i.e., substances with rat oral LD_{50} data).

2.3.14 Information and Data Included in the Test Report

Test and Control Substances

With the exception of the PC, the laboratories tested coded substances and had minimal information about the test substances' properties (see **Section 3.3** for the reference substance information provided to the laboratories). The following describes the test and test substance information that should be included in an NRU test method report.

- Chemical name(s) and synonyms, if known
- The CASRN, if known
- Formula weight, if known
- Purity and composition of the substance or preparation (in percentage[s] by weight)
- Physicochemical properties (e.g., physical state, volatility, pH, stability, chemical class, water solubility)

- Solubilization of the test/control substances (e.g., vortexing, sonication, warming, grinding) prior to testing, if applicable

Information Concerning the Sponsor and the Test Facility

- Name and address of the sponsor, test facilities, study director, and participating laboratory technicians
- Justification of the test method and specific protocol used

Test Method Integrity

- The procedure used to ensure the integrity (i.e., accuracy and reliability) of the test method over time (e.g., use of the PC data)

Criteria for an Acceptable Test

- Acceptable VC differences between each column of wells and the mean of both columns
- Acceptable concurrent PC ranges based on historical data (include the summary historical data)
- Number of toxic points on either side of the IC_{50} (i.e., number of points >0 and ≤50% viability and >50 and <100% viability)

Test Conditions

- Experiment start and completion dates
- Details of test procedures used
- Test concentration(s) used and how they were derived
- Cell type used and source of cells
- Description of modifications made to the test procedure
- Reference to historical data of the test model (e.g., solvent and PCs)
- Description of the evaluation criteria used

Results

- Tabulation of data from individual test samples (e.g., IC_{50} values for the reference substance and the PC, absolute and derived OD readings, reported in tabular form, including data from replicate repeat experiments as appropriate, and the means and standard deviations for each experiment)

Description of Other Effects Observed

- Cell morphology, precipitate, NR crystals, etc.

Discussion of the Results

Conclusion

Quality Assurance (QA) Statement for GLP-Compliant Studies

- A statement describing all inspections and other QA activities during the study, and the dates results were reported to the Study Director. This statement will also serve to confirm that the final report reflects the raw data.

During the validation study, the GLP-compliant laboratories, IIVS and ECBC, followed additional reporting requirements provided in the relevant GLP guidelines (e.g., OECD 1998; EPA 2003a, b; FDA 2003).

The SMT and laboratories developed standard forms for data collection (i.e., EXCEL® and PRISM® templates). The solubility test form was derived from a standard form provided by

IIVS. The EXCEL® template was an adaptation of a template format presented in the *Guidance Document* (ICCVAM 2001b).

2.4 Proprietary Components of the *In Vitro* NRU Test Methods

The only proprietary components used in these test methods are the NHK cells and the NHK basal culture medium obtained from CAMBREX Clonetics®. All other components are readily available through various scientific product suppliers.

Section 2.3.1.2 describes the NHK cells used in the study and provides the only commercial source. All laboratories throughout the entire study used cells from the same lot. Procedures used to verify the integrity of the NHK cells included comparison of positive control data across laboratories and observations of cell growth throughout the study. If a laboratory reported a problem with the cells, the SMT and Study Directors evaluated the testing parameters to decide if the problem was cell-oriented or if other factors influenced the problem. **Section 2.6.3.5** provides information concerning the resolution of cell-related issues and revisions made to the protocols to address such difficulties.

Section 2.10.1.1 and **Appendices B2** and **B4** provide information about the NHK growth medium, supplements, and commercial source. Problems arose with the keratinocyte growth medium during the study and resolutions and outcomes are addressed in **Sections 2.6.3.5**, **2.6.3.6**, **5.3.4**, and **11.1.2.2**.

Although this study used proprietary components for the NHK NRU test method, cells and medium from the commercial source used in the study are not required for implementation of this test method.

2.5 Basis for the Number of Replicate and Repeat Experiments for the 3T3 and NHK NRU Test Methods

The study protocols required each laboratory to test each coded reference substance in at least one range finding test using a log dilution factor, and in at least three definitive tests on three different days using a smaller dilution factor than used in the range finding test. Assays were performed over a number of days to evaluate day-to-day variation. Laboratories tested each coded reference substance until three definitive tests met the test acceptance criteria. Additional testing was often dictated by:

- Chemical issues (low toxicity, volatility, insolubility, and precipitation)
- PC failure
- Technical difficulties such as NR crystal formation

A stopping rule for insoluble reference substances was incorporated into the protocols for Phase III to limit the number of retests (see **Appendices B1** and **B2**):

> "If the most rigorous solubility procedures have been performed and the assay cannot achieve adequate toxicity to meet the test acceptance criteria after three definitive tests, then the Study Director may end all testing for that particular chemical."

2.6 Basis for Modifications to the 3T3 and NHK NRU Test Method Protocols

2.6.1 Phase Ia: Laboratory Evaluation Phase

All protocol revisions were implemented during Phase Ia unless otherwise stated.

2.6.1.1 *NR Dye Crystals*

NR dye crystals formed in the 96-well test plates when used at 50 μg/mL (OD values measured in the blanks increased from ~ 0.05 to 0.10) in both NRU test procedures. Troubleshooting efforts included incubating the NR medium overnight; centrifuging and filtering the NR medium prior to application to the 96-well plates; and reducing the concentration of NR dye. The laboratories performed tests using a reduced NR concentration of 33 μg/mL. Since there were no quantitative differences in results between tests with 50 μg/mL and tests with 33 μg/mL NR, the SMT accepted tests with both concentrations.

Protocol Revision: The NR dye concentration was reduced to 33 μg/mL for both cell types in subsequent test Phases.

2.6.1.2 *3T3 Cell Growth*

The growth rate of 3T3 cells (as determined by monolayer confluence) was slower than expected. As a result, the cells required more time in culture to obtain the proper density after seeding.

Protocol Revision: The 3T3 cells must be passaged 2-3 times after thawing before being used for the test. The protocol also emphasized attainment of the appropriate percentage of cell confluence (not more than 50% for 3T3 cells) required at the time the cells were exposed to the reference substance, rather than using the time in culture as the guide.

2.6.1.3 *NHK Cell Growth*

The NHK cells had an additional growth problem that manifested as a ring of dead/dying cells around the center of the wells. Troubleshooting efforts included evaluating various brands of 96-well plates (laboratories were not required to use the same brand of plates) and eliminating the change of medium prior to reference substance treatment. All laboratories participated in evaluating the effect of changing (i.e., refeeding) or not changing (i.e., no refeeding) the medium by performing a small study with the PC (SLS). Tests were performed: 1) after refeeding the cells with fresh medium, and 2) by adding SLS to the medium already on the cells. Control ODs were generally higher in the tests in which the medium was not replenished, but sensitivity to SLS was generally unchanged (see **Table 2-2**). FAL was experiencing difficulties in NHK cell growth at this stage of the study which may account for the difference in the refeeding and no refeeding SLS IC_{50} values. The SMT accepted tests with refeeding and those without refeeding (for Phase Ia) as long as they met the test acceptance criteria.

IIVS presented detailed information on the ring of dead cells issue (Raabe 2004). The laboratory showed that the ring of cell death coincided with the formation of a meniscus resulting from the residual medium left in the well after removal of the spent medium. The problem was resolved by eliminating the removal of medium before applying test chemical rather than requiring a standard brand of 96-well plates.

Protocol Revision: Step 2 of the NHK NRU test method was eliminated (change of medium prior to addition of reference substance). The volume of medium (with cells) was changed from 250 μL/well to 125 μL/well.

Table 2-2 Refeeding/No Refeeding Data for the NHK NRU Test Method

	ECBC		IIVS		FAL	
	Refeed	**No Refeed**	**Refeed**	**No Refeed**	**Refeed**	**No Refeed**
Number of Test Plates	4	4	6	6	2	4
Absolute OD[1] for VC	0.265 ±0.151	0.621 ±0.322	0.885 ±0.057	1.12 ±0.033	1.41 ±0.127	1.24 ±0.430
OD[1] for SLS IC_{50}	0.102 ±0.079	0.282 ±0.165	0.415 ±0.029	0.533 ±0.017	0.696 ±0.065	0.606 ±0.217
SLS IC_{50} (µg/mL)[1]	3.33 ±0.47	3.23 ±0.61	3.41 ±0.58	3.49 ±0.39	6.21 ±0.88	8.14 ±0.40

Abbreviations: NHK=Normal human epidermal keratinocyte; NRU=Neutral red uptake; ECBC=Edgewood Chemical Biological Center; IIVS=Institute for *In Vitro* Sciences; FAL=Fund for the Replacement of Animals in Medical Experiments Alternatives Laboratory; VC=Vehicle control; OD=Optical density; SLS=Sodium lauryl sulfate.
Note: OD values for SLS IC_{50} were extrapolated from the concentration-response curve data
[1]Mean ± standard deviation (uncorrected for blank absorbance

FAL, in contrast to the other two laboratories, used 80 cm^2 culture flasks for culturing the thawed cells from the ampules of cryogenically-preserved pool of cells and encountered difficulties in obtaining a satisfactory number of adhering NHKs.

Protocol Revision (FAL only): Culture flasks were coated with fibronectin-collagen to promote cell adherence.

2.6.1.4 *Vehicle Control OD Limits*

In Phase I, the acceptable range of VC OD values designated in the protocols ($0.3 \leq OD \leq 1.1$) were frequently unattainable in both test methods. Despite this, the Study Directors reported that the cells were adequately responsive. The SMT withdrew the VC OD limits as a test acceptance criterion.

Protocol Revision for Phase Ib: OD ranges were provided as guidelines for each cell type based on OD data from all laboratories, a review of the concentration-response data, and the ability of each test to pass the other test acceptance criteria. Each laboratory developed its own VC OD acceptability range based on its historical data.

2.6.1.5 *Precipitate Formation*

During solubility testing, it was observed that some substances, when tested at the same concentrations, precipitated in the 3T3 medium but not in the NHK medium. When a liquid reference substance (i.e., 2-propanol) produced this effect, the precipitate was attributed to the protein in the serum in the 3T3 medium rather than insolubility.

Protocol Revision: The reference substances were dissolved in 3T3 medium without NCS to make the 2X solutions. The dissolved 2X reference substance was added to medium containing 10% NCS to reach the final 5% NCS and 1X reference substance concentrations.

2.6.1.6 *Dilution Factor*

After a range finder test was performed, the definitive tests were to be performed using a $\sqrt[6]{10}=1.47$ dilution scheme centered on the IC_{50} that was calculated from the range finder. In Phase Ia, the Study Directors, for various reasons related to the specific substance being

tested, sometimes deviated from this requirement and used other dilution factors. The SMT agreed that the dilution factor requirements should be modified to allow more flexibility in setting up tests. The SMT accepted data generated using dilution factors other than the recommended 1.47 for definitive tests if all other test acceptance criteria were met. The use of smaller dilution factors generally increased the number of concentrations in the 10% to 90% viability range, which improved the precision of the IC_{50} calculation.

Protocol Revision: The $\sqrt[6]{10}=1.47$ dilution scheme was a suggested starting range, rather than a specific test acceptance criterion in subsequent test Phases.

2.6.1.7 *Test Acceptance Criteria*

The test acceptance criteria at the beginning of Phase Ia were:

- The IC_{50} for SLS had to be within the 95% CI of the historical PC mean established by the Test Facility (*rescinded after commencement of Phase Ia*)
- The OD_{540} of the VCs (with blank subtracted) had to be ≥ 0.3 and ≤ 1.1 (*rescinded after commencement of Phase Ia*)
- Mean OD values of the left and right VCs (columns 2 and 11 in the 96-well test plate) must not differ by more than 15% from the mean of all VC OD values
- At least two cytotoxicity values, one on either side of the IC_{50} but between 10% and 90% viability, needed to be present (*added after commencement of Phase Ia*)
- The Hill function curve fits ($R^2 > 0.9$ <u>or</u> $0.8 < R^2 < 0.9$) were evaluated on a case by case basis for acceptability by the SMT (*added after commencement of Phase Ia*).

2.6.2 <u>Phase Ib: Laboratory Evaluation Phase</u>

All protocol revisions developed during Phase Ia were implemented during Phase Ib unless otherwise stated.

2.6.2.1 *NR Crystal Formation*

FAL and ECBC routinely observed NR crystals forming in the 96-well test plates in the 3T3 NRU tests when 33 µg/mL NR was used. All laboratories tested 25 and 33 µg/mL NR concentrations and 2- and 3-hour NR incubation periods to determine which NR concentration and incubation period would provide optimal NRU measurements without crystal formation. In addition to determining whether NRU had reached a plateau at these concentrations and incubation times, the laboratories also determined whether the response to SLS differed under these conditions. Crystals were observed only at 33 µg/mL NR when present for three hours. **Figure 2-2** shows that the average OD results were similar for all NR concentrations and incubation periods tested. **Figure 2-3** shows that the SLS IC_{50} values were equivalent at the different NR concentrations and incubation periods. To minimize changes to the 3T3 protocol, the NRU concentration was lowered from 33 to 25 µg/mL, while the NR incubation period was maintained at three hours. The NR concentration and the incubation period for the NHK NRU test method remained at 33 µg/mL and three hours, respectively.

Protocol Revision for Phase II: The NR concentration for the 3T3 NRU test method was reduced to 25 µg/mL for the three-hour incubation period. Revised methods for preparation

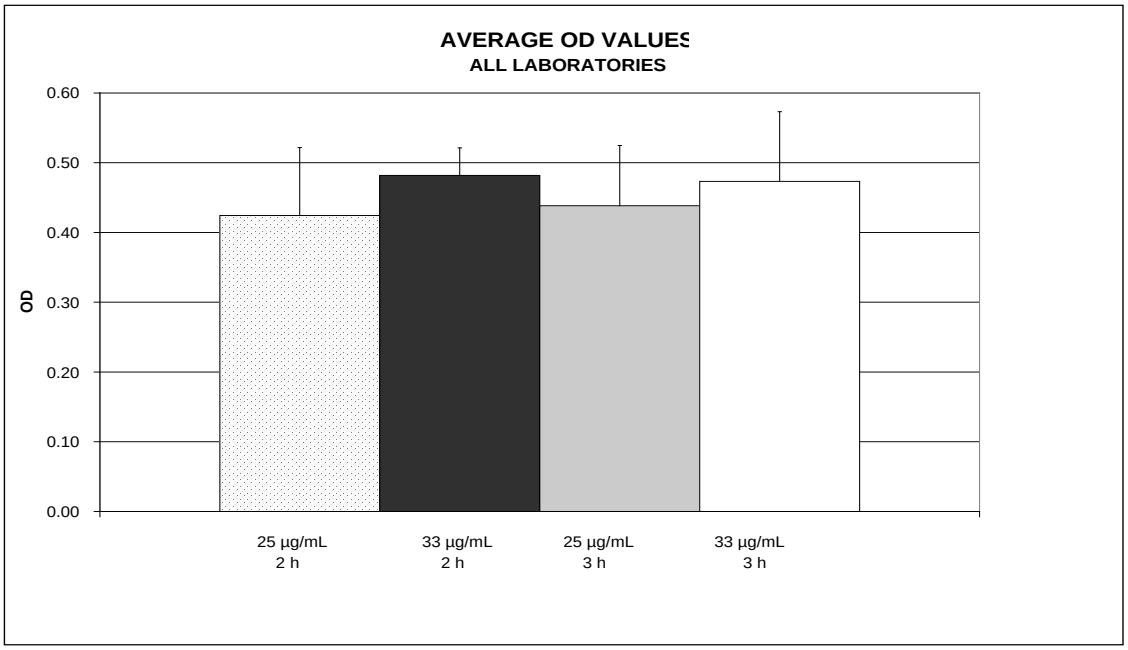
of the NR dye solution included filtration of the solution, maintenance of the solution at 37 ºC prior to application to the cells, and application of the NR solution to the cells within 15 minutes after removing it from 37 ºC. Also, cells were observed during the NR incubation period to monitor possible crystal formation.

2.6.2.2 *Heating of Reference Substance Solutions*

The laboratories had difficulty solubilizing arsenic trioxide, one of the reference substances used in Phase Ib. Heating and mechanical applications for increasing the laboratory's ability to solubilize substances into culture medium were reviewed and revised.

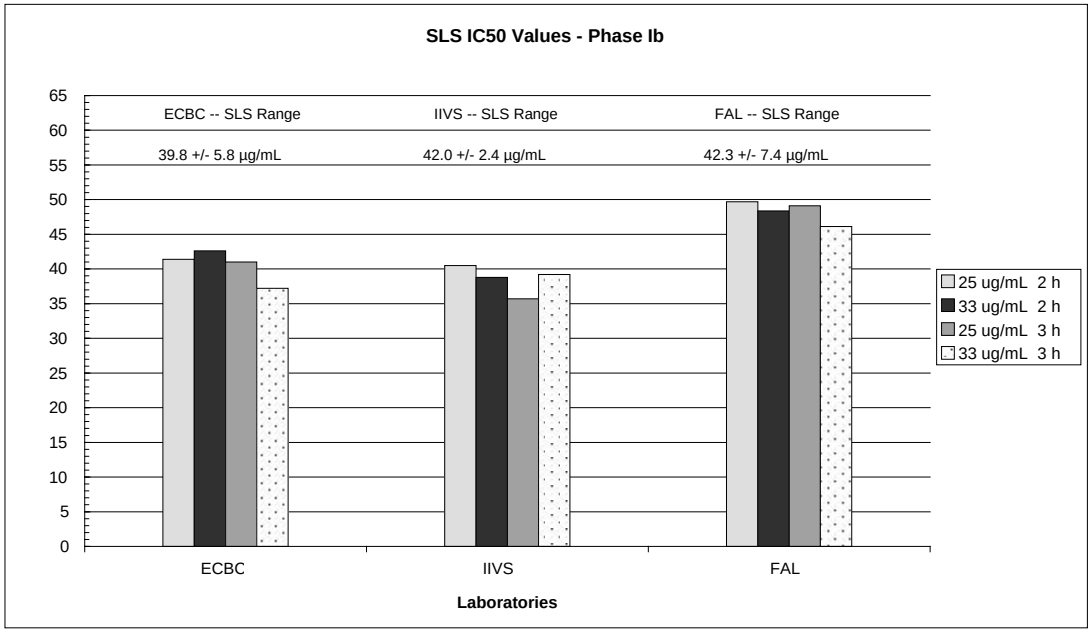
Protocol Revision for Phase II: The duration range for heating a stock solution at 37 ºC (if heating is needed) was increased from 5 to 10 minutes to 5 to 60 minutes.

Figure 2-2 3T3 NRU OD for SLS as a Function of NR Concentration and Duration



Abbreviations: OD=Optical density; NR=Neutral red; SLS=Sodium lauryl sulfate; 3T3=BALB/c 3T3 fibroblasts; NRU=Neutral red uptake; h=Hours.
Note: Error bars are one standard deviation.

Figure 2-3 SLS IC_{50} Values for Each NR Concentration and Incubation Duration (3T3 NRU)



Abbreviations: SLS=Sodium lauryl sulfate; IC_{50}=Test substance concentration that reduces cell viability by 50%; NR=Neutral red; 3T3=BALB/c 3T3 fibroblasts; NRU=Neutral red uptake; ECBC=Edgewood Chemical Biological Center; IIVS=Institute for *In Vitro* Sciences; FAL=Fund for the Replacement of Animals in Medical Experiments Alternatives Laboratory.
Note: SLS range is mean IC_{50} value ± one standard deviation.

Protocol Revision for Phase II: The duration range for heating a stock solution at 37 ºC (if heating is needed) was increased from 5 to 10 minutes to 5 to 60 minutes.

2.6.2.3 *Growth of Untreated Cells*

VC OD values were frequently lower than specified in the Phase I acceptance criteria. Phases Ia and Ib incorporated the acceptance limits shown in **Table 2-1** for the VC, but the limits were rescinded as test acceptance criteria for Phase II because the laboratories frequently failed to meet them even though cell growth and responsiveness to SLS was adequate.

Protocol Revision for Phase II: The specified VC OD range was eliminated as a test acceptance criterion. The OD data (all laboratories combined) from the VCs for both cell types was used to calculate OD ranges that would serve as guidelines for other tests (see **Section 2.2.9).**

2.6.2.4 *Correction of Reference Substance OD Values*

Each reference substance concentration was applied to six treatment wells and to two cell-free wells (i.e., blank wells) used to generate the background OD_{540} values to adjust for potential interference with the NR dye. The mean blank well OD (absolute OD) for each reference concentration was subtracted from the reference substance concentration ODs to provide the corrected OD for each replicate well.

2.6.2.5 *Laboratory Error Rates*

The SMT determined the Phase 1b error rates (number of tests with errors/total number of tests conducted) for each laboratory (**Table 2-3**) and compiled a list of the types of errors encountered. The vast majority of errors were transcriptional and typographical errors in the data sheets provided to the SMT.

Table 2-3 Error Rates[1] in Phase Ib by Laboratory and Test

Laboratory	NRU Test Method	
	3T3	NHK
ECBC	1/9 (10%)	4/17 (23%)
FAL	42/45 (93%)	12/29 (41%)
IIVS	1/20 (5%)	1/20 (5%)

Abbreviations: NRU=Neutral red uptake; 3T3=BALB/c 3T3 fibroblasts; NHK=Normal human epidermal keratinocytes; ECBC=Edgewood Chemical Biological Center; FAL=Fund for the Replacement of Animals In Medical Experiments Alternatives Laboratory; IIVS=Institute for *In Vitro* Sciences
Note: Most errors were transcriptional and typographical and not technical.
[1]Number of tests with errors/total number of tests (some data files had more than one error)

2.6.2.6 *Resultant Protocol Changes for Phase II*

Following the completion of Phases Ia and Ib, IIVS sponsored a weeklong laboratory training exercise for the cytotoxicity testing laboratories to help standardize the level of training among the technical staff and to identify any further 3T3 and NHK NRU protocol revisions that might be needed. Protocol revisions made because of this exercise included:

- Multi-channel repeater pipettes can be used for dispensing cells into the 96-well plates and dispensing plate rinse solutions, NR medium, and desorb solution but are not accurate enough to dispense the PC or the reference substances to the treatment wells.
- Use of 8-channel reservoirs for applying dosing solutions to the wells so that multi-channel single delivery pipettes could be used
- Use of a standardized length of time that the HBSS rinse remains on the cell monolayers in flasks during the cell subculture step
- Protection of plates from light during the shaking step for NR extraction; all laboratories will cover plates with a light-impermeable barrier (e.g., aluminum foil) during this step
- Allow plates to stand for at least five minutes after the shaking step is complete and eliminate any bubbles in media observed in the wells before measuring the OD
- Change the allowable seeding density range for 3T3 NRU test method from $2.5x10^3$ cells/well to $2 – 3x10^3$ cells/well
- Change the NHK culture flask size used at FAL for start-up of cryopreserved cells from 80 cm^2 to 25 cm^2 (the size the other laboratories had been using), and discontinue using a fibronectin-collagen coating.

2.6.2.7 *Test Acceptance Criteria*

The test acceptance criteria were revised as follows:

- The IC_{50} for SLS (PC) should be within 2 SDs (approximately 95%) of the historical mean established by each laboratory in Phase Ia.
- The mean OD values of the left and right VCs (columns 2 and 11 in the 96-well test plate) should not differ by more than 15% from the mean of all VC OD values on that plate.
- At least one calculated cytotoxicity value should be between 10% and 50% viability, and one value between 50% and 90% viability.
- The Hill function curve fit ($R^2 > 0.9$ or $0.8 < R^2 < 0.9$) should be evaluated on a case-by-case basis for acceptability by the SMT.
- VC OD criteria were based on Phase Ia data (mean ± two SDs): 0.3 to 0.8 for the 3T3 test method, and 0.6 to 1.7 for the NHK NRU test method (requirement for use of VC OD criteria as test acceptance criteria was rescinded after commencement of Phase Ib)

2.6.3 Phase II: Laboratory Qualification Phase

All protocol and acceptance criteria revisions were implemented during Phase II unless otherwise stated.

2.6.3.1 *Testing of Volatile Reference Substances*

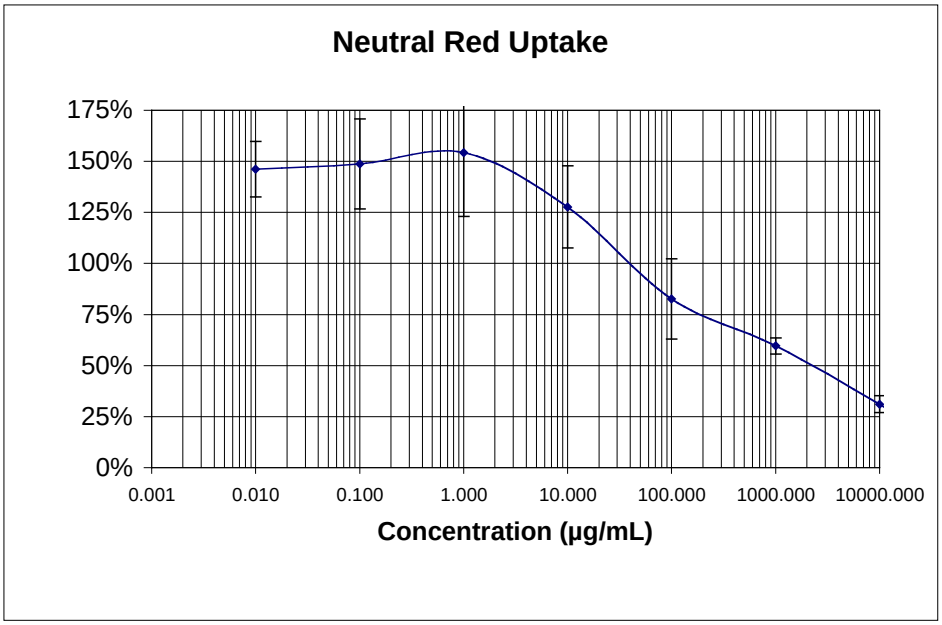
When 2-propanol was tested in 3T3 and NHK cells, vapors from the highest concentration wells contaminated the adjacent VC wells and also appeared to affect some lower concentration wells (i.e., the wells exhibited unexpectedly reduced levels of NRU). An example range finder concentration-response curve is shown in **Figure 2-4**. Such tests failed the VC criterion. When lower concentrations were used to avoid contaminating the VC wells adjacent to the highest concentration, the toxicity was inadequate to produce an IC_{50}. To address this problem, IIVS repeated their tests using film plate sealers, which isolated individual wells from one another; this was sufficient to prevent the cross-well contamination, and acceptable results were obtained. Based on these data, the SMT recommended to the other two laboratories that film plate sealers be used when testing 2-propanol.

FAL had previous experience layering mineral oil on the culture media in a well to prevent volatile substances from escaping, and provided 2-propanol test data where mineral oil had been added to each well. The data showed that the average oil vs. film IC_{50} values were not significantly different. However, there was less variability in the NRU data when using the film sealer so the SMT recommended this methodology.

A >15% difference between the mean VC OD of all VC cells and the mean OD of each VC columns on opposite ends of the test plate was used as a general indicator of substance volatility in the test if the VC adjacent to the highest test concentration had a significantly reduced OD value.

Protocol Revision: The SMT included the use of film sealers in the Phase III protocols when testing suspected volatile compounds.

Figure 2-4 Representative Concentration-Response for 2-Propanol in a 3T3 NRU Range Finder Test



96-WELL PLATE MAP

	1	2	3	4	5	6	7	8	9	10	11	12
A	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank
B	Blank	VC1	C1	C2	C3	C4	C5	C6	C7	C8	VC2	Blank
C	Blank	VC1	C1	C2	C3	C4	C5	C6	C7	C8	VC2	Blank
D	Blank	VC1	C1	C2	C3	C4	C5	C6	C7	C8	VC2	Blank
E	Blank	VC1	C1	C2	C3	C4	C5	C6	C7	C8	VC2	Blank
F	Blank	VC1	C1	C2	C3	C4	C5	C6	C7	C8	VC2	Blank
G	Blank	VC1	C1	C2	C3	C4	C5	C6	C7	C8	VC2	Blank
H	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank	Blank

CORRECTED ABSORBANCE (Sample OD_{540} - Mean Blank OD_{540})

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.000	-0.002	-0.001	-0.001	0.000	-0.003	0.001	0.002	0.002	-0.001	-0.002	-0.003
B	0.002	0.080	-0.001	0.070	0.124	0.206	0.296	0.389	0.291	0.301	0.343	0.002
C	-0.001	0.067	0.004	0.059	0.109	0.171	0.284	0.334	0.237	0.308	0.337	-0.004
D	0.003	0.058	0.003	0.056	0.110	0.163	0.243	0.271	0.246	0.251	0.283	0.002
E	0.003	0.077	0.001	0.067	0.106	0.092	0.218	0.252	0.328	0.250	0.290	0.003
F	-0.004	0.068	-0.002	0.050	0.110	0.164	0.216	0.289	0.336	0.267	0.281	-0.001
G	-0.004	0.071	0.003	0.053	0.122	0.147	0.204	0.226	0.263	0.295	0.330	-0.003
H	0.004	0.000	0.001	0.001	0.000	0.003	-0.001	-0.002	-0.002	0.001	0.001	-0.002

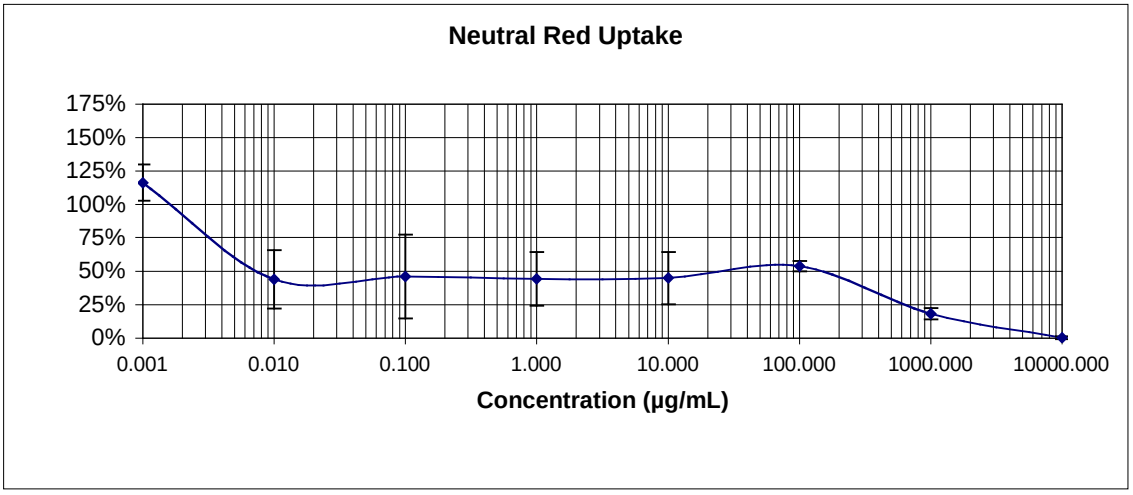
Abbreviations: 3T3=BALB/c 3T3 fibroblasts; VC=Vehicle control; C1 to C8=Test substance concentrations (C1-highest concentration, C8-lowest concentration); OD_{540}=Optical density at 540 nm; A to H=Row identification.
Note: %Difference of the two VC columns from the average VC was 63%. The mean corrected optical density (OD) for VC1, adjacent to the highest 2-propanol concentration, was 0.070, while that for VC2, adjacent to the lowest 2-propanol concentration, was 0.310. Setting the mean VC OD to 100% viability shifted the toxicity curve such that lower concentrations of 2-propanol seemed to be less toxic to the cells than the VCs (i.e., >100%).
Error bars are ±1 standard deviation.

2.6.3.2 *Atypical Concentration-Responses*

Atypical concentration-responses are defined for this study as response curves that differ from a basic sigmoidal shaped curve. Curves that show a biphasic response as well as those that exhibited a plateau-like response at toxicity levels than 100% were considered atypical.

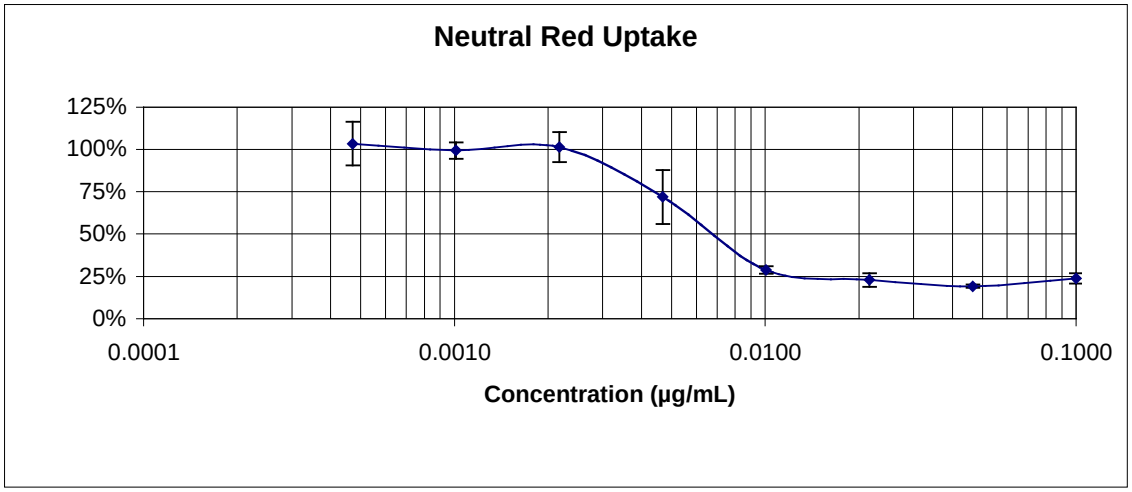
Two of the laboratories observed biphasic concentration-responses in the range finder tests for aminopterin and colchicine. When the range finder tests produced a biphasic response (see **Figure 2-5** for an example), the SMT advised the laboratories to focus the definitive tests on the lowest concentrations that produced at least a 50% loss in viability. Although doing so eliminated the biphasic response in the definitive tests, the highest tested concentrations did not reduce cell viability to 0% (see **Figure 2-6**). This effect with colchicine was very reproducible across laboratories in the NHK NRU test, but only FAL achieved this biphasic type of response with colchicine in the 3T3 NRU test. Aminopterin produced similar concentration-responses in the NHK NRU test at ECBC and FAL, but not at IIVS. In the 3T3 NRU test, only FAL obtained a biphasic response with aminopterin.

Figure 2-5 Representative Concentration-Response for Aminopterin in a NHK NRU Range Finder Test



Abbreviations: NHK=Normal human epidermal keratinocytes; NRU=Neutral red uptake.
Representative dose-response for aminopterin in a NHK range finder test. Laboratories were instructed to focus the definitive tests on the lowest concentration that produced a 50% reduction in viability in the range finder test.

Figure 2-6 Representative Concentration-Response for Aminopterin in a NHK NRU Definitive Test



Abbreviations: NHK=Normal human epidermal keratinocytes; NRU=Neutral red uptake.
Note that the maximum reduction in cell viability plateaued at about 75%

Biphasic concentration-responses are not uncommon. Calabrese (2005) states that numerous mechanistic explanations (including hormesis[4]) could account for biphasic response curves. Such concentration-responses could be because the substance acts through more than one mechanism of action (e.g., one mechanism that is active at low test substance concentrations and other mechanism[s]) that are effective at higher concentrations). Conolly and Lutz (2004) also provide examples of pharmacological and toxicological data sets of biologically based mechanisms that could explain biphasic responses. These examples include:

- Membrane receptor subtypes with opposite downstream effects
- Receptor-mediated gene expression
- Induction of DNA repair and "co-repair" of background DNA damage
- Modulation of the cell cycle

Although non-linear responses could also be due to technical error (e.g., improper dosing, unacceptable media, contamination), the responses seen in this study were reproducible, and there was no evidence to suggest that technical errors were involved. The SMT assumed that these responses were based on the chemicals' mechanisms of action. For example, colchicine binds to microtubular protein and interferes with function of mitotic spindles, which arrests cell division (NLM 2003). Aminopterin blocks the use of folic acid by the cells, inhibiting metabolism, RNA production, and protein synthesis, which is lethal during the S phase of the cell cycle by (NLM 2002). The variability of IC_{50} results for these substances among the laboratories may be due to different levels of cell confluence in the cultures at the time of treatment.

[4] Hormesis is a dose-response characterized by a compound's ability to produce an opposite effect at low doses compared with its effect at high doses (e.g., stimulatory at low doses and inhibitory at high doses).

2.6.3.3 *Hill Function*

The Hill function used in the various phases of this study was defined as follows:

$$Y = \text{Bottom} + \frac{\text{Top} - \text{Bottom}}{1 + 10^{(\log EC50 - \log X)\text{HillSlope}}}$$

where Y=response (i.e., % viability), X is the substance concentration producing the response, Bottom is the minimum response (0% viability, maximum toxicity), Top is the maximum response (maximum viability), EC_{50} is the substance concentration at the response midway between Top and Bottom, and HillSlope describes the slope of the curve. When Top=100% viability and Bottom=0% viability, the EC_{50} is the equal to the IC_{50}.

Responses that do not achieve 100% cytotoxicity with increasing substance concentration do not fit the Hill function well. The R^2 values from such tests often failed the acceptance criterion. To obtain a better model fit, the Bottom parameter was estimated without constraints (the previous practice was to use Bottom=0). However, when Bottom≠0, the EC_{50} reported by the Hill function was not the same as the IC_{50} because the Hill function relies on EC_{50}, which is defined as the point midway between the Top and Bottom responses. Thus, the Hill function calculation using the Prism® software was rearranged to calculate the IC_{50} as follows:

$$\log IC_{50} = \log EC_{50} - \frac{\log\left(\frac{\text{Top} - \text{Bottom}}{Y - \text{Bottom}} - 1\right)}{\text{HillSlope}}$$

where IC_{50} is the concentration producing 50% toxicity, EC_{50} is the concentration producing a response midway between the Top and Bottom responses; Top being the maximum response (maximum survival), Bottom is the minimum response (0% viability, maximum toxicity), Y=50 (i.e., 50% response), and HillSlope describes the slope of the response. The X from the standard Hill function equation is replaced, in the rearranged Hill function equation, by the IC_{50}.

IIVS performed the recalculations for their NHK NRU colchicine tests and the SMT performed the necessary recalculations for the other laboratories. Tests that were recalculated by the SMT are noted in the data summaries.

Protocol Revision: The protocol was revised to state that if a range finding test produces a biphasic response, then the concentrations selected for the subsequent tests should cover the most toxic dose-response range.

2.6.3.4 *Insoluble Reference Substances*

Lithium carbonate was insoluble in 3T3 medium. Only ECBC managed to expose 3T3 cells to sufficient lithium carbonate to produce three tests that met the acceptance criteria. Precipitate was reported for two of those tests at the three highest concentrations in the wells. Because the third highest concentration, 510.2 μg/mL, was approximately the IC_{50} (average was 564 μg/mL), the true IC_{50} for lithium carbonate may actually be lower than was

calculated, and therefore the LD_{50} value would be underestimated. However, the data were reproducible and were not discarded.

Protocol Revision for Phase III: The protocol was revised to allow an increase in the stirring/rocking duration in an incubator from one to three hours if cytotoxicity in the range finder test was limited by solubility. Also, a *Stopping Rule for Insoluble Chemicals* was added (see **Section 2.5** and **Appendices B1** and **B2**) so that the laboratories would not continue repeated testing of insoluble substances in order to obtain three acceptable definitive tests.

2.6.3.5 *Inadequate Cell Growth in NHK Medium*

IIVS and FAL had several NHK NRU test failures that were attributed to poor cell growth. The SMT compiled KBM® and SingleQuot® lot numbers that the laboratories were using, along with the laboratory assessments of NHK cell growth. The information was used to identify the lots that produced adequate growth. The SMT also obtained quality assurance and quality control test results from CAMBREX Clonetics® on the lots of KBM®, but the information provided was inadequate for determining how the medium would perform in the NHK NRU test method.

Resolution: A protocol for prequalifying the medium was developed (see **Appendix B4**). For Phase III, the SMT asked IIVS to prequalify new lots of KBM® and SingleQuots® for use by all laboratories.

2.6.3.6 *Performance Standards for Media to Support NHK Growth*

A prequalification-of-medium protocol (**Appendix B4**) was developed and used by IIVS to test several different lots of medium and supplements to find combinations that maintained the typical growth characteristics of the NHK cells used in this study. The laboratories then reserved samples of the acceptable lots at CAMBREX so that testing would not be interrupted due to unavailability of adequate materials.

Test Acceptance Criteria for Prequalifying Media Using SLS

- The fit of the SLS dose-response to the Hill model should be $R^2 \geq 0.85$ (i.e., from PRISM® software).
- The difference between the mean of all VCs and (a) the left mean VC, and (b) the right mean VC should be ≤15%.
- At least one concentration should exhibit >0% and ≤50% viability and at least one should exhibit >50% and <100% viability.
- After meeting all other acceptability criteria, the SLS IC_{50} must be within the historical range (±2.5 SD) established by the laboratory.

Other Criteria for Prequalifying Media (for consideration by a Study Director)

- General observations: rate of cell proliferation; percent confluence; number of mitotic figures per field; colony formation; distribution of cells in the flask; absence or presence of contamination
- Cell morphology observations should include overall appearance (e.g., good, fair, poor), and presence of abnormal cells

- Mean corrected OD_{540} of the VCs (e.g., are the values high/low when compared to historical data)
- Cell morphology and confluence of the VC wells at the end of the 48-hour treatment
- Cell doubling time, as compared to the doubling time with the previous batches of medium

2.6.3.7 *Test Acceptance Criteria for Phase II*

- The IC_{50} for SLS (PC) should be within 2.5 SDs of the historical mean established by the laboratory (*Phases Ia and Ib*)
- Mean OD values of the left and right VCs (columns 2 and 11 in the 96-well test plate) do not differ by more than 15% from the mean of all VC well OD values. At least one calculated cytotoxicity value ≥10% and ≤50% viability and at least one value >50% and ≤90% viability
- $R^2 \geq 0.90$. The test fails if $R^2 < 0.80$. If the $0.80 \leq R^2 < 0.90$, the SMT evaluates the model fit (Note: The Study Director makes this determination for non-validation studies.)

2.6.4 Phase III: Laboratory Testing Phase

The changes below were made in the Phase III protocols based on the data and results in Phase II.

2.6.4.1 *Required Cytotoxicity Values*

Obtaining at least one calculated cytotoxicity value >0% and ≤50% viability and at least one that is >50% and <100% viability may be difficult or unattainable for substances with steep dose responses.

Protocol Revision: The test acceptance criterion was qualified so that tests with only one concentration between 0 and 100% viability were acceptable if the smallest practical dilution factor (i.e., 1.21) was used and all other test acceptance criteria were met.

Tests for three reference substances were accepted that met this new criterion in the 3T3 NRU test method: diquat dibromide (1/9 tests); epinephrine bitartrate (2/9 tests); 1,1,1-trichloroethane (2/8 tests). No NHK tests required the use of these criteria (i.e., one point between 0% and 100% viability at the lowest dilution factor).

2.6.4.2 *Revisions to Data Analysis Procedures*

The following revisions to data analysis procedures were made in Phase III NRU protocols:

- If the Bottom parameter of the Hill function was fit to a value <0%, then the parameter was set to zero (0) for the IC calculations.
- If toxicity plateaued above 20% viability (i.e., toxicity was <80%), the IC_{80} was not determined. The IC_{20} and IC_{50} values were calculated from the range of available toxic responses.
- The requirement for substance dose-responses to fit the Hill equation with $R^2 \geq 0.90$ was rescinded. The Hill equation was used to characterize the shape of the response rather than to establish an acceptance criterion. The PC acceptance criterion was modified to $R^2 \geq 0.85$.

2.7 Differences Between the 3T3 and NHK NRU Protocols for the Validation Study and the *Guidance Document* Standard Protocols

As the validation study progressed through Phases I and II, the protocols provided in the *Guidance Document* (ICCVAM 2001b) were optimized to address problems that were encountered during the validation study phases. Changes to the *Guidance Document* protocols are described below.

- 3T3 cell seeding density for 96-well plates was decreased from $1x10^4$ cells/well to $2.0 – 3.0x10^3$ cells/well.
- The calcium concentration in NHK medium was changed from 0.15 mM to 0.10 mM. The test laboratories had expressed concern that cell differentiation would occur at the higher concentration and requested a lower concentration. CAMBREX Clonetics®, the supplier of the NHK cells and NHK medium used in this study, normally grows NHK cells in 0.15 mM calcium and has seen no differentiation. The supplier agreed that the cells would grow well at 0.10 mM but should not be cultured at concentrations <0.10 mM in order to avoid morphological and growth rate changes (CAMBREX technical division, personal communication).
- NHK cells were subcultured once prior to being distributed to the test wells, rather than for three passages. The laboratories expressed concern about the possibility of cell differentiation with subsequent passages in culture.
- The highest recommended final concentrations of DMSO and ETOH in the culture media were reduced from 1% to 0.5%. IIVS performed experiments with both cell types to determine the concentration necessary to avoid solvent toxicity. 3T3 cells were tested with 0.5, 1, and 2% ETOH and DMSO at 0.1, 0.2, 0.3, 0.4, 0.5, 1, and 2% concentrations. The 0.5% concentrations of both solvents were chosen as optimal because that concentration of ETOH produced no toxicity. Although 0.5% DMSO produced slight toxicity (i.e., cells were 91% viable as compared to the control cells; See **Appendix E1**), this concentration was chosen by the SMT and laboratories as an acceptable trade-off between slight toxicity and the ability to test substances at higher concentrations, and was used throughout the study for all reference substances that needed solvents other than culture medium (see Curren et al. 2003). DMSO was the preferred solvent if the test substance was not soluble in culture medium, and ETOH was not used in this study.
- The pH of the reference substance solutions was not adjusted with NaOH or HCl regardless of whether solutions became acidic or basic (optimum mammalian cell culture pH is approximately 7.4 [Freshney, 2000]) upon addition of the test substance because some of the basal cytotoxicity produced by test substances may be due to pH effects. See **Appendix F1** for pH values of the reference substances in culture medium.
- The CO_2 concentration in the incubator was reduced from 7.5% to 5.0% because the laboratories were already set up to use 5% CO_2, which is a typical optimum CO_2 concentration for mammalian cell culture.
- Washing and fixing the cells with a formaldehyde solution prior to NR elution from the cells was eliminated. Formaldehyde disposal was problematic in FAL's regulatory environment. The SMT and the laboratories agreed that the

use of formaldehyde was unnecessary because the NR desorb solution (1% glacial acetic acid, 50% ETOH, and 49% H_2O) adequately fixed the cells to the test plate (INVITTOX 1991).

- Reference substance exposure time for the 3T3 cells was extended from 24 hours to 48 hours (see **Section 2.2.4** and **Appendix E1**).
- Cell culture seeding densities for subculture were provided as guidance, rather than as strict cell number ranges. The laboratories determined adequate cell densities (see **Table 2-4**) based on their own experience with the growth of the cells in the wells, and the time needed to reach the appropriate level of confluence needed for addition of the test substance, the VC, and PC.

Table 2-4 Cell Seeding Densities[1]

Protocol	3T3 cells/cm^2 subculture to flasks	3T3 cells/well 96-well Plate	NHK cells/cm^2 subculture to flasks	NHK cells/well 96-well Plate
Guidance Document[2]	$1.25x10^4$	$2.5x10^3$	$3.5x10^3$	$2 – 2.5x10^3$
Phase Ia	$0.42 – 1.68x10^4$	$2.5x10^3$	$2.5 – 9x10^3$	$2 – 2.5x10^3$
Phase Ib	$0.42 – 1.68x10^4$	$2.5x10^3$	$2.5 – 9x10^3$	$2 – 2.5x10^3$
Phase II	$0.42 – 1.68x10^4$	$2 – 3x10^3$	$2.5 – 9x10^3$	$2 – 2.5x10^3$
Phase III	$0.42 – 1.68x10^4$	$2 – 3x10^3$	$2.5 – 9x10^3$	$2 – 2.5x10^3$

Abbreviations: 3T3=BALB/c 3T3 fibroblasts; NHK=Normal human epidermal keratinocytes
[1]Cell numbers determined by Coulter Counter or hemocytometer
[2]ICCVAM (2001b)

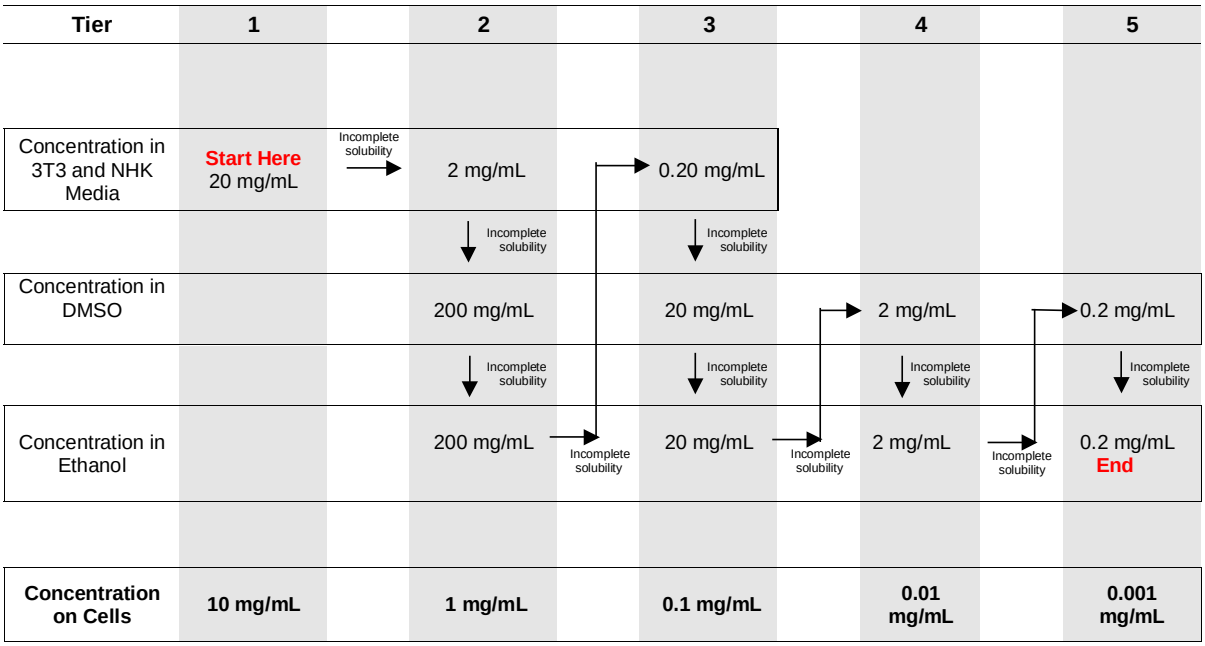
2.8 Overview of the Solubility Protocol

The SMT, with assistance from the laboratories, developed a solubility protocol to provide guidance for determining the most appropriate solvent for each test substance. The solubility protocol was based on an EPA guideline (EPA 1998) that involved testing for solubility in a particular solvent, beginning at a relatively high concentration and proceeding to successively lower concentrations by adding more solvent as necessary for dissolution. Testing stopped when, upon visual observation, the procedure produced a clear solution with no cloudiness or precipitate. The order of selection priority was culture medium, DMSO, and ETOH. Each laboratory tested the solubility of each reference substance using this protocol and provided the data to the SMT prior to initiating cytotoxicity testing. The SMT analyzed the solubility data provided by BioReliance and each testing laboratory, and designated the solvent to be used by all laboratories for each reference substance. This eliminated one potential variable in the NRU test results among laboratories.

The solubility protocol used by the *in vitro* laboratories required the sequential testing of reference substances in the various solvents at concentrations that would be equivalent to the concentration that would be applied to the cell cultures. The solubility flow chart in **Figure 2-7** shows, for example, that 2 mg/mL medium and 200 mg/mL DMSO or ETOH were equivalent concentrations because they yielded 1 mg/mL in cell culture. Medium was diluted by one-half when applied to cultures. The 0.5% [v/v] final concentrations were achieved by diluting DMSO and ETOH by 200-fold. At each concentration, the following mixing

procedures were employed, as necessary, to completely dissolve the reference substance in the sequence: vortex (1 to 2 minutes); sonication (up to 5 minutes); warming to 37 °C (5 to 60 minutes [NRU protocols allow warming to be extended to three hours if cytotoxicity in the range finder test was limited by solubility]). If the reference substance was still not dissolved, the next lower concentration, or a different solvent, was tested.

Figure 2-7 Flow Chart for Determination of Reference Substance Solubility in Medium[1], DMSO, or ETOH



Abbreviations: DMSO=Dimethyl sulfoxide; ETOH=Ethanol; 3T3=BALB/c 3T3 fibroblasts; NHK=Normal human epidermal keratinocytes.
Note: DMSO is U.S.P. analytical grade. ETOH is U.S.P. analytical grade (100% non-denatured).
[1]3T3 Medium - DMEM (Dulbecco's Modification of Eagle's Medium) with supplements; NHK medium - KBM® (Keratinocyte Basal Medium) with supplements (from CAMBREX Clonetics®).

2.9 Basis of the Solubility Protocol

The solubility protocol used by BioReliance, which tested solubility of the reference substances prior to testing by the *in vitro* laboratories, is provided in **Appendix G**. The protocol is based largely on information from the literature and Internet searches for solubility procedures, the experience of the SMT and IIVS, and solubility and IC_{50} information from the RC chemicals database (Halle 1998, 2003). The only formal solubility protocol discovered was the EPA Product Properties Test Guideline, OPPTS 830.7840 Water Solubility Column Elution Method; Shake Flask Method (EPA 1998).

2.9.1 Initial Solubility Protocol Development

BioReliance evaluated the solubility of each reference substance in cell culture media at 2000, 400, and 200 mg/mL, and if not soluble at those concentrations, in DMSO and then ETOH, at the same concentrations (initial protocol). It was apparent that these concentrations

were not low enough when the laboratory was unable to achieve solubility for arsenic trioxide. The solubility protocol was revised twice to lower the range of concentrations tested (see **Table 2-5**). An extra tier of concentrations ≤1 mg/mL was added for poorly soluble and insoluble substances. The protocol used by the laboratories was further revised to reduce the number of steps required (by testing in log units) and to test in tiers using concentrations that reflected the concentrations anticipated in the cell cultures (see **Figure 2-7**).

Table 2-5 Comparison of Concentrations Tested in the Various Solubility Protocols

Solubility Protocol Version	Concentrations Tested (mg/mL)					
	Step 1	**Step 2**	**Step 3**	**Step 4**	**Step 5**	**Steps 6-10**
BioReliance (1st) (4/26/02) and Phase Ia	2,000	400	200	NA	NA	NA
BioReliance (2nd) (9/17/02)	200	40	20	10	2	NA
BioReliance (3rd) (10/11/02)	200	40	20	10	2	1, 0.5, 0.25, 0.125, 0.05
Phases Ib, II, III for cytotoxicity laboratories	20 Medium	2 Medium 200 DMSO 200 ETOH	0.2 Medium 20 DMSO 20 ETOH	2 DMSO 2 ETOH	0.2 DMSO 0.2 ETOH	NA

Abbreviations: DMSO=Dimethyl sulfoxide; ETOH=Ethanol; Medium=Cell culture medium; NA=Not applicable
Note: DMSO is U.S.P. analytical grade. ETOH is U.S.P. analytical grade (100% non-denatured).

In Phases Ib and II, the SMT used the data from BioReliance to select the solvents to be used for testing the various chemicals. When it became apparent that the laboratories sometimes obtained different solubility results than those reported by BioReliance, the SMT used the cytotoxicity results from the laboratories to determine the solvents to be used for Phase III reference substances.

The final protocol provided a tiered approach for determining the 2X stock concentration for each reference substance (see **Figure 2-7**). This protocol had the advantage of reducing the number of steps for testing (compared to that used by BioReliance) (see **Appendix B3**).

2.9.2 Basis for Modification of the Phase II Protocol

All three testing laboratories found arsenic trioxide (tested in Phase Ib) less soluble (see **Table 5-10**) than was reported by BioReliance (BioReliance values: 0.25 mg/mL in 3T3 medium and 0.05 mg/mL in NHK medium). This chemical was not soluble using the procedures in the initial solubility protocol. IIVS warmed the stock solution (at least 200 μg/mL for 2X) for longer than the protocol specified (i.e., 30 to 50 minutes) but still had persistent, small, undissolved particles. ECBC obtained a clear solution (highest 2X concentration was 30 to 50 μg/mL), but found precipitated particles after the solution stood at room temperature. Sonication time was increased to 15 to 30 minutes, and heating time to approximately 30 minutes to get a finer suspension. This procedure achieved a more homogeneous mixture, resulting in more uniform serial dilutions and a more even application of the reference substance to the cells. FAL stirred the suspension (approx. 20 to 90 μg/mL) in the CO_2 incubator for 1.5 to 2 hours to get clear medium.

Protocol Revision for Phase II: The duration of the heating step was altered from 5 to 20 minutes to 5 to 60 minutes.

2.10 Components of the Solubility Protocol

2.10.1 Medium, Supplies, and Equipment Required

2.10.1.1 *Medium and Chemical Supplies*

- 3T3 culture medium: DMEM without L-glutamine and containing Hanks' salts and high glucose [4.5gm/l]; L-glutamine, 200 mM; NCS
- NHK culture medium: Keratinocyte Basal Medium without Ca^{++} (KBM®, Clonetics® CC-3104); KBM® SingleQuots® medium supplements (Clonetics® CC-4131): epidermal growth factor, insulin, hydrocortisone, bovine pituitary extract; Calcium SingleQuots® (Clonetics® CC-4202); penicillin/streptomycin solution (antimicrobial agents)
- United States Pharmacopoeia (U.S.P.) analytical grade DMSO
- U.S.P. analytical grade (100%, non-denatured) ETOH

2.10.1.2 *Equipment*

- Waterbath (37 °C)
- Sonication apparatus
- Vortex mixer
- Micropipettors
- Balance (capable of weighing 10 mg)
- pH meter

2.10.1.3 *Procedures*

The Phase III solubility protocol required the dissolving of approximately 10 mg of reference substance in approximately 0.5 mL medium (both 3T3 and NHK media were used) for a final concentration of 20 mg/mL (see **Appendix B3**). In order, the mixture was vortexed for 1 to 2 minutes, sonicated for up to 5 minutes, and warmed to 37 °C for 5 to 60 minutes, as necessary, to dissolve the substance. The endpoint for dissolution was a clear solution with no noticeable precipitate. If the reference substance was not soluble in medium at 20 mg/mL, then more medium was added to a concentration of 2 mg/mL (i.e., a total volume of approx. 5 mL) (Step 2). The mixing procedures were repeated as necessary to dissolve the reference substance. If the reference substance did not dissolve, approximately 10 mg reference substance was added to approximately 0.5 mL DMSO in an attempt to dissolve it at a concentration of 200 mg/mL (Step 3). If the reference substance was not dissolved, the same concentration was attempted in 100% ETOH (Step 4). Step 5 began in the same way, with 0.2 mg/mL medium and then progressed to 20 mg/mL DMSO, and then 20 mg/mL ETOH.

Determination of reference substance solubility was limited to visual observation of the resulting solution. If a solution appeared clear, then solubility testing ceased. If particles were visible or if the solution appeared cloudy, then more stringent mixing and/or heating procedures were employed. If necessary, the solubility procedure proceeded to the next solvent/concentration tier. The duration of the solubility test was dependent on the procedures used to achieve solubility. Some reference substances were immediately solubilized (e.g., liquids) and others required up to 60 minutes of heating and agitation or sonication.

2.10.2 Data Collection

All laboratories (including the reference substance distribution laboratory, BioReliance) used a worksheet designed to capture the solubility information for each reference substance. The endpoint for each step was a visual observation of the solution, a documented comment describing the observation, the concentration, and a conclusion of soluble or insoluble. Each worksheet contained:

- Reference substance code number and physical description
- Solvent used (3T3 medium, NHK medium, DMSO, ETOH)
- Amount of reference substance (mg) used in the initial stage
- Volume of solvent added and final volume (mL)
- Test substance concentration (µg/mL) in the solvent
- pH and color of the solution
- Mechanical procedures used (vortexing, sonication, heating), duration, and temperature
- Comments (soluble/insoluble at the particular concentration; visual observations; reactivity with solvent)

The solubility test information and data from the laboratories were transferred via email to the SMT and stored on the NICEATM server and as hard-copy printouts. Each laboratory also maintained electronic and hard-copy files of its data.

2.10.3 Variability in Solubility Measurements

Solubility determinations were not replicated because within-laboratory results were not expected to vary. Comparison of the results to determine inter-laboratory concordance for the 72 reference substances (see **Section 5.8** for results) provided a measure of variability among the laboratories and information about the reproducibility of the solubility determinations (see **Section 7.4**).

2.10.4 Solubility Issues During the Testing of the Reference Substances

Substance solutions were monitored throughout all aspects of the test procedures, and observations were documented. The lowest concentration of the substance in a 2X solution that contained observable precipitates, particles, globules, or oily droplets, was documented in the EXCEL® template. After substance exposure, all wells of the 96-well test plates were observed microscopically and scored using a visual observation code. The code addressed growth characteristics and the presence or absence of precipitates (see **Appendix B** [test method protocols] for the observation codes used). For solubility issues, the Study Directors made determinations of test acceptance based on the recommended concentration levels and the presence of precipitates, their scientific expertise, and test acceptance criteria.

2.10.5 Analysis of Solubility Data

During Phase III, the SMT used the solubility data from all laboratories to determine the solvents to be used for each chemical (see **Section 5.8** for solubility results and SMT selections). If the solubility of an individual reference substance was different in 3T3 medium and NHK medium, the same solvent would be used for both test methods, rather than having different solvents for each method. For example, if solubility in one culture medium was ≥2 mg/mL and solubility in the other was <2 mg/mL, and the substance was soluble in DMSO at 200 mg/mL, the SMT would select DMSO as the solvent for both test methods (each test method using its respective culture medium).

Solubilizing sufficient reference substance to produce cytotoxicity was challenging for relatively insoluble, low toxicity, substances such as lithium carbonate (in the 3T3 NRU test method) but generally was not a problem for toxic substances that did not require as high a concentration to kill cells. Some insoluble and highly toxic reference substances were problematic, however, because the amount of powdered reference substance added to solvent was very small, and laboratory personnel found it difficult to determine the presence of solute particles in solution. Arsenic trioxide is an example of such a solute (see **Section 2.9.2**).

2.11 Summary

The *Guidance Document* NRU protocols were used as the basis of the validation study protocols. The SMT and participating laboratories made initial modifications to the protocols prior to implementation of the study. Other protocol modifications were made after commencement of testing and were the result of recommendations from the laboratories and the SMT, based on their experience with the initial protocols. The resulting optimized protocols were used in the main testing phase (Phase III) of the study.

The protocol components used in the validation study were similar for the 3T3 and NHK cells. The following procedures were common to the NRU protocols for both cell types:

- Testing was performed in four phases (Phases Ia, Ib, II, and III)
- Preparation of reference substances and positive control
- Cell culture environment conditions
- Determination of test substance solubility
- Configuration of 96-well plates for testing samples
- 48-hour exposure to test substance
- Range finder and definitive testing
- Microscopic evaluation of cell cultures for toxicity
- Measurement of NRU
- Data analysis

The main differences in the test methods for the two cell types were:

- The conditions of propagation of the cells in culture
- The cell growth medium components
- The volumes of reference substance added to the 96-well plate

A solubility protocol was developed which allowed the laboratories to identify the most appropriate solvent and appropriate limit concentrations for each test substance.

Three laboratories participated in testing the 72 reference substances in both cell types and one additional laboratory procured and distributed the coded reference substances and performed solubility tests prior to their distribution to the testing laboratories.