# The *in vivo* erythrocyte *Pig-a* gene mutation assay

## Part 2: Validation report

Prepared by:

Robert H. Heflich*,[1], Stephen D. Dertinger[2], Takafumi Kimoto[3]

for the Organization for Economic Cooperation and Development Working Group of the

National Coordinators of the Test Guidelines Programme

[1]U.S. Food and Drug Administration, Jefferson, AR USA

[2]Litron Laboratories, Rochester, NY USA

[3]Teijin Pharma, Tokyo, Japan

*Direct correspondence to:

    Robert H. Heflich, Ph.D.
    Division of Genetic and Molecular Toxicology
    U.S. Food and Drug Administration
    National Center for Toxicological Research
    3900 NCTR Rd., Jefferson, AR USA
    Ph: +1-870-543-7493, Fax: +1-870-543-7393; email: robert.heflich@fda.hhs.gov

Contents

**Part 2: Validation Report**

## 1. Introduction

The elements of assay validation are outlined in OECD Guidance Document 34 (GD 34; OECD, 2005). The following sections address the validation elements in the GD. Because the *Pig-a* assay is an *in vivo* assay, the WNT agreed that it was prudent, and in the interests of reducing unnecessary animal use, that the *Pig-a* assay validation be conducted using a Retrospective Performance Analysis (RPA) rather than conducting additional animal testing intended specifically for justifying development of an OECD TG.

## 2. Rationale for the test method

The rationale and principle of the *Pig-a* assay are explained in Part I, Section 1 (the Detailed Review Paper (DRP)). As it is typically employed for evaluating the mutagenicity of test substances, the assay measures *in vivo* mutation induced in bone marrow erythroid cells by evaluating peripherial blood erythrocytes for the presence or absence of glycosylphosphatylinositol (GPI)-anchored cell surface proteins. The assay is performed using immunofluorescent identification of cells with GPI-anchored proteins and their enumeration by flow cytometry. Fluorescent cells are wild-type, while non-fluorescent cells are mutant.

## 3. Scientific basis

The scientific basis for the *Pig-a* assay is explained in Part 1, Sections 1 and 6. Because it is the only gene involved in GPI-anchor biosynthesis that is present as a single functional copy, inactivating mutations in the *Pig-a* gene are likely to be responsible for the mutant phenotype measured in the *Pig-a* assay. A variety of evidence, both direct and indirect, are consistent with this expectation. In all cases where it has been examined, increases in *Pig-a* mutant frequency

in the rodent *Pig-a* gene mutation assay have been accompanied by increases in *Pig-a* gene mutation.

## 4. Regulatory purpose and need

The *Pig-a* assay fills an unoccupied niche of providing a relatively rapid, inexpensive *in vivo* gene mutation assay that can be readily integrated in other general toxicology and genetic toxicology assays, that can be performed in non-transgenic (conventional) animals, and that has a minimum impact on animal well-being. A number of existing regulatory safety guidelines (*e.g.*, International Conference on Harmonization (ICH) S2(R1) and ICH M7; European Chemicals Agency (ECHA) Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) regulations) recommend *in vivo* genotoxicity testing, either as a part of the primary testing scheme or as a follow-up to positive *in vitro* test findings (see Cimino, 2006, for a somewhat dated, but comprehensive review; Galloway, 2017; ECHA, 2016). In addition, EFSA, ICH M7, and REACH regulatory guidance documents specifically recommend using an *in vivo* gene mutation assay to follow-up on positive responses in *in vitro* gene mutation assays.

The primary difficulty in addressing the existing regulatory requirements for evaluating *in vivo* gene mutation is that the TGR assay, the only *in vivo* gene mutation assay with an OECD TG and wide regulatory acceptance, is expensive, time consuming, and almost always requires performing a stand-alone assay, with resulting inefficient use of animal resources (a comparison of the TGR and *Pig-a* assay can be found in Section 3 of the DRP). As a result, the TGR assay is performed, and *in vivo* gene mutation assessed, only when absolutely necessary (Boverhof et al., 2011; Page et al., 2015). The *Pig-a* assay may be a practical alternative, that can substitute for the TGR assay under certain circumstances, and make *in vivo* gene mutation assessment more likely to be performed (Page et al., 2015). The IWGT report recommends that the *Pig-a* assay can be used for evaluating somatic cell mutation *in vivo*, provided that evidence of bone

marrow exposure by the test article or its metabolites can be established for supporting negative responses (Gollapudi et al., 2015).

ICH M7 (ICH, 2014) specifically recommends using the *Pig-a* assay for follow-up of positive response in *in vitro* gene mutation tests, and the U.S. Food and Drug Administration has accepted *Pig-a* gene mutation data to evaluate a drug contaminant (Custer, 2012). Galloway (2017), in commenting upon the ICH M7 recommendation, indicates 'Although an OECD guideline does not yet exist for the *Pig-a* assay, there is extensive experience, and protocol recommendations have been published by an IWGT Workgroup [Gollapudi et al., 2015].' In addition, a U.S. Environmental Protection Agency Policy and Guidance statement (US EPA OPP, 2012) says that, 'although the *Pig-A* gene mutation assay does not have an OECD Test guideline, it is a promising new *in vivo* mutation test that is sensitive and less costly than the transgenic rodent gene mutation assay, and it can be integrated into repeat-dose standard toxicology tests.' Finally, a European Food Safety Authority (EFSA) Scientific Opinion (EFSA, 2011) devotes a section to the assay, indicating its integration potential and the need for defining its sensitivity, and establishing standard protocols for analysis and data interpretation.

Another potential regulatory use of the assay is in establishing a cancer mode-of-action (MoA) as described by the U.S. Environmental Protection Agency's Guidelines for Carcinogen Risk Assessment (US EPA, 2005). Data indicate that the assay is capable of distinguishing between carcinogens operating through a genotoxic and a nongenotoxic MoA (Bemis et al., 2015). In addition, the characteristics of the *Pig-a* assay lend themselves to generating detailed longitudinal, dose-response data for gene mutation from a single set of animals of the species and strain used in the cancer bioassay. These types of data are important for addressing the

ssues of temporality and dose response concordance that contribute to the assessment of a potential mutagenic MoA (US EPA, 2005).

## 5. The relationship between the endpoint measured and *in vivo* biological effects

As outlined in Part 1, Section 1, "Principle of the assay", the *Pig-a* assay measures gene mutation induced in bone marrow erythroid cells. The assay generally uses fluorescent immuno-staining and flow cytometry to identify peripheral blood erythrocytes having the *Pig-a* mutant phenotype, *i.e.*, cells deficient in GPI-anchored surface markers. Studies described in Section 6 of the DRP directly demonstrate that mutant nucleated bone marrow erythroid precursor cells contain mutations in the *Pig-a* gene. Also, additional lines of evidence indicate that that the phenotype measured in the *Pig-a* assay is almost certainly due to mutation and most likely due to mutation in the *Pig-a* gene. This compelling body of evidence is reviewed in Section 6 of the DRP.

Mutation is a relevant toxicological endpoint for conducting safety assessments. Mutagenicity is commonly evaluated for determining whether or not a test article is a potential carcinogen (Cimino, 2006). Mutagenicity testing is also conducted to determine the potential of test articles for inducing germ line mutations (that may result in genetic diseases) and for inducing mutations that might result in somatic cell diseases other than cancer.

## 6. Protocol for conducting the assay

General recommendations for conducting the assay were made in the IWGT report (Gollapudi et al., 2015). Those recommendations were reviewed and updated in Part 1, Section 4 and were central to evaluating the *Pig-a* assay data used in the performance analysis of the assay described in Part 2, Sections 8 and 9 below (also see Part 1, Section 7 for more detail). For several reasons, including that *Pig-a* mutations accumulate as a result of repeated dosing and a general desire to reduce animal usage by integrating the assay into standard *in vivo* toxicity

testing, the *Pig-a* assay is generally conducted using a subchronic (*e.g.*, 28-day) repeated-dose treatment protocol. However, it is recognized that it may be advantages, in certain circumstances, to conduct studies employing acute or subacute treatments.  Review of the data associated with the assay indicate that there are a very few test substances whose mutagenicity is more readily detected with repeat-dose treatment protocols (see Sections 4a and 7 of the DRP). However, the overall performance of the assay is not greatly affected by treatment schedule (Part 1, Section 7).

# 7. Test variability, repeatability and reproducibility

## a. Interlaboratory trials

### i) International interlaboratory trial

The first inter-laboratory *Pig-a* mutation assay trial was organized through an NIH-NIEHS grant to Litron Laboratories (Rochester, NY). The trial's structure, aims, and scope benefited from many important intellectual contributions made by participating laboratory scientists, as well as the Health and Environmental Sciences Institute (HESI) consortium's *Pig-a* Working Group (Schuler et al., 2011). The trial was conducted with common protocols and reagents, although the rat strain and flow cytometer make/model were at the discretion of the participating laboratories. Stage I consisted of information gathering, while Stage II tested the transferability of an early anti-CD59 antibody-based scoring method using data from Litron Laboratories as the comparator. For this study, 14 laboratories treated rats acutely with specified doses of *N*-ethyl-*N*-nitrosourea (ENU). High concordance coefficient values indicated good transferability across sites (Dertinger et al., 2011b), and these results provided the impetus to continue investigating inter-laboratory performance. Stage III was also performed with rat models chosen by collaborators, but exposure was for 28-day consecutive days, with blood sampling and *Pig-a* analyses occurring at several time points. The following 5 diverse mutagenic agents were studied over the course of Stage III, with the results published in a special issue of *Environmental and Molecular Mutagenesis*: ENU (Cammerer et al., 2011), dimethylbenz[*a*]anthracene (DMBA; Shi et al., 2011), *N*-methyl-*N*-nitrosourea (Lynch et al.,

2011), benzo[a]pyrene (Bhalli et al., 2011), and 4-nitroquinoline-1-oxide (4NQO; Stankowski et al., 2011). Each chemical was studied in at least 2 laboratories, and the results once again showed a high level of agreement across sites.

While Stage II and III results suggested good interlaboratory transferability, the data also reinforced the notion that across laboratories and rat models, the frequency of mutant phenotype erythrocytes in untreated or vehicle control animals was generally on the order of 1 mutant phenotype cell per million total red blood cells (RBCs) or reticulocytes (RETs). The rarity of mutants suggested that the assay would benefit from increasing the number of cells evaluated per sample, thereby reducing the number of analyses that returned a zero mutant frequency value. Stage IV studies were therefore initiated only <u>after</u> new methodology—that is the use of immunomagntic separation—was successfully devised and instituted (Dertinger et al., 2011a). Given its higher statistical power relative to earlier methods without separation, this so-called *In Vivo* MutaFlow scoring approach was used for Stage IV studies (Dertinger et al., 2011c). Because of the high concordance this method demonstrated with expected outcomes, it continues to be responsible for a large fraction of rodent data originating from the US and Europe (see data deposited in the *Pig-a* mutant database described in Section 11).

### *ii) Japanese interlaboratory trial*

Two serial collaborative studies on the *Pig-a* gene mutation assay have been conducted in Japan. In the first of these collaborative studies, five participant laboratories evaluated the transferability of the rat *Pig-a* assay using both RBCs (hereafter referred to as the RBC *Pig-a* assay) and RETs (hereafter referred to as the PIGRET assay; Kimoto et al., 2011). The trial involved treatment of rats with a single dose of 3 typical mutagens (ENU, 4NQO, and DMBA). This trial successfully demonstrated that both the RBC *Pig-a* assay and the PIGRET assay were technically transferable and reproducible among participant laboratories (Kimoto et al., 2013). The subsequent trial of the collaborative study examined the utility of the *Pig-a* assay under a repeat dosing regimen. Four of the five laboratories assayed a series of test agents using a 28-day repeat dosing protocol. The responses were generally positive, with the PIGRET assay detecting the *in vivo* mutagenicity of chemicals earlier than the RBC *Pig-a* assay (Itoh et al.,

2014; Sanada et al., 2014; Muto et al., 2014; Kimoto et al., 2014). The fifth laboratory used *gpt-delta* rats that were treated with a single treatment of ENU. This study demonstrated a consistent relationship between increases in both RBC and RET *Pig-a* mutant frequency and increases in *gpt* mutant frequency in both bone marrow and liver (Horibata et al., 2014).

Subsequently, the Mammalian Mutagenicity Study Group of the Japanese Environmental Mutagen Society (MMS/JEMS) organized a collaborative research group to conduct a more extensive multi-laboratory study of the PIGRET assay, which was expanded to a total of 16 collaborators and 24 test chemicals (Kimoto et al., 2016). The MMS/JEMS-collaborative study consisted of 2 steps. In Step 1, a total of 13 laboratories that did not participate in the previous collaborative study conducted both the PIGRET and RBC *Pig-a* assays in rats given a single administration of ENU, to confirm assay transferability and proficiency. The concordance correlation coefficients for the PIGRET and RBC *Pig-a* assays ranged from 0.89 to 0.99 and from 0.91 to 0.99, respectively, indicating that both assays were successfully transferred to these additional laboratories. The outcome of Step 1 also demonstrated the reproducibility of dose responses by comparing data generated by each of the participant laboratories using a potency ranking procedure that employs covariate benchmark dose analysis (Johnson et al., 2016). In Step 2, the 24 test chemicals were tested for mutagenicity using a single-administration protocol. The outcomes were compared with the results of the MutaFlow method that were reviewed in the IWGT report, and which described the US/European lab experiences. Consistent results were obtained for 21 of the 24 chemicals (see Figure 3 in Kimoto et al., 2016), with the exceptions being the three chemicals, 2-AAF, azathioprine, and DEN. Although further research is necessary to understand the difference in outcomes, it was concluded that it may be difficult to detect the mutagenicity of some chemicals such as those described above under a single-dosing protocol (see analysis in Part I, Section 4). Otherwise, concordance was seen between the Step 2 study and the US/European lab experiences, suggesting that detecting a positive or negative *Pig-a* mutant frequency is reliable regardless of the approach. Of note, the study collaborators also concluded that the PIGRET assay offers strong potential as a short-term test for *in vivo* mutation using a single-dose protocol.  The PIGRET procedure can detect an increase

in RET *Pig-a* mutant frequency beginning 1 week after the dosing for most of the chemicals that tested positive.

### iii) Korean interlaboraty study

The most recent interlaborary trial was conducted by the Korean Ministry of Food and Drug Administration (MFDA) in three laboaorties. As a preliminary step, each laboratory demonstrated their proficieny by performing an assay measuring *Pig-a* mutant frequencies in ENU-treated rats (Chung et al., 2017). For the main study, each laboratory tested three expected genotoxins having different mechanisms of action, 1,3-propanesultone, mitomycin C, and *N*-propyl-*N*-nitrosourea (Chung et al., 2018).   The study design employed treatment on 3 consecutive days, sampling on Days -1, 15 and 29, and mutant analysis using the Mutaflow immunomagnetic enrichment procedure. In addition, one laboratory integrated the *in vivo* micronucleus and comet assay endpoints into the study. All the labs detected the three test articles as positive and produced quantiatively similar *Pig-a* mutant RET and total RBC frequencies, demonstrating the transferability and reproducibility of the assay. In additional, the trial confirmed that the *Pig-a* assay could readily be integrated into GLP-compliant genotoxicity studies employing the comet and micronucleus assays.

### iv) Additional interlaboratory studies

Instead of assessing transferability through inter-laboratory experiments that include every aspect of a full *Pig-a* study, there are advantages to splitting the same coded blood samples and analyzing them at two or more analytical sites. This design reduces the number of animals required, and better focuses the assessment on the sample processing and analysis phases. One such study was described by Gollapudi *et al.* (2015), whereby blood samples where collected from 15 Sprague Dawley rats exposed for 28 consecutive days to either vehicle or various dose levels of thiotepa. Blood samples were collected and processed by Litron personnel (Rochester, NY) for same-day determination of mutant RET and total RBC frequencies. A second aliquot of each whole blood sample was maintained in a refrigerator for next-day labeling and analysis at

Litron, while a third set of coded aliquots was shipped overnight to Groton, CT so that collaborators at Pfizer could perform labeling, magnetic separation, and flow cytometric analysis. Next-day, same-site analyses produced *Pig-a* mutant frequencies similar to those obtained with fresh samples, demonstrating good intra-laboratory reproducibility. Next-day, off-site analyses conducted at Pfizer also showed close agreement with mutant cell frequencies obtained at Litron. These data demonstrate good inter-laboratory transferability of the scoring method and illustrate the compatibility of blood sample shipments between sites.

Raschke et al. (2016) reported on a 3Rs friendly experimental design that allows for critical assessment of inter-laboratory reproducibility. This work involved reconstruction experiments, also known as spiking experiments, whereby blood from a negative control rat was combined with known proportions of blood from either an ENU-treated rat or else a blood sample comprised of mutant mimics (i.e., CD59 epitopes blocked with non-fluorescent antibodies) (see also Part 1, Section 10). These reconstruction experiments were performed independently in four laboratories, using the MutaFlow immunomagnetic enrichment method and showed good overall scoring precision (correlation coefficients >0.99) as well as accuracy (estimated slope 0.71–1.09). These results provide further evidence of the transferability and reproducibility of the *Pig-a* assay.

### b. Intralaboratory studies

Although not as well documented as the organized interlaboratory trials, there are a few published reports on the intralaboratory reproducibility of the *Pig-a* assay.

Gollapudi et al. (2015) provide several examples indicating a high degree of reproducibility when the same laboratory analyzed technical replicates from both mutagen treated and negative control animals on the same day and on different days. These assays were performed using an earlier version of the assay, not involving immunomagnetic separation (referred to as the basic method in the DRP).

Godin-Ethier *et al*. (2015) estimated intra-run (single labeling of a single sample, with multiple analytical runs) and inter-run (multiple independent labelelings and analysis of a single sample)

coefficients of variation of <4.8 and 20.8% for negative control and ENU-treated rat total RBC and RET mutant frequencies determined using the Mutaflow method.  Assay reproducibility was not affected by up to 4 days of storage, the analyst performing the assay, or small changes in technique that were tested (*e.g*, tube vs plate analysis, difffernt blood volumes, increased antibody reaction time).

A study described by Avlasvich et al. (2019) describes RET and total RBC mutant frequencies in blood samples from ENU- and carboplatin-treated rats and negative control rats. The study, which involved three different laboratories, analyzed mutant frequencies in portions of samples that were analyzed at the time of collection and that were frozen and analyzed months later. No effect of freezing and reanalysis was found among the mutant frequencies in the repetative analysis of the same samples, whether the frozen samples were analyzed in the same laboratory or shipped and analyzed in a second laboratory.

## 8. Test method performance using a series of reference chemicals

Test method performance was evaluated using a Retrospective Performance Analysis (RPA) as described in OECD GD 34. The process followed for conducting the RPA is detailed in Part 1, Section 7. Briefly, two semi-independent subcommittees were formed that were made up of experts in the assay and in genotoxicity testing in general. The *Pig-a* subcommittee assembled the available *Pig-a* assay data, evaluated the data for quality and conformity to protocol guideline recommendations, performed statistical evaluations where necessary, and evaluated assay responses. The guidelines for this analysis are described in Part 1, Section 7. Responses were categorized as positive (P), negative (N), equivocal (E), or inconclusive (I) using three sets of criteria that used data from studies with acute and subacute treatment protocols to a greater or lesser degree (see Table V of the DRP).  The test responses determined for individual data sets from mouse and rat studies, overall calls for each test agent, and the reasons for making E or I calls, are shown in Tables VI, VII and VIII of the DRP (Part I).

In addition to *Pig-a* assay data on the test articles shown in Part 1, Tables VI, VII and VIII, vechicle substances used for *Pig-a* assays were evaluated to determine their mutagenicity in the assay.  The approach used for analyzing these data was similar to that used for the conventional test substances.  These analyses were all based on the primary data in the *Pig-a* assay data websites located at the University of Maryland (see Part 2, Section 11, below, and Part 1, Section 7) and the citeria used for data acceptance and interpretation are given in Part 1, Section 7.

A second subcommittee reviewed literature and data bases to generate P, N, E, or I responses for rodent cancer and relevant genotoxicity assays for each of the test and vehicle control substances for which P or N *Pig-a* responses were available. Consensus response calls were generated for rodent cancer, rodent hematopoietic system cancer, the TGR assay, the bone marrow TGR assay, and the *in vivo* MN assay. The process used for this analysis is outlined in Part 1, Section 7 and the resulting consensus responses are shown in Part 1, Table XI.

GD 34 recommends that it is 'perferable' (but not required) that test agents be coded in order to remove bias in the results of the performance analysis due to prior knowledge of expected responses. Although none of the *Pig-a* data used for the performance analysis were produced using coded samples, most of the data generation in the *Pig-a* assay is automated by the flow cytometric analysis used for the endpoint, clear data acceptance and data interpretation standards have been used for evaluating responses in the assay, and *Pig-a* response calls were made independently of response calls in the conventional assays used to evaluate *Pig-a* assay performance. Thus, significant efforts were made to minimize any bias in the assay performance evaluation.

## 9. Evaluation of performance relative to relevant toxicity data

GD 34 indicates that a performance analysis should 'describe the accuracy (*e.g.*, sensitivity, specificity, positive and negative predictivity, false positive and false negative rates) of the proposed test method and it should be compared to that obtained for the reference test

method currently accepted by regulatory agencies….' Although the *Pig-a* assay is not a replacement for the TGR assay, data from the TGR assay can be used as a reference test for evaluating the performance of the *Pig-a* assay as an *in vivo* gene mutation assay. The *Pig-a* assay is designed to measure *in vivo* gene mutation originating in erythroid precursor cells of the bone marrow, while the TGR assay can measure gene mutation in virtually any tissue. TGR responses were not found for erythroid precursor cells; however, the *Pig-a* performance analysis used TGR mutant responses in bone marrow, which incudes responses from erythroid cells.  In addition to comparing *Pig-a* responses to TGR responses in bone marrow, the analysis compared *Pig-a* responses to TGR responses in all tissues in order to evaluate the accuracy of the *Pig-a* assay for predicting *in vivo* mutation in general. For these comparisons the degree of assay sensitivity, positive predictivity, specificity, negative predictivity and concordance was calcululated. In addition, Cohen's kappa and the Prevalence And Bias Adjusted Kappa (PABAK) tests were used to further evaluate the relatedness of the responses. Note that because of the interest in conducting the assays using subacute and acute, as well as subchronic treatment protocols, analyses were conducted using three sets of response evaluation criteria that made greater or lesser use of *Pig-a* data from acute and subacute treatment protocols. Summary results from these comparisons are shown in Tables XII, XIII, and XIV of the DRP. Additional details of the analyses can be found in Annex III of the DRP.

GD 34 also states that 'In cases where the proposed test method is mechanistically and functionally similar to a validated test method with established performance standards, the accuracy of both test methods should be compared.' In order to address this recommendation, *Pig-a* responses were evaluated using the *in vivo* MN assay and tumor induction in the rodent hematopoetic system and rodent cancer in any tissue as reference assays. The MN assay measures an endpoint that is, in large part, structurally and functionally distinct from *Pig-a* gene mutation; however, the hematopoetic tissue MN assay is the most commonly employed *in vivo* genotoxicity assay, it evaluates genotoxicity in the bone marrow, and it measures an endpoint that is often involved with or coincident with the induction of gene mutation.  As for using tumor induction to define true positives and negatives, gene mutation is mechanistically involved with carcinogenesis, although nongenoxicants can be carcinogens due to secondary

mechanisms.  Because of these considerations, lower accurary was expected when MN or rodent tumorigenicity were used as reference assays.

The analyses indicate near perfect accuracy for the *Pig-a* assay when bone marrow TGR responses or cancer in hematopoetic tissues was used to define true positive and negative responses. Slightly lower accuracy was found when responses in the *in vivo* MN assay or TGR responses in any tissue were used to define true positive and negative responses, and when cancer in any tissue was used as the reference assay. Where the accuracy of the *Pig-a* assay was less than perfect, it was usually due to a reduction in negative predictivity, *i.e.*, test subtances that were negative in the *Pig-a* assay were positive in the reference assay.

In general, the discordant responses were due to 1) mechanisms for induction of the reference assay responses that did not involve gene mutation (or were not coincident with the induction of gene mutation) or 2) responses that did not require bone marrow exposure.  Discordant responses when the MN assay was used as the reference assay generally fell into the first category, while discordant responses when TGR in any tissue was used as the reference assay generally fell into the latter. Some discordant responses appeared to fall into both catagories, especially for cancer in any tissue where tissue specificity and nongenotoxic carcinogens contributed to discordant responses. In addition, there were a very few instances (*e.g.*, for diethylnitrosamine and acrylamide) where the *Pig-a* assay may have been more or less sensitive for the detection of *in vivo* gene mutation in bone marrow than the TGR assay. A more detailed discussion of the discordant responses is given in Part 1, Section 7.

Note that the precision of these analyses, especially those involving hematopoietic tissue TGR and tumor data, may have been compromised by the relatively low number of observations associated with these responses. However, within the limitations of the data set, the analyses indicate that the *in vivo Pig-a* assay has a high degree of accuracy for identifying test substances that induce *in vivo* gene mutation and rodent cancer, and that it does particularly well at identifying test substancess that induce mutation and rodent cancer in hematopoietic tissues.

## 10. Generation of the data using Good Laboratory Practices principles

A minority of the *Pig-a* data evaluated in this report were generated as part of studies formally complying with the principles of Good Laboratory Practice (GLP). The known exceptions are studies conducted by contract research organiztions (*e.g.*, Covance, BioReliance, Biotoxtech), the multilab study conducted by the Korean MFDA, and the NNK study conducted by the US FDA at NCTR. However, many *Pig-a* studies, because of the extensive planning and resources involved in executing them, are conducted 'in the spirit of GLP' to maximize the value of the results. Also, as these studies are conducted with animals, they are evaluated for their appropriate use of animal resources by institutional animal use and care committees, which only approve the use of animals in well-designed studies that comply with known standards for conducting assays that have a high probably of producing useful results.

## 11. Availability of data for expert review

The *Pig-a* data that were used to prepare Part 1 (the DRP), including those used for the RPA of the assay, were derived from two sources. One primary source was the peer reviewed literature that is cited in Part 1, Tables VI and VII. The second is primary data that has been collected on a website at the School of Pharmacy, University of Maryland, Baltimore USA. Primary data from published manuscripts, or data otherwise cleared for general release, are available at https://www.pharmacy.umaryland.edu/centers/cersi-files/. A companion website is maintained containing data that is not available for public release but that may be made available for expert review upon request. In addition, statistical analyses performed on these primary data sets, especially on data from the non-public website, are available for review upon request.

## 12. Summary and conclusions

A validation of the *in vivo Pig-a* gene mutation assay was conducted using the principles outlined in OECD GD 34.  The rationale, scientific basis, regulatory need, relationship between

the endpoint measured, and the *in vivo* biological effect and the assay protocol were established in Part 1 of this document, the DRP, and are only summarized here. Both interlaboratory and intralaboratory studies, involving laboratories in Asia, North America, and Europe, are described that establish the variability, repeatability, and reproducibility of the assay.

An RPA of the assay was performed as described in the DRP (Part 1, Section 7) and indicates that the assay is highly accurate for detecting rodent mutagens and carcinogens that affect the bone marrow and the hematopoietic system, respectively. The assay is also reasonably accurate when TGR mutation in any tissue, *in vivo* MN induction, and rodent cancer in any tissue is used to define true positive and negative responses.

Most *Pig-a* data were generated under conditions that assure their validity, either as part of studies conducted under GLP guidelines or under the spirit of GLP guidelines. All the data used for the RPA are available for expert review in published reports and/or as original data deposited in databases on public and private websites.

It is the opinion of the authors of the DRP and validation report (Parts 1 and 2 of this document) that suffienct information is available to support development of an OECD TG for the *in vivo Pig-a* gene mutation assay.

## References for Part 2

Avlasevich SL, DK Torous, JC Bemis, JA Bhalli, CC Tebbe, J Noteboom, D Thomas, DJ Roberts, M Barragato, B Schneider, J Prattico, M Richardson, BB Gollapudi, SD Dertinger (2019) Suitability of long-term frozen rat blood samples for the interrogation of *Pig-a* gene mutation by flow cytometry. Environ Mol Mutagenesis 60:47-55.

Bemis JC, C Labash, SL Avlasevich, K Carlson, A Berg, DK Torous, M Barragato, JT MacGregor, SD Dertinger (2015) Rat *Pig-a* mutation assay responds to the genotoxic carcinogen ethyl carbamate but not the non-genotoxic carcinogen methyl carbamate. Mutagenesis 30:343-347.

Bhalli JA, JG Shaddock, MG Pearce, VN Dobrovolsky, X Cao, RH Heflich, H-W Vohr (2011) Report on Stage III *Pig-a* mutation assays using benzo[*a*]pyrene, Environ Mol Mutagenesis 52:731-737.

Boverhof DR, MP Chamberlain, CR Elcombe, FJ Gonzalez, RH Heflich, LG Hernandez, AC Jacobs, D Jacobson-Kram, M Luijten, A Maggi, MG Manjanatha, J van Benthem, BB Gollapudi (2011) Transgenic animal models in toxicology: Historical perspectives and future outlook, Toxicol Sci 121:207-233.

Cammerer Z, JA Bhalli, X Cao, SL Coffing, D Dickinson, KL Dobo, VN Dobrovolsky, M Engel, RH Fiedler, WC Gunther, RH Heflich, MG Pearce, JG Shaddock, T Shutsky, CJ Thiffeault, M Schuler (2011) Report on Stage III *Pig-a* mutation assays using *N*-ethyl-*N*-nitrosourea - Comparison with other in vivo genotoxicity endpoints, Environ Mol Mutagenesis 52:721-730.

Chung YS, BS Pak, SH Han, JY Lee, JY Kim, SM Back, CR Park, SH Kim, JK Lee (2017) Evaluation of the transferability of in vivo Pig-a gene mutation assay across three laboratories, J Altern Anim Exp 11:19-26.

Chung Y-S, B Pak, S Han, J Lee, J Kim, S-M Back, C-R Park, S-H Kim, J-K Lee (2018) Multi-laboratory evaluatin of 1,3-propane sultone, N-propyl-N-nitrosourea, and mitomycin C in the Pig-a mutation assay in vivo, Mutat Res 831:62-68.

Cimino MC (2006) Comparative overview of current international strategies and guidelines for genetic toxicology testing for regulatory purposes, Environ Mol Mutagenesis 47:362-390.

Custer LL (2012) Impact of new mutagenicity data on marketed drugs: A case study, Annual Meeting of the Genetic Toxicology Association, October 24-25, Newark, DL. Available at: http://www.gta-us.org/scimtgs/2012Meeting/speakers2012.html.

Dertinger SD, SM Bryce, S Phonethepswath, SL Avlasevich (2011a) When pigs fly: Immunomagnetic separation facilitates rapid determination of *Pig-a* mutant frequency by flow cytometric analysis, Mutat Res 721:163-170.

Dertinger SD, S Phonethepswath, P Weller J Nicolette, J Murray, P Sonders, H-W Vohr, J Shi, J Krsmanovic, C Gleason, L Custer, A Henwood, K Sweder, LF Stankowski Jr, DJ Roberts, A Giddings, J Kenny, AM Lynch, C Defrain, F Nesslany, B-jM van der Leede, T Van Doninck, A Schuermans, K Tanaka, Y Hiwata, O Tajima, E Wilde, A Elhajouji, WC Gunther, CJ Thiffeault, TJ Shutsky, RD Fiedler, T Kimoto, JA Bhalli, RH Heflich, JT MacGregor (2011b) International *Pig-a* gene mutation assay trial: Evaluation of transferability across 14 laboratories, Environ Mol Mutagenesis 52:690-698.

Dertinger SD, S Phonethepswath, P Weller, S Avlasevich, DK Torous, JA Mereness, SM Bryce, JC Bemis, S Bell, S Portugal, M Aylott, JT MacGregor (2011c) Interlaboratory *Pig-a* gene mutation assay trial: Studies of 1,3-propane sultone with immunomagnetic enrichment of mutant erythrocytes, Environ Mol Mutagenesis 52:748-755.

ECHA (European Chemicals Agency) (2016) Guidance on information requirements and chemical safety assessment; Chapter R.7a: Endpoint specific guidance. R.7.7.1 Mutagenicity. Version 5.0. pp 528-561.

https://echa.europa.eu/documents/10162/13632/information_requirements_r7a_en.pdf/e4a2a18f-a2bd-4a04-ac6d-0ea425b2567f.

EFSA (European Food Safety Authority) (2011) Scientific opinion on genotoxicity testing strategies applicable to food and feed safety assessment, EFSA Journal 9:2379.

Galloway SM (2017) International regulatory requirements for genotoxicity testing for pharmaceuticals used in human medicine, and their impurities and metabolites, Environ Mol Mutagenesis 58:296-324.

Godin-Ethier J, F Leroux, N Wang, S Thébaud, F Merah, A Nelson (2015) Characterizsation of an *in vivo Pig-a* gene mutation assay for use in regulatory toxicology studies. Mutagenesis 30:359-363.

Gollapudi BB, AM Lynch, RH Heflich, SD Dertinger, VN Dobrovolsky, R Froetschl, K Horibata, MO Kenyon, T Kimoto, DP Lovell, LF Stankowski Jr, PA White, KL Witt, JY Tanir (2015) The *in vivo Pig-a* assay: a report of the International Workshop on Genotoxicity Testing (IWGT) Workgroup, Mutat Res 783:23-35.

Horibata K, A Ukai, M Honma (2014) Evaluation of rats' in vivo genotoxicity induced by N-ethyl-N-nitrosourea in the RBC *Pig-a*, PIGRET, and *gpt* assays, Genes Environ 36:199-202.

ICH (International Council on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use) (2014) Assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk: M7 (Step 4). Available at: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Multidisciplinary/M7/M7_Step_4.pdf.

Itoh S, M Nagata, C Hattori, W Takasaki (2014) *In vivo* mutagenicity of ethyl methanesufonate detected by *Pig-a* and PIGRET assays. Genes Environ 36:174-178.

Johnson GE, M Yamamoto, Y Suzuki, H Adachi, T Kyoya, H Takasawa, K Horibata, E Tsutsumi, K Wada, R Kikuzuki, I Yoshida, T Kimoto, A Maeda, K Narumi (2016) Measuring reproducibility of dose response data for the *Pig-a* assay using covariate benchmark dose analysis, Mutat Res 811:135-139.

Kimoto T, S Chikura, K Suzuki, Sm Kobayashi, Y Itano, K Horibata, M Honma, VN Dobrovolsky, RH Heflich, D Miura, Y Kasahara (2011) Further development of the rat Pig-a mutation assay: measuring rat Pig-a mutant bone marrow erythroids and a high throughput assay for mutant peripheral blood reticulocytes, Environ Mol Mutagenesis 52:774-783.

Kimoto T, K Horibata, S Chikura, K Hashimoto, S Itoh, H Sanada, S Muto, Y Uno, M Yamada, M Honma (2013) Interlaboratory trial of the rat *Pig-a* mutation assay using an erythroid marker HIS49 antibody, Mutat Res 755:126-134.

Kimoto T, S Chikura, K Suzuki-Okada, X Kobayashi, Y Itano, D Miura, Y Kasahara (2014) The rat *Pig-a* mutation assay in single and 28 day-repeated dose study of cyclophosphamide: The PIGRET assay can detect the *in vivo* mutagenicity earlier than the RBC *Pig-a* assay, Genes Environ 36:191-198.

Kimoto T, K Horibata, D Miura, S Chikura, Y Okada, A Ukai, S Itoh, S Nakayama, H Sanada, N Koyama, S Muto, Y Uno, M Yamamoto, Y Suzuki, T Fukuda, K Goto, K Wada, T Kyoya, M Shigano, H Takasawa, S Hamada, H Adachi, Y Uematsu, E Tsutsumi, H Hori, R Kikuzuki, Y Ogiwara, I Yoshida, A Maeda, K Narumi, Y Fujiishi, T Morita, M Yamada, M Honma (2016) The PIGRET assay, a method for measuring *Pig-a* gene mutation in reticulocytes, is reliable as a short-term *in vivo* genotoxicity test: Summary of the MMS/JEMS-collaborative study across 16 laboratories using 24 chemicals, Mutat Res 811:3-15.

Lynch AM, A Giddings, L Custer, C Gleason, A Henwood, M Aylott, J Kenny (2011) International *Pig-a* gene mutation assay trial (Stage III): Results with *N*-methyl-*N*-nitrosourea, Environ Mol Mutagenesis 52:699-710.

Muto S, K Yamada, T Kato, Y Iwase, Y Uno (2014) *Pig-a* gene mutation and micronucleus induction in rat peripheral blood by methyl methanesulfonate, Genes Environ 36:186-190.

OECD (Organisation for Economic Cooperation and Development) (2005) Guidance document on the validation and international acceptance of new or updated test methods for hazard assessment. Series on Testing and assessment, No 34. Available at: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2005)14&doclanguage=en.

Page TJ, VL Reynolds, JB Phelps (2015) Preclinical assessment of genotoxic impurities, an overview of current regulatory guidance, available assays, and data interpretation. In: Pharmaceutical Industry Practices on Genotoxic Impurities, CRC Press, Boca Raton, FL, pp 209-234.

Raschke M, B-W Igl, J Kenny, J Collins, SD Dertinger, C Labash, JA Bhalli, CCM Tebbe, KM McNeil, A Sutter (2016) In vivo *Pig-a* gene mutation assay: Guidance for 3Rs-friendly implementation, Environ Mol Mutagenesis 57:678-686.

Sanada H, M Okamoto, T Ohsumi, T Nakamura (2014) Evaluation for a mutagenicity of 4,4'-methylenedianiline on hematopoietic cells by *Pig-a* gene mutation assay in rats, Genes Environ 36:179-185.

Schuler M, BB Gollapudi, V Thybaud, JH Kim (2011) Need and potential value of the Pig-a in vivo mutation assay - a HESI perspective, Environ Mol Mutagenesis 52:685-689.

Shi J, L Krsmanovic, S Bruce, T Kelly, M Paranjpe, K Szabo, M Arevalo, S Atta-Safoh, F Debelie, MK LaForce, J Sly, S Springer (2011) Assessment of genotoxicity induced by 7,12-dimethylbenz(*a*)anthracene or diethylnitrosamine in the *Pig-a*, micronucleus and Comet assays integrated in 28-day repeat dose studies, Environ Mol Mutagenesis 52:711-720.

Stankowski LF Jr, DJ Roberts, H Chen, T Lawlor, M McKeon, H Murli, A Thakur, Y Xu (2011) Integration of Pig-a, micronucleus, chromosome aberration and Comet assay endpoints in a 28-day rodent toxicity study with 4-nitroquinoline-1-oxide, Environ Mol Mutagenesis 52:738-747.

US EPA (US Environmental Protection Agency) (2005) Guidelines for carcinogen risk assessment. Available at: https://www.epa.gov/sites/production/files/2013-09/documents/cancer_guidelines_final_3-25-05.pdf.

US EPA OPP (US Environmental Protection Agency, Office of Pesticide Programs) (2012) Advances in genetic toxicology and integration of in vivo testing into standard repeat dose studies. Available at: http://www.epa.gov/pesticides/science/integrating-gentox-studies.html.