IL-2 Luciferase (IL-2 Luc) Assay Report of the Peer Review Panel

on

a JaCVAM co-ordinated study programme addressing the validation status of the IL-2 Luc assay for the prospective identification of immunotoxic substances on T cells

Report completed by the Peer review Panel on June 30th, 2020

Table of Contents

Executive Summary

The IL-2 luciferase (Luc) assay has been proposed as an in vitro alternative to animal testing. The assay provides information on adverse outcome pathways for immunotoxicity to T cells.
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 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 the use of compounds with a better-defined immunotoxicity mechanism will help to improve the accuracy of the assay. While the predictive capacity was not satisfactory for a stand-alone method, the IL-2 Luc assay is acceptable for use in an Integrated Approach to Testing and Assessment.

Peer Review Panel Composition

Fujio Kayama (Chair)	Jichi Medical University, Japan
Henk van Loveren	Maastricht University, The Netherlands
Haley Neff-LaFord	Seattle Genetics, Inc., USA
Barbara Kaplan	Mississippi State University, USA
Xingchao Geng	National Center for Safety Evaluation of Drugs (NCSED), China
Takao Ashikaga	JaCVAM, Kawasaki, Japan
Sang-Hyun Kim	Kyungpook National University, Korea

Background

Immune dysregulation can have serious adverse health consequences and it could be caused by many types of chemicals, such as environmental contaminants, food additives, and drugs. It ranges from reduced resistance to infection and neoplasia to allergic and autoimmune conditions. For many years, to identify such immunotoxic chemicals depended on animal models. For animal welfare, ethical and scientific reasons there has been a desire to replace in vivo methods with non-animal alternatives (1). The immune system comprises innate and adaptive immunity. Both arms of the immune response function differently and are driven by different population of cells. A variety of intracellular signaling pathways also play roles in innate and adaptive immune responses. Given the complexity of the immune system, it is unlikely that a single in vitro method will cover all immunotoxicants. Therefore, Integrated Approaches to Testing and Assessment (IATA) are flexible approaches for safety assessment based on the integration and translation of the data derived from multiple methods and sources. For example, target compounds are tested using various approaches such as the human whole blood cytokine release assay (HWBCRA), lymphocyte proliferation assay, mixed lymphocyte reaction and fluorescent cell chip assay after the evaluation of myelotoxicity (2). There are many in vitro methods aiming at evaluating various aspects of molecular and cellular events in the adverse outcome pathway (AOP) for immunotoxicity. However, no in vitro method is validated formally.

The IL-2 Luc assay was developed to be part of a high-throughput screening system that enables evaluation of chemical immunotoxicity. This screening system was named Multi-ImmunoTox assay or MITA (3) (4) (5). The IL-2 Luc assay, using a human cell line transfected with luciferase genes under control of the IL-2 promoter, identifies the effect of chemicals on the IL-2 activity in the 2H4 cells in the presence of stimulants (3).

The PRP first met in February 2019 to review a progress report on the IL-2 Luc assay prepared by the Validation Management Team (VMT). Following commentary on this work by the PRP, the VMT refined the validation report.
The PRP engaged in follow-up telephone conferences in October 2019, November 2019 and June 2020. With the provision of all of the amended, updated and additional material, including the final VMT report, this PRP Report was prepared.

IL-2 Luc Test Method Definition

The PRP confirmed that the IL-2 Luc assay test method has been fully described in the

report of the VMT and in the associated detailed test protocol. During the validation study, the test developer changed their prediction model. A clear definition of the 35% threshold and its reason was explained. The VMT report describes the need for the assay in the current regulatory context (6). Furthermore, a clear rationale for the assay has been given (the rationale for the test method is that drugs and chemicals, environmental contaminants, food additives, and drugs can target the immune system, resulting in immune dysregulation). It is known that IL-2 exerts pleiotropic actions on CD4+ T cell differentiation via its modulation of cytokine receptor expression. It promotes Th1 differentiation by inducing IL-12Rb2 (and IL-12Rb1), promotes Th2 differentiation by inducing IL-4Ra, inhibits Th17 differentiation by inhibiting gp130 (and IL-6Ra), and drives Treg differentiation by inducing IL-2Ra. IL-2 also potently represses IL-7Ra, which decreases survival signals that normally promote cell survival and memory cell development (7). Therefore, it is reasonable that the test developer focused on the regulation of IL-2 transcription and attempted to construct an AOP of immunotoxicity with transcriptional dysregulation of IL-2 as a central key event. The VMT report mentioned that IL-2 Luc assay will be part of a broader tiered approach to eventually include IL-8 and IL-1β, that corresponds to the AOPs for immunotoxicity and as such constitutes an IATA.

The PRP agreed that the mechanistic basis of the method and how it related to the T-cell specific endpoint also was well described in the VMT report.

Within Laboratory Reproducibility

The PRP agreed that the results which emerged have demonstrated a sufficient degree of within laboratory reproducibility. For achieving such conclusion, the PRP focused on results obtained with the final protocol and prediction model.
A total of 5 coded chemicals (4 T-cell targeting and 1 non T-cell targeting) were evaluated by 3 experimental sets. Based on such assumptions, the success criterion of >80% within-laboratory reproducibility was achieved in each of the three participating laboratories (Lab. A: 80.0% (4/5), Lab. B: 100% (5/5), Lab. C: 80.0% (4/5)).

The PRP notes that extensive documentation of within-laboratory reproducibility data for the final and all the development phases of the IL-2 Luc assay has been displayed in the VMT report and its appendices. Taken into account, the data lends support to the view that the assay has a sufficient level of reproducibility within laboratories.

Interlaboratory Transferability

The PRP noted that the technical transfer of the IL-2 Luc assay involved training and successful assessment of 3 experiments of 5 test substances (not blinded) by each of the participating laboratories. That work was prior to their approval to participate in the subsequent validation work.

Between Laboratory Reproducibility

With regard to between-laboratory reproducibility, the PRP recognized that the test results gave 80% (20/25) and met the success criterion of >80% between-laboratory reproducibility. The number of test chemicals is combined of the Phase I (5) and Phase II (20) .

Again, the PRP notes that extensive and transparent documentation of between-laboratory reproducibility data for all phases of the IL-2 Luc assay has been displayed in the VMT report and its appendices.

The PRP concluded that the assay demonstrated successful between-laboratory reproducibility.

Predictive Capacity

To determine the predictivity of the IL-2 Luc assay, it is crucial to understand the immunotoxic characteristics of chemicals used in the study. The PRP agreed that classification chemicals into those that affect T cell function (T cell-targeting chemicals, TTC) and those that do not directly affect T cell function (non-T cell-targeting chemicals, NTTC) was needed. The PRP confirmed the rationale for classifying immunotoxic chemicals are clearly described in the validation report.

Demonstration of a test method's performance should be based on the testing of representative, preferably coded, reference chemicals. The PRP concluded that the validation study used an appropriate level of test chemical coding to ensure fully blinded evaluation. With respect to chemical selection, the PRP confirmed that the criteria for chemical selection were clearly outlined. On the other hand, it should be noted that there is a question as to whether or not the number of true negatives (8/25) in the set was sufficient.

The immunotoxic characteristics of each chemical used in the Phase I and Phase II are

shown in the VMT report and based on the criteria total 25 chemicals were classified into 16 positives, 8 negatives and 1 unclassified. According to the classification, accuracy is 75% (18/24), specificity is 75% (6/8) and sensitivity is 75% (12/16). The PRP concluded the predictive capacity of the test is not sufficient to detect all immunotoxic chemicals if used as a stand-alone test.

The PRP basically agreed with the test developer's opinion that there are at least 2 reasons that the predictivity did not meet the success criterion. First, the classification of immunotoxic chemicals is sometimes uncertain, because mechanistic information is often limited. Second, the IL-2 Luc assay does not cover every aspect of the effects of chemicals on T cell function.

Regarding the second point, the PRP noted again this assay should be used in combination with other assays, as in the context of IATA. With respect to the first point, the PRP noted the use of compounds with a better-defined immunotoxicity mechanism will help to ultimately improve the accuracy of the validation study and also recommended a Detailed Review Paper (DRP) on in vitro immunotoxicity testing.

Following phases I and II, the assay was then applied to over 60 chemicals that were previously evaluated by the test developer (8). The accuracy was calculated as 82.4%, which is in line with results obtained in phases I and II.

Applicability Domain

The PRP concluded that the applicability domain should be better defined. First, the method cannot detect immunotoxicity associated with inhibition of DNA synthesis and cell division. Second, the assay might not detect compounds that require metabolic activation to a toxic intermediate. Third, the use of PMA/Io as a stimulant bypasses signaling through the T cell receptor. Finally, the IL-2 Luc assay shares limitations common to many suspension cell-based techniques when testing highly hydrophobic substances.

Performance Standards

The PRP concluded that the list of performance standard (PS) substances placed in Appendix 16 to the VMT report was satisfactory. Mutual Acceptance of Data (MAD) will only be guaranteed for test methods validated according to the PS, if these test methods have been reviewed and included in this Test Guideline by the OECD. The PS is

supplemented by a list of proficiency chemicals, listed in Appendix 15, to be used as a routine check on performance of the assay.

Additional Comments

The PRP concluded that the validation study management and conduct met the criteria set out in OECD GD 34 (2005). The PRP noted that the study was conducted not under GLP certification but in the spirit of GLP.

The PRP appreciated the transparency with which all the IL-2 Luc assay material was presented. The PRP noted that during the review they were able to access the raw data files associated with the IL-2 Luc assay development/validation work.

The PRP also noted that AOP networks and a DRP in this field are essential for the construction of an IATA.

Conclusions and Recommendations

The PRP concluded that, even though the predictive capacity was not sufficient to allow use as a stand-alone test, the IL-2 Luc assay validation has demonstrated that the method should be acceptable as a part of IATA for the predictive screening of T-cell targeted immunotoxicity.

Acknowledgements

The PRP is also grateful to the members of the VMT for supplying the report on the validation study of the IL-2 Luc Assay. This research was supported by the WHLW grant.

References

1) Adler et al. (2010) Alternative (non-animal) methods for cosmetics testing: current status and future prospects. Arch Toxicol. 2011 85(5):367-485
2) Corsini E and Roggen EL., 2017. Overview of in vitro assessment of immunotoxicity. Curr Opin Toxicol. 5: 13-18.

3) Saito, R., Hirakawa, S., Ohara, H., et al., 2011. Nickel differentially regulates NFAT and NF-kappaB activation in T cell signaling. Toxicology and applied pharmacology, 254, 245-255.
4) Takahashi, T., Kimura, Y., Saito, R., et al., 2011. An in vitro test to screen skin sensitizers using a stable THP-1-derived IL-8 reporter cell line, THP-G8. Toxicological sciences: an official journal of the Society of Toxicology 124, 359-369.
5) Kimura, Y., Fujimura, C., Ito, Y., et al., 2014. Evaluation of the Multi-ImmunoTox Assay composed of 3 human cytokine reporter cells by examining immunological effects of drugs. Toxicol In Vitro 28, 759-768.6) (WHO)/, W.H.O., Meeting, I.P.o.C.S.I.S., 2012. Guidance for immunotoxicity risk assessment for chemicals. WHO press.
7) Liao, W., Lin, J.X., Wang, L., et al., 2011, Modulation of cytokine receptors by IL-2 broadly regulates differentiation into helper T cell lineages. Nat Immunol. 12: 551-559.
8) Kimura, Y., Fujimura, C., Ito, Y., et al., 2018. Profiling the immunotoxicity of chemicals based on in vitro evaluation by a combination of the Multi-ImmunoTox assay and the IL-8 Luc assay. Arch Toxicol 92, 2043-2054.