

VI. Report on Peer Review Panel Meeting: Evaluation of an *In Vitro* Estrogen Receptor (ER) Transcriptional Activation (TA) Test Method for Endocrine Disruptor Chemical (EDC) Screening

Introduction and Overview of Proposed Methods and Applications: The BG1Luc ER TA (LUMI-CELL[®]) Test Method to Identify Substances with Estrogen Agonist and/or Antagonist Activity

NICEATM Deputy Director Dr. Warren Casey briefed SACATM on the proposed endocrine disruptor test method. The EPA has been mandated to develop a screening program to detect EDCs so it asked ICCVAM to evaluate existing validated *in vitro* EDC screening tests. ICCVAM found none, leading ICCVAM and SACATM to make validation of such a test a high priority. In response, there was a nomination from Xenobiotic Detection Systems (XDS) for its LUMI-CELL[®] assay, a luciferase reporter assay that detects estrogen-binding activity. The assay is based in human ovarian carcinoma (BG-1) cells, with endogenous ER-alpha and ER-beta. The test provides a concentration-response, and so can assess both potency and efficacy. There are nearly identical protocols for both agonists and antagonists. The agonist assay involves gain of function, while the antagonist test measures loss of function, both based upon luciferase levels.

Dr. Casey provided a timeline for the project, beginning in January 2004 with the nomination of the assay by XDS, through the public peer review meeting in Bethesda in March 2011. He reviewed the definition of validation and ICCVAM's validation criteria, as well as the four phases of the international validation study, which was sponsored by NICEATM-ICCVAM, JaCVAM, and ECVAM.

When the testing was completed, accuracy and reproducibility were assessed. The agonist test method was 97% accurate, had 96% sensitivity and 100% specificity. The antagonist method was 100% accurate, with 100% sensitivity and 100% specificity. The agonist method showed 100% intra-laboratory reproducibility of the substances tested independently three times. Inter-

Summary Minutes from the June 16-17, 2011 SACATM Meeting Arlington Hilton, Arlington, VA

laboratory reproducibility was 81%. For the antagonist methods, intra-laboratory reproducibility was 100%, while inter-laboratory reproducibility was 89%. In a comparison of the BG1Luc ER TA with the ER binding assay, there was 97% concordance. Compared with the Chemical Evaluation and Research Institute (CERI) Stably Transfected Human Estrogen Receptor Transcriptional Activation (STTA) assay, overall there was 86% concordance using 26 reference substances. Based on the validation program, ICCVAM recommended the use of the BG1Luc ER TA as a screening test to identify substances with estrogen agonist and antagonist activity, with the highest test substance concentration limited to 10 µM for the antagonist assay. ICCVAM also developed and released performance standards for the assays.

ICCVAM conducted a peer review panel meeting March 29-30, 2011, to consider the recommendations, performance standards, and background data. The panel consisted of 16 scientists from 6 countries. Following the SACATM meeting, the Endocrine Disruptor Working Group will consider SACATM comments and the panel report and finalize ICCVAM's test method evaluation report. Ultimately, in fall 2011, the ICCVAM recommendations will be forwarded to Federal agencies, and a draft test guideline will be forwarded to OECD.

Summary of the Independent Scientific Peer Review Panel Evaluation of the Validation Status of the LUMI-CELL ER® (BG1Luc ER TA) Test Method

Dr. John Vandenberg of North Carolina State University (retired), who chaired the Peer Review Panel ("the Panel"), briefed SACATM on the meeting.

He reviewed ICCVAM's charges to the Panel and its recommendations. The Panel agreed with ICCVAM that the BG1Luc ER TA could be used as a screening tool to identify substances with *in vitro* estrogen agonist and antagonist activity. It considered the test method protocol to be complete and adequate in detail, and agreed with ICCVAM about the needs for future studies. The Panel also suggested that such future studies could address metabolic activation, that the reference substance list and associated database could be expanded with additional negative agonist and positive antagonist substances as they are identified, and that efforts could be made to identify a quantitative cytotoxic method. It also concurred with the draft ICCVAM performance standards and some modifications to expand applicability of the performance standard.

Public Comments

Dr. Niemi called for public comments and noted written comments had been submitted from CertiChem, Inc.

Dr. Catherine Willett, Associate Director of Regulatory Testing for People for the Ethical Treatment of Animals (PETA), reported that PETA lauded the Panel and supported the recommendations, both the main finding recommending the test method and the other recommendations. She congratulated the Panel on its review, saying it was "an incredibly thorough, well-done, well-reviewed validation study." She listed several panel recommendations that PETA supported: (1) designation of the assay as an alternative for the CERI STTA assay and the rat uterine cytosol assay, (2) development and validation of ER

Summary Minutes from the June 16-17, 2011 SACATM Meeting Arlington Hilton, Arlington, VA

binding assays using recombinant receptors for both humans and other animals, (3) development and use of a metabolism component, (4) inclusion of potency evaluations to quantify activity, (5) evaluation of the quality of the data used to classify the original ICCVAM reference substances (6) discussion of the use of assay, and (7) discussion of the animal reduction potential.

She conveyed several additional PETA recommendations: (1) revise the chemical list to follow up on the evaluation and updating of the chemical reference list, and adding the new information to a publicly searchable database; (2) ensure that the best characterized chemicals are used for future assay evaluations; (3) identify new reference chemicals in underrepresented chemical classes; (4) consider the use of the assay to reduce animal testing, such as its use in addition to screening and prioritization, revising the structure of the Endocrine Disruptor Screening Program (EDSP) Tier 1 assessment by performing *in vitro* assays prior to animal testing, and the adoption of a weight-of-evidence approach that could be used to further reduce or eliminate estrogen receptor-related animal tests; and (5) evaluate the data quantitatively using a Relative Potency Index relative to a standard reference chemical, to allow quantitative comparison to the CERI STTA and to other assays

She noted the study had taken 7 years to complete, and so was not included in Phase I of the EDSP. She said a more efficient process is needed in light of the large number of new assays emerging. She recommended the Panel note issues that contributed to the length of the review in its report, and include recommendations for avoiding those issues in future reviews.

Dr. Niemi recognized Dr. Fowle, who was at that point prepared to respond to Dr. Hansen's request regarding data on adoption of alternative test methods.

Dr. Fowle said data had last been collected August 26, 2010, regarding 12 assays, which were grouped from the larger assay population: LLNA: 241, Corrositex: 0, Up and Down Assay: 1,139, EpiSkin/EpiDerm: 2, BCOP: 14, ICE: 0, *In Vitro* pyrogen tests: 0, Cytosensor: 0, EpiOcular: 3, LumiCell: 0, CertiChem: 0, Total: 1,399.

He mentioned that those figures may make it appear that EPA and others are not committed to reducing, refining and replacing animal use, and asked that he be allowed to comment at some point about some of the things EPA is doing to achieve the 3Rs. Dr. Niemi asked Dr. Fowle to hold those comments for later in the meeting.

SACATM Discussion

Dr. Corcoran, lead discussant, said the EDC method evaluation seemed to be "a tour de force," and commended the work of the Panel. He said he would like more information about the quality of data issue that had been commented upon in the Panel's report, specifically the criteria involving ranking and sensitivity analysis, or tests for trends in terms of the criteria for evaluating positive and negative compounds. He asked Dr. Vandenberg to comment on whether the Panel was proposing a higher and new standard for all assays of this nature. Dr. Vandenberg said it would be presumptuous for the Panel to do so, in terms of attempting to direct what other panels might do. On the other hand, he said, it would be fine for other panels

Summary Minutes from the June 16-17, 2011 SACATM Meeting
Arlington Hilton, Arlington, VA

to adopt the standards described by this one. Dr. Corcoran asked for clarification on the Panel's conclusion that there were insufficient data to term the evaluation a "thorough" analysis, although it was termed as "adequate." This was, he said, due to the use of a descriptive versus a formal, inferential assessment of the data. Dr. Vandenberg said it was hard for the Panel to consider the analysis to be thorough, since there would always be things that had not been thought of. Thus, their description of the analysis was adequate. Statistically, he said the analysis of the data was considered to be adequate, with no fault found. Dr. Casey added it was always difficult to get statisticians to agree on anything, so some of the comments pointed to ways things could have been done differently, statistically, particularly EC₅₀ calculations. Dr. Corcoran said he had been hoping to hear that ICCVAM was moving toward a new standard for quality of data.

Dr. Corcoran added he would like to have seen more information in the document on the implications of the assay for use in Europe and Japan. Dr. Vandenberg said that was not specifically discussed as it related to the background document, but it did come up during the discussion, and there were foreign representatives present who brought some of those issues. Dr. Corcoran said he would like to have seen validation conducted in one set of known agonists and antagonists, and then movement into a second set of yet-untested agonists and antagonists, thus incorporating a two-step process. He recognized it had already been a 7-year, \$3 million process, but nonetheless objected to validation based on only one set of compounds. Dr. Casey said every positive and every negative they could find had been tested, but the chemical space was very small for well-referenced compounds; just 38 compounds fit the criteria. Dr. Corcoran maintained since the protocol was changed over the course of the 7 years, having two sets of data would have helped, even if it involved splitting up the known compounds. Despite his comments, Dr. Corcoran said the review was "a very impressive body of work."

Dr. Elmore, lead discussant, agreed with the previous comments, as well as the conclusions and recommendations contained in the report. He felt, however, the BG1 cell line needs to be better characterized. He recommended the cell line be placed in a repository to ensure access and availability in the future.

Dr. Meyer, lead discussant, was also impressed with the work of the Panel, calling it "very comprehensive and very clear." She strongly supported the idea that cytotoxic changes be quantified. She noted that although validation normally means the replacement of an *in vivo* method with an *in vitro* method, in this case, an *in vitro* method is to be replaced by another *in vitro* method. She questioned the priority of whether ICCVAM should be funding such an effort, given the large number of animals still being used in other areas. Dr. Meyer noted the introduction of the non-radioactive LLNAs would actually replace animal use, but that the EDC assay is a screening method, and that she was uncomfortable with expending too many resources on such an approach. She wondered whether the current method could not be further developed to work on antagonists. She also asked about harmonization for *in vitro* methods. She mentioned it would be helpful to have a formula in the document on how the fold-reduction was calculated and commented on a lack of clarity for expressing the performance standard.

Summary Minutes from the June 16-17, 2011 SACATM Meeting
Arlington Hilton, Arlington, VA

Regarding priority, Dr. Stokes said the developer nominated the method for validation studies in 2005, and at that time it was given a very high priority by SACATM. He said in the case of a positive, that such information could be used along with other mechanistic data to move forward with characterizing whether or not the compound is in fact an *in vivo* endocrine disruptor. Dr. Stokes said regarding the comparison with the current method that has been adopted by the EPA and is in their guidelines, this was done because the adoption had occurred after the validation study was initiated. He said, "But we didn't even know about that method in 2005, that it even existed, but it was moving along and a couple of years later, yes, we did find out that it was going through validation as well. This study was nominated and a validation study was initiated before there was any knowledge of the other method."

Dr. Fowle said in terms of maximizing the utility of tests, clearly things have evolved, and some of the earlier screens that were developed for validation occurred a number of years ago. He said it's really important, if these screens get used, that they get linked very closely in terms of working with the regulatory agencies and the users who'll be using them, to make sure these assays will be used, and will be used for purposes which will help advance the mission. He said Dr. Meyer raised some very good points in terms of the resources available. ICCVAM focuses on validation of alternative methods to animal tests, and he thinks it's very important to focus on replacements for animal tests. EPA's policy and approach for using the EDSP is such that it probably will not be using this assay. He said he thought it just sort of underlines the importance of having very close communications at the beginning, middle, and end. He alluded to the history of EPA discussions with Drs. Stokes, Bucher, and Birnbaum as they tried to build on the lessons learned to try to do a better job in the future. He suggested having a retreat or similar meeting to look at the good things ICCVAM has done, see what might be improved, and figure out how to move forward. Dr. Birnbaum agreed with Dr. Fowle, but reminded everyone that the purpose of some *in vitro* tests is to answer a very specific question. She said this test determines whether a substance is an agonist or antagonist for ER α and ER β , but there are other ways that chemicals can be endocrine disruptors, e.g., of the estrogen signaling system, and this test is not identifying them.

Dr. Wilson, lead discussant, concurred with previous comments, as well as the need for a follow-up meeting with ICCVAM to focus on trying to determine an overview of the various assays currently in use. He noted to run an assay is as much an art as a science, and that it should be moved more toward the science. So a focused discussion with experts to understand the limitations of the current assays and see whether any stand out would be helpful to further the state of the science. For the EDC assay, he agreed with Dr. Elmore regarding better characterization of the cell line. He cautioned that use of the phrase "endocrine disruptor" carries an obvious stigma, and suggested a careful definition of what is or is not an endocrine disruptor be put into the background information of the document.

Dr. Casey noted the figures regarding accuracy, reliability, and reproducibility had been approached in a thoughtful manner, with choices having been made among potential approaches. Dr. Meyer suggested revising the specific section she had earlier referred to as problematic. Regarding usability of the assay in high throughput screening, Dr. Casey said it was currently being evaluated at NCGC, and that it works well in a 384-well format, but it may

Summary Minutes from the June 16-17, 2011 SACATM Meeting
Arlington Hilton, Arlington, VA

not provide adequate signal-to-noise in a 1536-well format. He continued by noting the cells for the assay are co-owned by XDS and Dr. Michael Denison at the University of California, Davis, and that Dr. Denison was reluctant to put the cells into a repository because he wishes to maintain control of them. He does make them freely available to academic and government labs through a formal licensing process. Dr. Vandenberg said the Panel had discussed the issue of cell line availability at length, and did all they could to ensure access to the cell line. Dr. Toth expressed concern about drift in the cell line over time, asking whether there are quality control measures to ensure such drift would not take place. Dr. Casey said positive and negative controls are run with each test, but that currently there is not a way to track the genetic stability of the line. Dr. Stokes said all of the *in vitro* assays use acceptance criteria for the positive controls, so there must be a response within that acceptance range. Thus, if the cells have changed and the response has been decreased to below that threshold for an acceptable positive control response, or if it exceeds the upper limit of it, it would not be a good run and it would indicate that perhaps the cells had changed, become contaminated, or were the wrong cells.

Regarding the history of the assay, Dr. Stokes noted the EDSP was mandated by laws in 1996. The *LUMI-CELL ER*[®] was developed in response to a Small Business Innovative Research (SBIR) topic issued by NIEHS in the late 1990s in response to considerable interest at the time. The SBIR grant to develop the EDC method was supported by NIEHS and NIH grant funds. Dr. Birnbaum added that since NIH supported the development of cell lines, they should be fully available. Relevant to agencies' involvement, Dr. Stokes said there is an Endocrine Disruptor Working Group that includes representatives from all of the ICCVAM agencies. Dr. Stokes said the working group had EPA representatives on it who were kept abreast of the study design, chemical selection, and protocols, which were all run by that group before this testing went forward. He clarified that all of the agencies in ICCVAM had the opportunity for input into this validation study. New members have been integrated into ICCVAM and SACATM, and the work of the previous members may have been forgotten. He said NICEATM-ICCVAM is trying to make sure as much information as possible is reflected in the final evaluation reports that go out to the agencies and to the public.