NICEATM Pre-Screen Evaluation of the *In Vitro* Endocrine Disruptor Assay (Robotic MCF-7 Cell Proliferation Assay of Estrogenic Activity)

EVALUATION

NICEATM

16 October 2006

Table of Contents

EXECUTIVE SUMMARY

On June 4, 2005, the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) received a nomination from CertiChem, Inc. (CCi) for the validation of a cell based estrogen receptor (ER) transcriptional activation method. The proposed test method evaluates the potential endocrine disruptor activity of substances by measuring whether and to what extent a substance induces cell proliferation via ER dependent pathways.

In support of this nomination, NICEATM received a background review document (BRD) from CCi entitled, "Test Method Nomination: MCF-7 Cell Proliferation Assay of Estrogenic Activity" on January 19, 2006. The BRD contains the historical development and rationale for the test method, test method protocol and supporting materials. The test method was developed and refined by CCi in collaboration with laboratories at the University of Missouri and Northwestern University. Results from studies done at all three facilities were presented in support of the currently proposed test method.

In accordance with the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) nomination process, NICEATM conducted a pre-screen evaluation of the CCi BRD to determine the extent that it addresses ICCVAM prioritization criteria, ICCVAM submission guidelines, and ICCVAM recommendations for standardization and validation of *in vitro* endocrine disruptor test methods. The performance of the proposed test method based on pre-validation data was also reviewed to determine if this performance warrants consideration for further validation. The CCi BRD was reviewed for completeness and to identify aspects or omissions that could impede further review. The CCi BRD was not reviewed with respect to data quality or presentation.

The areas considered in evaluating information provided by CCi in their BRD and the extent to which the criteria are met are as follows:

- the extent to which the BRD addresses ICCVAM prioritization criteria
- the extent to which the BRD provides the information requested in the

ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods (NIH Pub. No. 03-4508)

- the extent to which the proposed test method adheres to the Recommendations of the *ICCVAM Evaluation of In Vitro Test Methods for Detecting Potential Endocrine Disruptors* (NIH Pub. No. 03-4503), especially those regarding essential test method components (previously known as minimum procedural standards) and recommended validation substances
- the extent to which the proposed test method shows adequate performance (reliability and accuracy) during pre-validation to warrant consideration for validation studies

The CCi BRD addressed the specified ICCVAM prioritization criteria. However, there were a number of deficiencies in the organization and content of the BRD that made it difficult to conduct a comprehensive pre-screen evaluation and that could potentially impede a review by an expert peer-review panel. These deficiencies were communicated to CCi and NICEATM subsequently received a revised BRD from CCi on April 27, 2006. The revised CCi BRD adequately addresses the ICCVAM submission guidelines and recommendations for standardization and validation of *in vitro* endocrine disruptor test methods.

1.0 INTRODUCTION

On April 21, 2004, a *Federal Register* (*FR*) Notice (Vol. 69, No. 77, p. 21564) entitled "*In Vitro* Endocrine Disruptor Test Methods: Request for Comments and Nominations" was published that stated:

- The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the Scientific Advisory Committee on Alternative Toxicological Methods (SACATM) had identified *in vitro* endocrine disruptor screening methods as a priority for validation.
- ICCVAM had published guidelines for development of *in vitro* endocrine-disruptor estrogen and androgen receptor (ER and AR) binding and transcriptional activation (TA) assays. In these guidelines, ICCVAM recommended that priority be given to assays that do not require the use of animal tissue as the receptor source, but rather use recombinant proteins, and do not use radioactive materials.
- On behalf of ICCVAM, the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) invited the nomination for validation studies of *in vitro* test methods that meet these recommendations and for which there are standardized test method protocols, pre-validation data, and proposed validation study designs.
- ICCVAM would consider nominations and comments received in response to this *FR* notice and develop recommended review and validation priorities for endocrine disruptor screening methods.
- Prior to the initiation of such studies, the proposed test method protocols would be evaluated for adherence to relevant recommendations in the report, ''ICCVAM Evaluation of *In Vitro* Test Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays'' (NIH Publication No. 03–4503; http://iccvam.niehs.nih.gov/methods/endocrine.htm) by NICEATM and the ICCVAM Endocrine Disruptor Working Group (EDWG).

On June 4, 2005, NICEATM received a nomination from Certichem, Inc. (CCi) to validate an *in vitro* ER TA text method developed by CCi to quantify the ER activity of unknown chemicals.

On January 19, 2006, NICEATM received a background review document (BRD) from CCi in support of this nomination entitled, "Test Method Nomination: MCF-7 Cell Proliferation Assay of Estrogenic Activity" (CCi submission). The BRD contains the historical development and rationale for the proposed test method, the test method protocol, and supporting materials. The test method was developed and refined by CCi in collaboration with laboratories at University of Missouri (UM) and Northwestern University (NWU). Results from studies conducted at all three facilities were presented in support of the currently proposed test method, which evaluates the estrogenic activity of substances by measuring whether, and to what extent, a substance induces cell proliferation via ER-dependent pathways. Results from studies supporting the reliability of the test method were obtained by repeated testing of ten substances at the three facilities (CCi, UM, and NWU). Accuracy of the test method was supported by comparing ICCVAM published and ranked EC50 values for 18 reference substances recommended by ICCVAM for endocrine disruptor (ED) validation studies with CCi experimentally obtained and ranked EC50 values for these same substances. Accuracy of the test method was further supported by comparing ICCVAM published estrogenic activities (positive or negative) for a larger set of 40 substances on the ICCVAM list with CCi experimentally obtained estrogenic activities for these same substances. The submission does not provide information or supporting data for the test methods ability to detect the anti-estrogenic activity of substances.

2.0 ASSESSMENT OF THE PROPOSED TEST METHOD

The CCi BRD was reviewed for completeness and to identify aspects of omissions that could impede an expert peer review. The BRD was not reviewed with respect to data quality or presentation. Rather, the content of the BRD was evaluated based on the following criteria:

1) the extent to which the BRD addresses ICCVAM prioritization criteria
2) the extent to which the BRD includes the information requested in the

ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods (NIH Pub. No. 03-4508)

3) the extent to which the proposed test method protocol adheres to the recommendations of the *ICCVAM Evaluation of In Vitro Test Methods for Detecting Potential Endocrine Disruptors* (NIH Pub. No. 03-4503), especially those regarding essential test method components (previously known as minimum procedural standards)
4) the extent to which the proposed test method shows adequate performance (reliability and accuracy) during pre-validation to warrant consideration for validation studies

The CCi BRD received by NICETAM on January 19, 2006 addressed the specified ICCVAM prioritization criteria. However, there were a number of deficiencies in the organization and content of the BRD that made it difficult to conduct a comprehensive pre-screen evaluation and that could potentially impede a review by an expert peer-review panel. These deficiencies were communicated to CCi and NICEATM subsequently received a revised BRD on April 27, 2006 that was further evaluated for the criteria outlined in 2-4 above.

2.1 Extent to Which the Background Review Document Addresses the ICCVAM Prioritization Criteria

2.1.1 Applicability to Regulatory Testing Needs

The U.S. Environmental Protection Agency Endocrine Disruptor Steering and Testing Advisory Committee (EDSTAC) identified a need for high throughput screening (HTS) systems to test substances for their potential ED activity. The EDSTAC concluded that prior methods were not adequate to detect many types of ED activity and recommended the development of a tiered Endocrine Disruptor Screening Program. This program would consist of two tiers, with the first tier (Tier 1) consisting of *in vitro* and *in vivo* tests designed to identify substances that have the potential to interact with the endocrine system, with a preference for automated HTS test methods as a way to streamline the detection process.

The second tier (Tier 2) would contain *in vivo* multigenerational tests to confirm endocrine ED activity and effects. The CCi test method is intended as an automated HTS Tier 1 screening system to test substances for their potential ED activity.

2.1.2 Applicability to Multiple Agencies or Programs

The CCi test method should be applicable to the needs of all agencies and programs that require testing for potential ED activity in pharmaceuticals, chemicals, complex mixtures, animal feeds, human foodstuffs and packaging.

2.1.3 Extent of Expected Use or Application and Impact on Human Health

The association of exposure to estrogenic substances and adverse health effects in human and wildlife populations has led to worldwide concern. Some of the health effects that have led to this concern include global increases in reproductive cancers, regional declines in sperm counts, altered sex ratios in wildlife populations, and accelerated puberty in females. The CCi test method is intended to be included in a battery of Tier 1 tests for the detection of potential ED activity.

2.1.4 Potential for the Proposed Test Method, Compared to Current Accepted Test Methods, to Refine, Reduce, or Replace Animal Use

There are no currently accepted test methods for screening potential ED substances. The CCi test method could allow rapid screening of potential ED substances toward more refined testing of substances in animal models. This could result in a significant reduction in animal use in Tier 1 and Tier 2 testing.

2.1.5 Potential for the Proposed Test Method to Provide Improved Prediction of Adverse Health Effects, Compared to Current Accepted Test Methods

There are no test methods approved for the detection of potential EDs by any regulatory agency. The CCi test method has potential for improving the prediction of adverse health effects when used in a battery of Tier 1 screening tests for the detection of potential ED activity.

2.1.6 Extent to Which the Proposed Test Method Provides Advantages (e.g., Reduced Cost and Time to Perform) Compared to Current Methods

The CCi test method uses a HTS automated system and can rapidly test the estrogenic activity of substances in a relatively short amount of time (days) and at low cost. The test method is sensitive and robust, capable of detecting substances with high estrogenic activity at less than picomolar concentrations and substances with low estrogenic activity at less than micromolar concentrations. Because the test method is transcriptionally based, it is potentially capable of screening for either estrogenic or anti-estrogenic activity.

2.2 Extent to Which the CertiChem Background Review Document Provides the Information Requested in the *ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods*

2.2.1 Introduction, Including the Scientific and Regulatory Rationale for the Proposed Test Method

The introduction contains scientific and regulatory rationales for the proposed test method. It states that there are no known limitations to the type of materials that can be tested, however, the testing of particulates and solids is limited to testing solvent-extractions. The BRD also discusses concerns regarding the testing of volatile substances.

2.2.2 Information on the Development of the Proposed Test Method Protocol and its Key Components

Information on the development of the proposed test method and its key components was addressed but a rationale for the selected number of replicate samples per study and the number of repeat experiments (indicated as ranging between 3 and 9) was not provided.

2.2.3 Characterization of the Substances Used for Validation Studies on the Proposed Test Method

The rationale for selection of test substances is based on their inclusion in the ICCVAM ED Reference Substances list. A rationale is also given for cytotoxic substances and antioxidants that were used in test method development. The information provided for substances

included Chemical Abstracts Service Registry Number (CASRN), chemical class, purity, supplier and concentrations tested.

2.2.4 *In Vivo* Reference Data Used to Assess the Accuracy and Reliability of the Proposed Test Method

This was not provided as there is no accepted animal or human data set to serve as a reference for determining the accuracy of *in vitro* test methods for identifying substances with estrogenic activity *in vivo.*

2.2.5 Test Method Data and Results

Test method protocols used to generate each submitted set of data are described, and an explanation of protocol differences is included. The statistical approach used to evaluate data and to indicate test method repeatability is also included. Tested substances were not coded.

2.2.6 Assessment of the Accuracy of the Proposed Test Method

An assessment of accuracy is conducted and possible test method limitations regarding the testing of volatiles is discussed. Measures of variability for historical positive and negative control data is limited.

2.2.7 Assessment of the Reliability (Repeatability/Reproducibility) of the Proposed Test Method

The selection rationale indicates to what extent the chosen set of substances represents the range of possible test outcomes. A statistical assessment of repeatability and reproducibility is provided.

2.2.8 Assessment of Test Method Data Quality

This section is complete. Studies were not Good Laboratory Practice (GLP) compliant and substances tested were not coded for identity.

2.2.9 Other Scientific Reports and Reviews Pertinent to the Proposed Test Method

This section is complete.

2.2.10 Assessment of Animal Welfare Considerations (Refinement, Reduction, and Replacement)

This section is complete.

2.2.11 Practical Considerations (e.g., Test Method Study Costs, Time Needed to Perform Study, Ease of Transferability of the Test Method Among Laboratories)

Training requirements needed for personnel to demonstrate proficiency with the test method are addressed and detailed cost and time considerations are provided.

2.2.12 A Comprehensive and Complete List of References

References are provided for materials ranging from the test method rationale to the summary of positive and negative control data.

2.2.13 Supporting Materials (Appendices)

This section is complete.

2.3 Extent to Which the Proposed Test Method Adheres to the Recommendations of the *ICCVAM Evaluation of In Vitro Test Methods for Detecting Potential Endocrine Disruptors* (NIH Pub. No. 03-4503), Especially Those Regarding Essential Test Method Components

2.3.1 Reference Estrogen and Associated TA Response

The ICCVAM recommended reference estrogen for agonism is 17β-estradiol (E2). The proposed test method uses E2 run as a full dose-response curve in each experiment.

2.3.2 Preparation of Test Substances and the Volume of Administered Solvent

ICCVAM recommends water, ethanol, or DMSO as solvent with the volume of solvent in the reaction mixture ranging from 0.1 to 1%. In the CCi test method, test substances are

dissolved in 100% EtOH, with an administered concentration of 0.5%, and volume and concentration of solvent is constant and identical for all wells.

2.3.3 Concentration Range of Test Substances that Should be Tested

ICCVAM recommends a limit dose of 1 mM in the absence of solubility or cytotoxicity constraints and the testing of seven concentrations of substance spaced at log intervals up to the limit dose should be tested. An evaluation of cell cytotoxicity is also recommended. The maximal test concentration used in the CCi test method is 0.5 mM, with range finder testing concentrations spaced at log intervals from 0.5 mM to 50 pM. Cytotoxicity is visually evaluated and concentrations of test substances that cause overt toxicity are not considered in the evaluation of the data.

2.3.4 Negative, Solvent and Positive Controls

ICCVAM recommends the use of concurrent solvent controls in each experiment. The CCi test method uses concurrent solvent controls in each experimental 96 well plate and volume and concentration of solvent is constant and identical for all wells. ICCVAM also suggests the use of a positive control estrogen with a maximal TA response two to three orders of magnitude lower than the reference estrogen. The CCi test method does not use a positive control other than the E2 reference estrogen.

2.3.5 Number of Within-Test Replicates

ICCVAM recommends that all concentration levels of controls, the reference estrogen, and test substance are to be run in triplicate. Within-test replicates are run in quadruplicate in the CCi test method.

2.3.6 Methods for Data Analysis

ICCVAM does not recommend a specific method for data analysis. The current proposed CCi test method protocol specifies the use of a four parameter Hill equation to calculate EC_{50} values. However, some data provided in the BRD used a Michaelis-Menton analysis to calculate EC_{50} values.

2.3.7 GLP Compliance

ICCVAM recommends that studies should be performed in compliance with GLP guidelines. The CCi testing facilities, test method, and protocol are not currently GLP compliant.

2.3.8 Study Acceptance Criteria

ICCVAM recommends that the response for the reference estrogen and controls should be within the appropriate historical acceptance range. The CCi test method protocol defines acceptance criteria as follows:

- EC_{50} of the positive control (E2) must be within 2.5 standard deviations (SD) of the historical mean established by the test laboratory and have an r^2 (coefficient of determination) value ≥ 0.9 calculated by four parameter Hill equation using GraphPad Prism.
- At least one data point on each EC_{50} plot of the positive control and the test chemical must be 10% - 50% of the maximum response and at least one data point must be 50% - 90% of the maximum response.
- At least two data points must be < 10% of the maximum response and these points constitute the bottom plateau of the concentration response curve.
- The standard deviation of all vehicle controls should not be more than 15% of the mean.
- The standard deviation of all negative controls should not be more than 15% of the mean.
- Any EC_{50} plot must be repeated using different dilutions of the appropriate positive control or test substance until all these conditions are met.

2.3.9 Interpretation of Results

Substances are classified as ER agonists when the response (i.e., cell proliferation) elicited by the substance is increased significantly above the concurrent solvent control level, as determined by a statistical test in the CCi test method. Specificity for ER agonism is confirmed by testing all substances that are positive for agonism against a reference antagonist (ICI 182,780).

2.3.10 Repeat Studies

Repeat studies have been conducted. Repeatability of the CCi test method was evaluated by comparing the correlation of variation for ranked EC_{50} values between data from ICCVAM and from the repeat studies conducted at CCi, NWU and UM.

2.3.11 Study Report

No examples of study reports were provided in the BRD

2.4 Extent to Which the Proposed Test Method Shows Adequate Performance (Reliability and Accuracy) During Pre-Validation to Warrant Consideration for Validation Studies

2.4.1 Performance Evaluation

2.4.1.1 *Reliability (Intralaboratory Repeatability and Intra- and Inter-Laboratory Reproducibility)*

The CCi BRD provided limited information on well-to-well and experiment-to-experiment repeatability[1] for the test method, and a discussion of intralaboratory reproducibility was not provided. Different versions of the CCi protocol were tested in three independent laboratories. None of the three laboratories involved in the testing used coded test substances.

- *Intralaboratory Repeatability* – A specific discussion regarding the closeness of agreement between test results obtained within a single laboratory when the procedure is performed on the same substance under identical conditions within a given time period was not provided. The BRD references studies that were conducted at CCi, NWU, and UM to support overall test method reliability. Information was provided for 10 substances that were tested repetitively. Nine of the substances are on the ICCVAM recommended validation substances list. The antioxidant, propyl gallate, which is not on the

[1] In the CCi test method, a substance is tested at up to 11 concentrations, with each concentration tested in quadruplicate wells, on separate 96-well plates, with one well per plate. There is no information on how sample variability across plates would compare to well variability for replicates run within a single plate.

ICCVAM list, was included to provide a known ER-negative substance for testing. Data for the substances that were positive for ER activity are presented as the mean and standard error of the mean (SEM) of calculated EC_{50} values. These EC_{50} values were ranked for relative agonist activity and then evaluated by comparing the correlation of variation for the ranked EC_{50} values between data from ICCVAM and from the repeat studies conducted at CCi. A comparison of these ranked values was then made for values between CCi and NWU, and CCi and UM. The analysis indicated that the relative agonist activities are not significantly different for any of the comparisons and the submission uses these results to support the overall reliability of the test method. A statistical analyses specifically addressing well-to-well repeatability was not provided. Graphical and tabular presentations of data are used to support experiment-to-experiment reproducibility but a statistical analysis specifically addressing this is not provided.

- *Test Method Intralaboratory Reproducibility* – Although a review of data tables provided as supporting materials for the BRD indicated that experiments were conducted by different individuals, a determination of whether qualified experimenters within the same laboratory can successfully replicate results using the test method protocol was not discussed nor specifically analyzed.
- *Test Method Interlaboratory Reproducibility* – The studies conducted at CCi, NWU and UM used different versions of the basic test method protocol. Therefore, an assessment of interlaboratory reproducibility is somewhat qualified. However, the analysis of ranked EC_{50} values indicating that the relative agonist activities are not significantly different suggest that adequate interlaboratory reproducibility is demonstrated. An overview of test method protocol differences is provided in **Table 1**.

2.4.2 The Accuracy of the Proposed Test Method for Detecting Agonist Activity

There is no accepted animal or human data set to serve as a reference for determining the accuracy of *in vitro* test methods for identifying substances with estrogenic activity *in vivo*.

As an alternative, CCi evaluated relative agonist activity by comparing ranked EC_{50} values for 18 substances on the ICCVAM recommended reference substances list with their ranked EC_{50} values for these same substances using a Least Squares Regression analysis. The

Table 1 Differences in Test Method Protocols Between Laboratories

Protocol Properties	Northwestern University	University of Missouri	Certichem
Format	Manual	Robotic	Robotic
Seeding Density (96 well plate)	500 cells/well	2200 cells/well	300 cells/well
Frequency of Media Change	Every other day	Daily	Daily
Length of Exposure to Substance	7 Days	4 Days	7 Days
Robot Model	Not Applicable	Tomtec Quadra 96 Robotic Workstation	epMotion 5070 Robotic Workstation
DNA Quantification	CyQuant Proliferation Assay Kit (Invitrogen)	Burton Diphenylamine Assay	Burton Diphenylamine Assay
Output of DNA Quantification Method	Fluorescence	Absorbance	Absorbance
EC_{50} Calculation	Hill Equation	Michaelis-Menton	Hill Equation

analysis indicated that the relative agonist activities are not significantly different. CCi also tested 34 substances for which no estrogenic activity has been reported and for which QSAR analyses predict as negative. CCi testing also indicated that these substances are negative.

2.4.3 Evaluation of Concordance

CCi tested 40 substances on the ICCVAM recommended reference substances list and compared ICCVAM published ER activities for concordance with their experimental results. ICCVAM had classified 29 of these substances as positive (or presumed positive) and 11 as negative (or presumed negative) for *in vitro* ER TA activity. The comparative results are as follows:

- CCi positive and ICCVAM Positive 27 substances
- CCi negative and ICCVAM Positive 2 substances
- CCi negative and ICCVAM Negative 10 substances
- CCi positive and ICCVAM Negative 1 substances

The CCi test method correctly identified as positive 27 out of the 29 substances indicated as positive by ICCVAM. Of the two substances that were indicated as false negatives in the CCi test method, one (4-hydroxytamoxifen) is a well-known antagonist that is positive in some ER agonist tests and the other (2,4,5-trichlorophenoxyacetic acid) is a substance for which the ICCVAM classification is based on data from a single study. The one false positive substance was mifepristone, which had been classified by ICCVAM as a presumed negative for ER agonism. Updated information from other laboratories has indicated that mifepristone is actually positive for ER agonism, suggesting that the CCi result was correct. **Table 2** provides the 2 x 2 table used to conduct an analysis for concordance, sensitivity, specificity, positive and negative predictivity, and false negative and false positive rates.

Table 2 Analysis of Accuracy Between ICCVAM Data and Certichem Results

		ICCVAM Classification		
		Positive	Negative	Total
CCi Classification	Positive	27	1	28
	Negative	2	10	12
	Total	29	11	40

Concordance = 93%
Sensitivity = 96%
Specificity = 83%
Positive Predictivity = 93%
False Negative Rate = 4%
False Positive Rate = 17%
Negative Predictivity = 91%

3.0 SUMMARY

The CCi BRD adequately addresses the ICCVAM prioritization criteria and the ICCVAM submission guidelines. The proposed test method protocol is consistent with the recommendations for standardization and validation of *in vitro* endocrine disruptor test methods. The CCi BRD also adequately addresses the performance of the CCi test method.