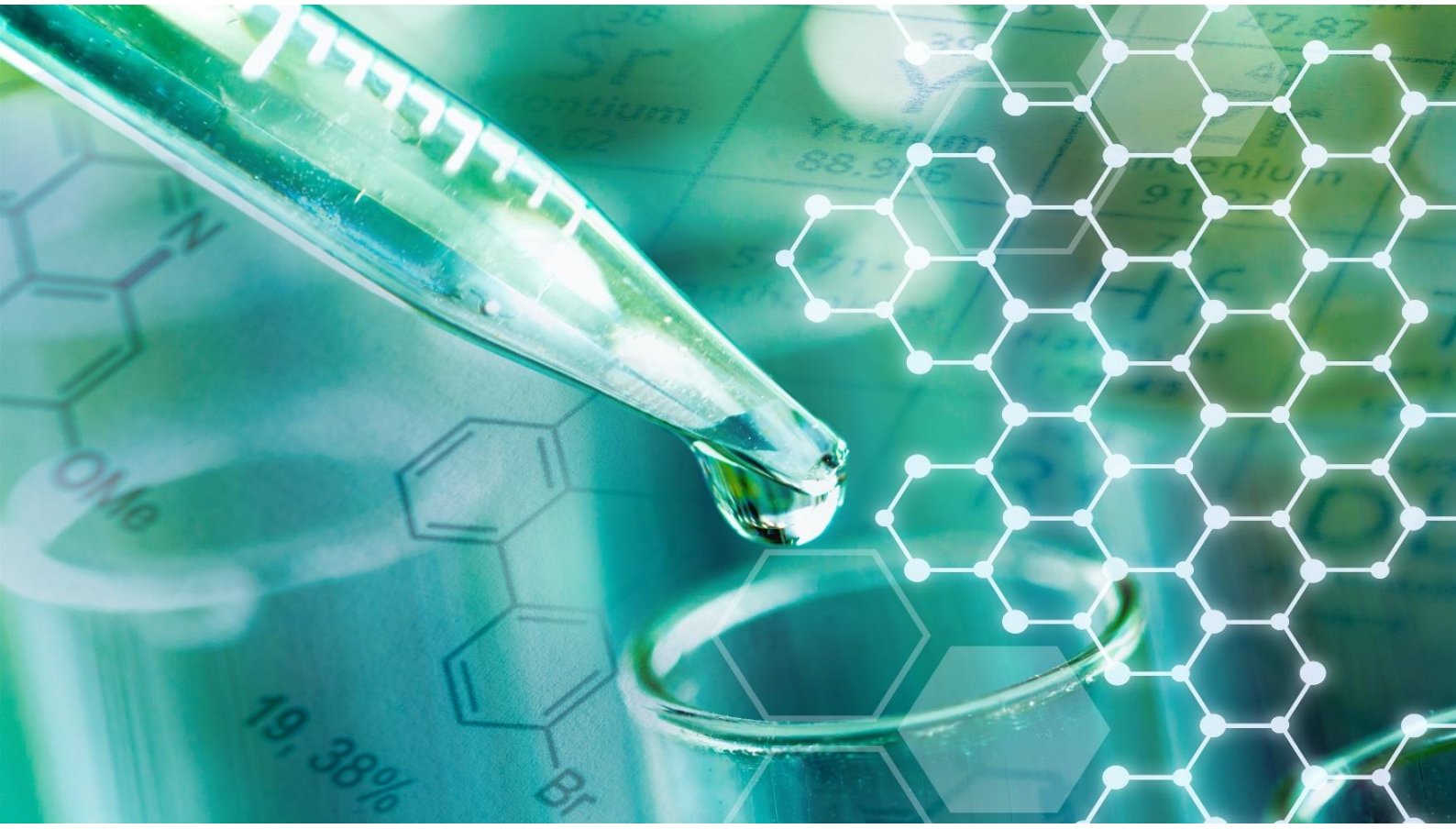
Peer Review Report of the Validation of the IL-2 Luc Leukocyte Toxicity Test (IL-2 Luc LTT) Assay on In Vitro Immunotoxicity

Series on Testing and Assessment No. 392



OECD

Series on Testing and Assessment

No. 392

Peer-Review Report of the Validation of the IL-2 Luc Leukocyte Toxicity Test (IL-2 Luc LTT) Assay on In Vitro Immunotoxicity

2024

Please cite this publication as:

OECD (2024), *Peer-Review Report of the Validation of the IL-2 Luc Leukocyte Toxicity Test (IL-2 Luc LTT) Assay on In Vitro Immunotoxicity*, OECD Series on Testing and Assessment, No. 392 (if applicable), OECD Publishing, Paris.

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 38 countries in North and South America, Europe and the Asia and Pacific region, as well as the European Union, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several Partner countries and from interested international organisations attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in twelve different series: **Testing and Assessment**; **Good Laboratory Practice and Compliance Monitoring**; **Pesticides**; **Biocides**; **Risk Management**; **Harmonisation of Regulatory Oversight in Biotechnology**; **Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; Safety of Manufactured Nanomaterials;** and **Adverse Outcome Pathways.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (https://www.oecd.org/en/topics/chemical-safety-and-biosafety.html).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organizations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank, Basel, Rotterdam and Stockholm Conventions and OECD. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

IL-2 Luc LTT assay validation - Report of the Peer Review Panel

A JaCVAM coordinated study program addressing the validation status of the IL-2 Luc LTT for prospective identification of immunotoxic chemicals.

Report completed by the Peer Review Panel on May 2, 2023.

Foreword

This document contains the report of the peer-review organized by Japan for the validation of the IL-2 Luc LTT assay on in vitro immunotoxicity, a project on the work plan of the Test Guidelines Programme.

When the project proposal was discussed in 2022, the WNT recommended to take a step-wise approach and consult the WNT after the validation report and peer-review reports are reviewed by the Expert Group to decide on the pertinence of developing a Test Guideline.

The validation report and the peer-review report were circulated in September and again in December 2023 for review. Feedback received was minimal (Canada, the Netherlands, Germany and the United States commented). The Working Party of the National Coordinators of the Test Guidelines Programme endorsed the conclusions of the peer-review report and recommended to proceed with the development of a draft Test Guideline for the IL-2 LTT assay as a non-standalone assay for in vitro immunotoxicity.

The peer-review report was endorsed by the Working Party of the National Coordinators of the Test Guidelines Programme (WNT) at their 36th meeting in April 2024. This document is published under the responsibility of the Chemicals and Biotechnology Committee.

Executive Summary:

The interleukin-2 (IL-2) luciferase lymphocyte toxicity test (IL-2 Luc LTT) assay has been proposed as an in vitro alternative to animal testing. The assay is related to the recently reviewed IL-2 luciferase assay (IL-2 Luc). While the IL-2 Luc assay provides information on adverse outcome pathways for immunotoxicity to T cells, the IL-2 Luc LTT identifies those T cell targeting agents that suppress IL-2 through an antimitotic mechanism, which was a gap identified with the original IL-2 Luc assay.

The Peer Review Panel (PRP) found the Validation Management Team's report presented the necessary information for an independent review.

The PRP concluded that the IL-2 Luc LTT assay was well defined and has a clear protocol and criteria for data interpretation. All necessary information, including performance standards, was sufficiently detailed. Both within- and between-laboratory reproducibility were satisfactory. The PRP noted that it is important to clearly delineate the IL-2 Luc LTT from the IL-2 Luc assay. While the predictive capacity was not satisfactory for a stand-alone method, the IL-2 Luc LTT assay is acceptable for use in an Integrated Approach to Testing and Assessment. In fact, the report included data showing that the combination of the IL-2 Luc with the IL-2 Luc LTT increased predictive capacity.

Peer Review Panel Composition:

Barbara Kaplan, chair	Mississippi State University, USA
Chiyomi Kubo	Chugai Pharmaceuticals, Japan
Henk van Loveren	Maastricht University, The Netherlands
Takayuki Yoshimoto	Tokyo Medical University, Japan
Haley Neff-LaFord	Seagen Inc., USA
Support from Takao Ashikaga	JaCVAM, Japan

Background:

A healthy immune system must be in balance such that it can identify and eradicate pathogens when needed, but not react to self-proteins. Thus, an imbalance of the immune response in either direction can produce morbidity and mortality; an insufficient immune response can cause infections or cancers, while an overactive immune response can contribute to hypersensitivity or autoimmune diseases.

Exposure to drugs or chemicals can alter the immune response in either direction. Therefore, there exists a need to develop and validate assays to detect immunotoxic compounds, especially those that produce immune suppression. While an immune response is a coordinated effort among several different cell types, the T cell is a critical initiator of adaptive immunity that can provide help to other immune cells by producing cytokines or mediate killing of infected cells. An important cytokine produced by T cells is interleukin-2 (IL-2), which promotes proliferation of T cells and helps maintain specific T cell subpopulations that contribute to immune homeostasis.

The IL-2 Luc assay was developed as a screening method to identify chemicals that target T cells. In fact, it was the first validated in vitro assay to identify immunosuppressive chemicals. Interestingly, the IL-2 Luc assay did not detect well-characterized immunosuppressive drugs that acted by inhibiting proliferation of T cells. Thus, the IL-2 Luc LTT assay was proposed to identify those T cell targeting chemicals that acted through an anti-mitotic mechanism. Indeed, it was demonstrated that the IL-2 Luc LTT assay accurately identified immunosuppressive chemicals (mostly pharmaceutical drugs) that were not identified with the IL-2 Luc assay alone. These results reinforce the idea that immunotoxic chemicals will not likely be identified with a single assay and that assays to identify immunotoxic chemicals will be part of integrated approaches to testing and assessment (IATA).

The purpose of this document is to provide a constructive review of the IL-2 Luc LTT. The PRP initially met virtually in October 2022 and then in person in Shizuoka, Japan in November 2022 to continue to evaluation. Subsequent virtual meetings occurred in January, February, March, and May 2023 to prepare this peer review report.

Evaluation criteria and peer review panel comments to each (provided in blue text) are detailed below:

Evaluation Criterion 1: A rationale for the test method should be available, including a description of the human health effect, a clear statement of scientific need, and regulatory application.

The human health effect problem statement, test method, rationale for scientific need and intent for regulatory application are stated in the document, although it is not necessarily found in a single section (and not all sections are at the beginning of the document). Perhaps a single section summarizing the rationale could be included somewhere near the beginning that includes these points as part of the rationale: 1. Environmental contaminants, food additives, and drugs can affect the immune system, resulting in immune dysregulation; 2. Immune dysregulation can have serious adverse health consequences, ranging from reduced resistance to infection and neoplasia to allergic and autoimmune conditions; 3. In vitro methods are needed due to the high cost, ethical concerns, and questionable relevance to risk assessment for humans using animal cells and model systems; 4. The fact that antimitotic effects of chemicals could not be detected by the IL-2 Luc assay; and 5. That the regulatory application is eventual generation of an OECD test method for immunotoxicity. There was also concern that there needs to be more background description on the meaning of measuring IL-2 for determining the immune suppression and antimitotic effect would be appreciated for reader to better understand the assays as follows (maybe in the first paragraph of the section 3-7). IL-2 is one good marker for immune suppression and cell proliferation. IL-2 plays a critical role in the proper activation of T cells and the lack of IL-2 production is well known to induce T cell anergy, leading to immune tolerance. In addition, IL-2 is also a potent growth factor for T cells and inhibition of IL-2 reduces T cell proliferation. Finally, there was some concern that the human relevance was not entirely clear, especially regarding the advantages of distinguishing between antimitotic and immunosuppressive effects.

Evaluation Criterion 2: The toxicological mechanisms and the relationship between the test method endpoint(s) with the biological effect as well as the toxicity of interest should be addressed, describing limitations of the test method.

The rationale for developing the test method and the toxicity of interest were clearly stated. The limitations of the test method were also laid out with discussion of the applicability domain and the combination of the IL-2 Luc LTT with the original IL-2 Luc assay to strengthen predictivity.

Evaluation Criterion 3: A detailed test method protocol should be available.

The detailed test method protocol was provided. There were 2 points that were inconsistent: there was a difference between total amount and sum of required amount in Table 6 of 8-2-2 (page 56), and the centrifuge speed for thawing 2H4 cells was wide range (120~350xg) in 8-2-4 (page 58). These points were properly corrected in the final version of the validation report.

Evaluation Criterion 4: The within and between laboratory reproducibility of the test method should be demonstrated.

The initial pre-validation trial (phase 0) was conducted with 3 non-coded chemicals (bleomycin sulfate, dexamethasone and 6-thioguanine). Within- and between-laboratory reproducibility of this assay (phase I) was checked using five coded chemicals in three test facilities (mycophenolic acid, indomethacin, cyclosporin A 5-FU, and mannitol). Between-laboratory reproducibility (phase II) was determined using 20 coded chemicals in three test facilities. Importantly, there were different chemicals tested in each of the 3 phases (except for use of a negative control chemical). Reproducibility was demonstrated. Within-laboratory reproducibility in the Phase I trial in Lab A, Lab B, and Lab C demonstrated 100%, 100%, and 100% reproducibility, respectively. Between-laboratory reproducibility result for Lab A, Lab B, and Lab C was 92.0 % for the combined data of the Phase I and Phase II trials. These results satisfied the acceptance criteria for the validation study of a within-laboratory reproducibility of at least 80% and a between-laboratory reproducibility of at least 80%.

Evaluation Criterion 5: Demonstration of the test method's performance should be based on testing of representative, preferably coded reference chemicals.

There was a Chemical Selection Committee (CSC) that determined the chemicals that should be used in all three phases, including use of positive and negative control chemicals. Chemicals were selected based in part on those chemicals determined by the literature (i.e., Luster et al, 1998) and other sources (e.g., National Toxicology Program) to be immunotoxic. Moreover, there was purposeful selection of immunotoxic chemicals that are known to act through an antimitotic mechanism as well as others that do not depend on antimitotic action to exhibit immunotoxicity. All chemicals were coded.

Evaluation Criterion 6: Predictive capacity should be demonstrated using representative chemicals.

Predictivity was provided for several endpoints. First, predictivity was provided based on the IL-2 Luc LTT for phases I and II, considering whether the known mechanism of the chemical had been determined to involve an antimitotic mechanism. Predictivity for the chemicals evaluated in phases I and II for the IL-2 Luc LTT was 76%. Second, predictivity was provided considering use of the IL-2 Luc LTT in combination with the IL-2 Luc, which was also 76% for the chemicals evaluated in phases I and II. Third, the predictivity was provided based on assessment of 85 test compounds (46 pharmaceutical drugs and 39 non-pharmaceutical chemicals). Predictivity for IL-2 Luc LTT for antimitotic effects was 80.8%. When the applicability domain was applied, the predictivity was slightly reduced to 79.5%. Finally, when combined with the IL-2 Luc, the predictivity was 77.2%. Predictivity is also provided for the pharmaceutical drugs for the IL-2 Luc compared to the IL-2 Luc in combination with the IL-2 Luc LTT and for the non-pharmaceutical chemicals for the IL-2 Luc compared to the IL-2 Luc in combination with the IL-2 Luc LTT (i.e., the 85 test compounds were separated into drugs versus chemicals and the predictivity calculated). Overall, there is extensive information provided on the predictivity of the IL-2 Luc LTT and how that might change when used in combination with the IL-2 Luc.

Evaluation Criterion 7: All data should adequately support the assessment of the validity of the test method for peer review.

Data were appropriately included and described in a way that allowed for peer review.

Evaluation Criterion 8: All data from the validation study supporting the validity of a test method should be obtained in accordance with the principles of Good Laboratory Practice (GLP).

Assays and quality assurance were carried out in the spirit of GLP, as not all the participating laboratories routinely worked under GLP certification. Overall, the principles of GLP were followed for this validation.

Evaluation Criterion 9: Applicability domain of the test method should be defined.

Applicability domain was described clearly in section 10-5, although it is discussed before this section so perhaps this could be moved further up in the document. Overall, the PRP agrees with the applicability domain as defined: 1.) chemicals that interfere with luciferase or luminescence that

confound its activity/measurement are out of the applicability domain; 2) the use of PMA/Io as a stimulant bypasses signaling through the T cell receptor and the subsequent intracellular signaling events that precede activation of phospholipase C, and therefore precludes detection of chemicals that act on those upstream signaling molecules; and 3) the Jurkat T cell line (from which 2H4 cells are derived) might lack several key proteins involved in the activation of normal T cells in response to TCR stimulation, and therefore may not be able to detect effects of chemicals that act on those key proteins.

Evaluation Criterion 10: Proficiency chemicals should be set up in the proposed protocol.

Bleomycin sulfate and cyclosporin are used as positive and negative controls for antimitotic effects, respectively. There were used to demonstrate proficiency across laboratories. NOTE: page 36 states that dexamethasone was negative control for antimitotic effects, but all other tables and verbiage state cyclosporin; needs to be edited.

Evaluation Criterion 11: Performance standards should be set up with the proposed protocol.

Various criteria to establish the success of the assay were stated several times, although they were not called “performance standards”. Moreover, some criteria or performance standards were not clear (i.e., a concrete attainment goal of the predictivity was not mentioned). Specific success criteria are not mentioned because this assay is not a stand-alone method but should be combined with IL-2 Luc assay. The predictivity was improved when IL-2 Luc LTT was used in combination with the IL-2 Luc assay plus the applicability domain. For instance the predictivity was highest when examining the combined results from the IL-2 Luc and IL-2 Luc LTT for those drugs or chemicals within the applicability domain.

Evaluation Criterion 12: Advantages in terms of time, cost, and animal welfare.

These advantages were described, although the cost advantage wasn’t completely evident. This assay was considered to be suitable for addressing 3Rs of animal research. However, the advantage was not as high when compared to similar in vitro assays. It might be better to compare the cost with other in vitro assays after the cost was calculated including human resources, availability of the cells, time for the cell maintenance, and number of samples that can be conducted in one experiment.

Evaluation Criterion 13: Completeness of all data and documents supporting the assessment of the validity of the test method.

Quality assurance issues were clear, and 3 areas were identified (temperature of luciferase activity measurement; time of P/I stimulation not consistent in one lab; and that positive and negative chemical controls need to be done at the same time as experiment). A detailed summary of the data was provided. Although we did not have access to raw data or paperwork, there were extensive analyses provided.

Evaluation Criterion 14: Validation Study Management (VSM) and Conduct.

The activity of the VSM and conduct of the study was described. A timeline for the assay development was also provided. A list of the various participants in the study was also included.

Other considerations

There is still some concern that there will continue to be confusion with IL-2 Luc versus IL-2 Luc LTT as it is currently presented. Even in the introduction and as part of the rationale, there are data shown from the IL-2 Luc, which might be confusing if people are not reading the document carefully. Perhaps consider presenting only the IL-2 Luc LTT and associated data first then dedicate a separate section at the end to an "IATA effort" using the combination of the IL-2 Luc and IL-2 Luc LTT.

Conclusions

In summary, the IL-2 Luc LTT assay was well described, well defined, and has a clear protocol and criteria for data interpretation. All necessary information, including performance standards, was sufficiently detailed. Both within- and between-laboratory reproducibility were satisfactory. While the predictive capacity was not satisfactory for a stand-alone method, the IL-2 Luc LTT assay is acceptable for use in an Integrated Approach to Testing and Assessment for immunotoxicity.

Acknowledgments

The peer review panel would like to acknowledge the hard work provided Dr. Aiba and Dr. Kimura in leading the laboratory effort, and Dr. Kojima and Dr. Ashikaga in coordinating the trials through JaCVAM. The peer review panel also acknowledges the validation management team, the

chemical selection committee, data management team, and members of the laboratories in the participating test facilities. The report is extensive and there has been a strong effort in not only conducting the necessary studies to get the IL-2 Luc LTT submitted as a test guideline but presenting additional data and analyses to use it in combination with IL-2 Luc as part of an IATA for immunotoxicity. The peer review panel also thanks Dr. Aiba for being very responsive to our concerns throughout the review process.

Peer Review Report of the Validation of the IL-2 Luc Leukocyte Toxicity Test (IL-2 Luc LTT) Assay on In Vitro Immunotoxicity

Series on Testing and Assessment No. 392

This document contains the report of the peer review organised by Japan for the validation of the IL-2 Luc LTT assay on in vitro immunotoxicity, a project on the work plan of the Test Guidelines Programme until 2024.