UNITED STATES ENVIRONMENTAL PROTECTION
AGENCY WASHINGTON, D.C. 20460

OFFICE OF
RESEARCH AND DEVELOPMENT

September 21, 2018

Jerry Cook
Technical Director
Chemical Products Corporation
102 Old Mill Road
Post Office Box 2470
Cartersville, Georgia 30120

Dear Mr. Cook:

This letter is in response to the Request for Correction (RFC) received by the U.S. Environmental Protection Agency (EPA) on April 6, 2018, which was assigned RFC #18001 for tracking purposes. In the RFC letter, Chemical Products Corporation (CPC) states that *Provisional Peer Reviewed Toxicity Values for 9,10-Anthraquinone (CASRN 84-65-1)*, disseminated by EPA's Office of Research and Development (ORD) in 2011 (referred to herein as the "9,10-AQ PPRTV"), and the toxicity values for 9,10-anthraquinone (9,10-AQ) provided in EPA's Regional Screening Level (RSL) Tables, disseminated by EPA's Office of Land and Emergency Management (OLEM), do not reflect "sound and objective scientific practices" as required by EPA's *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility and Integrity of Information Disseminated by the Environmental Protection Agency* (IQGs) and requests correction.

Summary of the Request

The CPC RFC requests the 9,10-AQ PPRTV be withdrawn and revised and that the toxicity values for 9,10-AQ be removed from the RSL Tables pending revision of the 9,10-AQ PPRTV.
To support this RFC, CPC provided a letter detailing purported deficiencies in a 2-year bioassay of 9,10-AQ conducted by the National Toxicology Program (NTP) (this NTP bioassay is referred to herein as "TR 494"). The CPC RFC asserts that TR 494 does not represent "sound and objective scientific practices" as required by EPA's IQGs. Specifically, the RFC further asserts that the 9,

10-AQ PPRTV and the 9,10-AQ toxicity values provided in the RSL Tables do not comply with EPA's IQGs due to the use of information from TR 494 in their development.

Background

NTP routinely develops and disseminates scientific information about hazardous and potentially toxic chemicals. Bioassays conducted by NTP are conducted in compliance with Food and Drug Administration (FDA) Good Laboratory Practice (GLP) Regulations and undergo quality assurance audits, public comment and external peer review. TR 494 includes the results of toxicity testing conducted by NTP in male and female F344/N rats and B6C3F1 mice exposed to 9,10 -AQ in the diet for 14 weeks or 2 years. TR 494 also contains the results of genetic toxicology testing conducted in *Salmonella typhimurium*, mouse bone marrow cells, and mouse peripheral blood erythrocytes.

In developing the 9,10-AQ PPRTV, EPA carefully evaluated all available toxicity information for 9,10-AQ. As part of this evaluation, TR 494 was determined to be an appropriate study for use in developing toxicity values and was utilized for the derivation of screening subchronic and chronic non-cancer oral Reference Doses (RfDs) and an oral slope factor (OSF) for 9,10-AQ. The 9,10-AQ PPRTV was developed following all applicable EPA guidelines and was externally peer reviewed by independent scientific experts.

Following completion of the 9,10-AQ PPRTV, the 9,10-AQ toxicity values derived in the PPRTV were included in the RSL Tables.

The EPA Response to CPC Request for Correction

In the Attachments to this response, EPA addresses the assertions raised in the RFC that are relevant to the science evaluation and information presented in the 9,10-AQ PPRTV and the RSL Tables under EPA's IQGs.

In Attachment 1, EPA addresses the following issues as detailed in the CPC RFC:

A. TR 494 presents conclusions which are not scientifically sound and do not comply with EPA's IQGs

1. Mutagenicity testing of Sample A07496
2. Storage of TR 494 test article

B. There is no scientifically sound basis for concluding that non-mutagenic 9,10-AQ caused cancers in NTP TR 494

C. There is no scientifically sound basis for considering non-mutagenic 9,10-AQ likely to be carcinogenic to humans

D. EPA's screening levels for 9,10-AQ do not reflect sound and objective scientific practices

1. The 9,10-AQ PPRTV should be withdrawn
2. 9,10-AQ should be removed from the RSL Tables

In Attachment 2, EPA provides the results of a study quality evaluation of TR 494 conducted in response to this RFC.

Conclusion

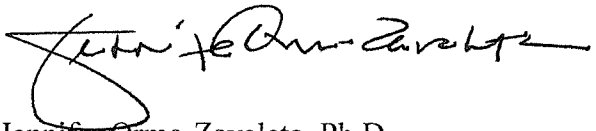
After carefully reviewing the RFC submitted by CPC and conducting a study quality evaluation of TR 494, EPA has concluded that the underlying information and conclusions presented in *Provisional Peer Reviewed Toxicity Values for 9,10-Anthraquinone (CASRN 84-65-1)* and the toxicity values for 9,10-AQ found in the Regional Screening Level Tables are consistent with EPA's IQGs.

Additionally, each of the purported deficiencies in TR 494 detailed in the CPC RFC has been specifically addressed by the National Institute of Environmental Health Sciences (NIEHS) in response to Requests for Correction and Requests for Reconsideration previously submitted by CPC to NIEHS. No new relevant information regarding TR 494 was provided in this RFC. NIEHS has concluded that there is no evidence of noncompliance with National Institutes of Health (NIH) and Health and Human Services (HHS) Information Quality Guidelines for TR 494.

Your Right to Appeal

If you are dissatisfied with the response, you may submit a Request for Reconsideration (RFR) as described in EPA's Information Quality Guidelines. The EPA requests that any such RFR be submitted within 90 days of the date of the EPA's response. If you choose to submit an RFR, please send a written request to the EPA Information Quality Guidelines Processing Staff via mail (Information Quality Guidelines Processing Staff, Mail Code 282IT, USEPA, 1200 Pennsylvania Avenue, NW, Washington, DC 20460); or electronic mail (quality@epa.gov). If you submit a RFR, please reference the case number assigned to this original Request for Correction (RFC #18001). Additional information about how to submit an RFR is listed on the EPA Information Quality Guidelines website at http://epa.gov/quality/informationguidelines/index.html.

Sincerely,



Jennifer Orme-Zavaleta, Ph.D.
Principal Deputy Assistant Administrator for Science
Office of Research and Development

Cc: Vaughn Noga, OEI
Vincia Holloman, OEI
Kevin Kirby, OEI
Tina Bahadori, NCEA
Mary Ross, NCEA
Annette Gatchett, NCEA
Belinda Hawkins, NCEA
Kevin DeBell, AO

Attachment 1: U.S. EPA Response to the Chemical Products Corporation (CPC) Request for Correction (RFC) of *Provisional Peer Reviewed Toxicity Values for 9,10-Anthraquinone (CASRN 84-65-1)* and the Regional Screening Levels for 9,10-Anthraquinone

Attachment 2: Study Quality Evaluation of *NTP Technical Report on the Toxicology and Carcinogenesis Studies of Anthraquinone (CAS No. 84-65-1) in R344/N Rats and B6C3F1 Mice (Feed Studies)* ("TR 494")

Attachment 1

U.S. EPA Response to the Chemical Products Corporation (CPC) Request for Correction (RFC) of *Provisional Peer Reviewed Toxicity Values for 9,10-Anthraquinone (CASRN 84-65-1)* and the Regional Screening Levels for 9,10-Anthraquinone

September 2018

National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency

The Request

The Chemical Products Corporation (CPC) Request for Correction (RFC) requests the 9,10-anthraquinone (9,10-AQ) Provisional Peer Reviewed Toxicity Value (PPRTV) assessment "be immediately withdrawn and revised to provide toxicity values … based upon sound science" and that 9,10-AQ "be immediately removed from EPA's Regional Screening Level Tables" pending revision of the 9,10-AQ PPRTV.

To support the RFC, CPC asserts that EPA "should not consider the conclusions presented in National Toxicology Program Technical Report 494 (TR-494) to represent valid peer-reviewed toxicity values or sound science because peer reviewers were presented false information by NTP staff which prevented the Peer Review Panel from rendering a sound scientific judgement."

Response

In this response, EPA is addressing the assertions raised in the RFC that may be relevant to the derivation and dissemination of EPA toxicity values under EPA's *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility and Integrity of Information Disseminated by the Environmental Protection Agency (IQGs).*

Specifically, EPA is addressing the following topics as raised in the CPC RFC:

A. TR 494 presents conclusions which are not scientifically sound and do not comply with EPA's IQGs

1. Mutagenicity testing of Sample A07496
2. Storage of TR 494 test article

B. There is no scientifically sound basis for concluding that non-mutagenic 9,10-AQ caused cancers in NTP TR 494

C. There is no scientifically sound basis for considering non-mutagenic 9,10-AQ likely to be carcinogenic to humans

D. EPA's screening levels for 9,10-AQ do not reflect sound and objective scientific practices

1. The 9,10-AQ PPRTV should be withdrawn
2. 9,10-AQ should be removed from the RSL Tables

In considering this RFC, EPA reviewed the following:

USEPA (2002). "Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency" (https://www.epa.gov/quality/epa-information-quality-guidelines)

USEPA (2011). "Provisional Peer Reviewed Toxicity Values for 9, 10-Anthraquinone (CASRN 84-65-1)" (https://hhpprtv.ornl.gov/issue_papers/Anthraquinone910.pdf)

NTP (2005). "NTP Technical Report on the Toxicology and Carcinogenesis Studies of Anthraquinone (CAS No. 84-65-1) in F344/N Rats and B6C3F1 Mice (Feed Studies)" [TR 494] (https://ntp.niehs.nih.gov/ntp/htdocs/lt_rpts/tr494.pdf)

HHS (2002). "Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated to the Public" (https://aspe.hhs.gov/report/hhs-guidelines-ensuring-and-maximizing-quality-objectivity-utility-and-integrity-information-disseminated-public)

USEPA (2005). "Guidelines for Carcinogen Risk Assessment" (https://www.epa.gov/risk/guidelines-carcinogen-risk-assessment)

May 21, 1999 NTP Board of Scientific Counselors Summary Minutes from Peer Review of Draft Technical Reports of Long-Term Toxicology and Carcinogenesis Studies by the Technical Reports Review Subcommittee (https://ntp.niehs.nih.gov/ntp/about_ntp/bsc/trrs/1999/may/trrs21may1999mins_508.pdf)

February 17-18, 2004 NTP Board of Scientific Counselors Technical Reports Review Subcommittee Meeting Summary Minutes. (https://ntp.niehs.nih.gov/ntp/about_ntp/bsc/trrs/2004/feb/trrs17feb2004mins_508.pdf)

December 9, 2004 NTP Board of Scientific Counselors Technical Reports Review Subcommittee Meeting Summary Minutes (https://ntp.niehs.nih.gov/ntp/about_ntp/bsc/trrs/2004/dec/trrs9dec2004mins_508.pdf)

Your November 17, 2002 Request for Correction to HHS concerning TR 494 (https://aspe.hhs.gov/information-requests-corrections-and-hhs-responses)

The March 19, 2003 HHS response to your Request for Correction (https://aspe.hhs.gov/information-requests-corrections-and-hhs-responses)

Your March 28, 2003 Request for Reconsideration to HHS (https://aspe.hhs.gov/information-requests-corrections-and-hhs-responses)

The September 8, 2003 HHS response to your Request for Reconsideration (https://aspe.hhs.gov/information-requests-corrections-and-hhs-responses)

Your February 24, 2004 Request for Correction to HHS concerning TR 494 (https://aspe.hhs.gov/information-requests-corrections-and-hhs-responses)

Your March 14, 2005 Request for Correction of Information to HHS concerning TR 494 (https://aspe.hhs.gov/information-requests-corrections-and-hhs-responses)

Your May 31, 2006 Request for Correction to HHS concerning TR 494 and the July 13 and July 17, 2006 Addenda (https://aspe.hhs.gov/information-requests-corrections-and-hhs-responses)

The December 22, 2006 HHS response to your Request for Correction (https://aspe.hhs.gov/information-requests-corrections-and-hhs-responses)

Your January 5, 2007 Request for Reconsideration and the March 1, 2007 Addenda (https://aspe.hhs.gov/information-requests-corrections-and-hhs-responses)

The September 22, 2008 HHS response to your Request for Reconsideration https://aspe.hhs.gov/information-requests-corrections-and-hhs-responses)

Boobis, AR et al. (2009). A Data-Based Assessment of Alternative Strategies for Identification of Potential Human Cancer Hazards. Toxicologic Pathology, 37: 714-732.

Butterworth, BE et al. (2004). Contamination Is a Frequent Confounding Factor in Toxicology Studies with Anthraquinone and Related Compounds. International Journal of Toxicology, 23: 335-344.

Doi, AM et al. (2005). Influence of Functional Group Substitutions on the Carcinogenicity of Anthraquinone in Rats and Mice: Analysis of Long-Term Bioassays by the National Cancer Institute and the National Toxicology Program, Journal of Toxicology and Environmental Health, Part B, 8:2, 109-126.

A. TR 494 presents conclusions which are not scientifically sound and do not comply with EPA's IQGs

The CPC RFC asserts that the National Toxicology Program (NTP) Technical Report (TR 494) does not represent "sound science" and that "peer reviewers were presented false information by NTP staff which prevented the Peer Review Panel from rendering a sound scientific judgement." The RFC contends that NTP staff presented false information to a peer review panel "in order to achieve acceptance of the conclusions presented" in TR 494.

More specifically, the RFC asserts that NTP provided peer reviewers with false mutagenicity testing results of the TR 494 test article and false information regarding storage of the TR 494 test article. These issues are addressed separately below.

1. Mutagenicity testing of Sample A07496

The CPC RFC asserts that: "Someone at NTP arranged for the shipment of "Sample A07496" to BioReliance Testing Laboratories and authorized mutagenicity testing of a sample labeled "Sample A07496" by BioReliance with full knowledge that this AQ sample was not the TR-494 test article." The RFC includes two attachments (Attachments 2 & 3) to support this assertion. It also asserts that "peer reviewers were under the false impression that the TR-494 AQ test article had been determined to be non-mutagenic when they approved the conclusions in TR-494."

HHS specifically addressed the issues raised in the current CPC RFC related to the 9,10-AQ samples tested by BioReliance Corp. in their December 22, 2006 response to your May 31, 2006 Request for Correction (as amended on July 13 and July 17) and in their September 22, 2008 Response to your January 5, 2007 Request for Reconsideration (as amended on March 1, 2007). Attachments 2 & 3 in the current RFC were also submitted as attachments to HHS in your January 5, 2007 HHS Request for Reconsideration (as amended on March 1, 2007) and were addressed by HHS in their responses (detailed below).

In their 2006 response, HHS indicated that: "The Methods and Materials section for TR494 identifies the source of the anthraquinone used in the NTP 2-year studies as lot no. 5893. The NTP conducted follow-up genetic toxicology studies on a sample from lot no. 5893 as well as samples of anthraquinone produced by other processes. Appendix E, which contains the results from these follow-up studies, identifies the source of each sample noting that Sample A07496 is from lot no. 5893. Also, in response to Freedom of Information Act requests by CPC filed on March 28, 2006 and July 19, 2006, the NTP sent

you records documenting shipment of lot no. 5893 to BioReliance Corporation for genetic toxicology testing and verifying Sample A07496 as an aliquot from lot no. 5893 (Enclosures 1 and 2, respectively)."

In their 2008 response, HHS indicated that: "… we reviewed the records related to handling of samples in this matter. Those records indicated that Battelle, the analytical chemistry laboratory, shipped 2 g of 4 different lots of anthraquinone to BioReliance, the study laboratory, on June 1, 2004 … The samples were labeled by lot number, the standard information that is included on samples. BioReliance received the materials on June 2, 2004. You were sent this record previously. Battelle sent NTP the Bulk Chemical Shipment Report dated June 22, 2004 verifying it had shipped samples of the 4 anthraquinone lots to BioReliance on June 1, 2004. This document also was provided by NTP to you in NTP's Response to the Request for Correction. BioReliance confirmed assignment of each test article aliquot number to the correct lot of anthraquinone. You were sent this document previously. The review of these records provided assurance that the samples were handled appropriately and in conformity with routine procedures."

In their responses outlined above, HHS affirmed that the samples were appropriately labeled and that Sample A07496, tested by BioReliance Corp., was an aliquot of the TR 494 test article. In the Genetic Toxicology section, TR 494 states (p246): "Sample A07496 (lot no. 5893) from Zeneca Fine Chemicals … was from the lot used in the 2-year studies …" and that this sample was tested in *Salmonella typhimurium* strains TA98, TA100 and TA1537. TR 494 clearly reports the mutagenicity testing results of Sample A07496 (p248): "Sample A07496, the compound used in the 2-year studies (99.8% pure), was negative in TA98, TA100, and TA1537, with and without 10% and 30% rat S9 at concentrations up to 10,000 µg/plate with both solvents (Table E3)."

Based on the information provided in TR 494, in conjunction with HHS's responses to your prior Requests for Correction and Reconsideration concerning mutagenicity testing of Sample A07496, EPA concludes that there is no evidence of noncompliance with EPA IQGs for TR 494.

2. Storage of TR 494 test article

During the December 9, 2005 meeting of the Board of Scientific Counselors Technical Subcommittee, Dr. Cynthia Smith mistakenly indicated that the 9,10-AQ sample that had been tested for mutagenicity had been taken from an archived sample stored "frozen under argon" when it had instead been taken from archived bulk material stored at room temperature (in an amber glass bottle). The CPC RFC asserts that "… the possibility of decomposition of biologically significant mutagenic impurities in the TR-494 test article over time confounds interpretation of a 2004 negative mutagenicity assay, even if the assay had been performed on an aliquot of the TR-494 test article."

HHS specifically addressed the issue of the 9,10-AQ sample storage and Dr. Smith's statement in their December 22, 2006 Response to your May 31, 2006 Request for Correction (as amended on July 13 and July 17) and in their September 22, 2008 Response to your January 5, 2007 Request for Reconsideration (as amended on March 1, 2007). Specifically, in their 2006 response, HHS agreed to address the misstatement by Dr. Smith by electronic and text erratum and indicated that storage in amber glass bottles at room temperature is the recommended storage conditions for 9,10-AQ. Also in the 2006 response, HHS addressed the issue of possible degradation of impurities. HHS indicated that: "The purity analyses described above and in Appendix J of TR494 were all conducted on aliquots of the anthraquinone test article lot no. 5893 stored at room temperature. Each of these analyses conducted

at different times over a 10-year period gave purity values for the anthraquinone test article lot no. 5893 that are in agreement and do not show evidence of degradation of the bulk test article".

EPA concludes that the misstatement by Dr. Smith had no bearing on the scientific conclusions in TR 494. HHS appropriately corrected this misstatement by a published erratum. The test article was stored according to recommended storage conditions and purity analyses conducted over a ten-year time period showed no evidence of degradation. Importantly, the potential contribution of contaminants to the overall carcinogenicity findings in TR 494 was discussed in detail by the Board of Scientific Counselors Technical Subcommittee. EPA agrees with the conclusions presented by NTP (Dr. Irwin) at the December 9, 2005 Board of Scientific Counselors Technical Subcommittee meeting that "the low exposure levels, bioavailability, and relative mutagenicity make it unlikely that 9-nitroanthracene contributed significantly to the results of the carcinogenicity studies." As such, the "possibility of decomposition" of mutagenic contaminants (specifically 9-nitroanthracene) in Sample A07496 prior to mutagenicity testing does not alter the overall conclusions regarding the carcinogenicity of 9,10-AQ as it is unlikely that the contaminants contributed significantly to the carcinogenic responses reported in TR 494.

In summary, TR 494 was conducted in compliance with Good Laboratory Practice (GLP) regulations as defined by the Food and Drug Administration (FDA), was subjected to quality assurance audits and received appropriate peer review. TR 494 clearly acknowledges the issue of potential contamination of the test article with 9-nitroanthracene and transparently states (p93): "Based on the information currently available, it is not possible to determine to what extent, if any, 9-nitroanthracene influenced the carcinogenic response in the 2-year studies."

Based on the information provided in TR 494, in conjunction with HHS's responses to your prior Requests for Correction and Reconsideration concerning Dr. Smith's misstatement and the potential for degradation of contaminants in the TR 494 test article, EPA concludes that there is no evidence of noncompliance with EPA IQGs for TR 494.

<u>B. There is no scientifically sound basis for concluding that non-mutagenic 9,10-AQ caused cancers in NTP TR 494</u>

As mentioned above, TR 494 was conducted in compliance with Good Laboratory Practice (GLP) regulations as defined by the Food and Drug Administration (FDA), was subjected to quality assurance audits and received appropriate peer review.

The issue of contamination of the 9,10-AQ test article with 9-nitroanthracene was discussed by the Board of Scientific Counselors Technical Subcommittee and was transparently acknowledged in TR 494. In fact, TR 494 clearly states (p92): "The NTP was unable to confirm the bacterial mutagenicity of the anthraquinone used in the NTP studies described in this Technical Report." The TR 494 peer reviewers were provided the mutagenicity testing results of the TR 494 test article and discussed the results in relation to the carcinogenic findings. They agreed that the carcinogenic results reported in TR 494 were valid.

EPA agrees with the conclusions in TR 494 regarding the carcinogenic activity of 9,10-AQ in male and female F344 rats and B6C3F1 mice. EPA also agrees with the following (p93) in TR 494: "Based on the information currently available, it is not possible to determine to what extent, if any, 9-nitroanthracene

influenced the carcinogenic response in the 2-year studies. The anthraquinone tested, greater than 99.8% pure, produced a carcinogenic response consistent with that observed with other anthraquinones. The biotransformation of anthraquinone to mutagenic metabolites with systemic concentrations at least five times greater than is possible for 9-nitroanthracene indicate that anthraquinone is potentially carcinogenic."

The lack of bacterial mutagenicity does not equate to a lack of carcinogenicity *in vivo*. There are several potential modes of action for carcinogenic compounds, mutagenicity is only one. Neither NTP (in TR 494) nor EPA (in the 9,10-AQ PPRTV) asserts that 9,10-AQ induced tumors in TR 494 through a mutagenic mode of action.

EPA concludes that there is no evidence of noncompliance with EPA IQGs for TR 494 or the 9,10-AQ PPRTV.

C. There is no scientifically sound basis for considering non-mutagenic 9,10-AQ likely to be carcinogenic to humans

In determining the cancer Weight of Evidence (WOE) descriptor in the 9,10-AQ PPRTV, EPA reviewed all available information from epidemiological, toxicological and mode of action studies in accordance with EPA's *Guidelines for Carcinogen Risk Assessment*. The 9,10-AQ PPRTV details the information that the descriptor "likely to be carcinogenic to humans" is based on, noting the strengths and weaknesses of the evidence. The information considered in determining the cancer WOE descriptor for 9,10-AQ included information from all lines of evidence and was not "solely from" TR 494 as asserted in the CPC RFC.

The 9,10-AQ PPRTV does not determine a mode of action for 9,10-AQ and clearly states (p26): "The majority of data on 9,10-anthraquinone indicate that 9,10-anthraquinone is not mutagenic."

The evaluation of the carcinogenicity evidence and the conclusion that 9,10-AQ is likely to be carcinogenic to humans in the 9,10-AQ PPRTV was reviewed by independent scientific experts (external peer reviewers) following applicable EPA peer review guidelines. No new scientific information was provided in the CPC RFC that would alter the conclusion in the 9,10-AQ PPRTV that 9,10-AQ is appropriately classified as likely to be carcinogenic to humans.

EPA concludes that there is no evidence of noncompliance with EPA IQGs for the 9,10-AQ PPRTV.

D. EPA's screening levels for 9,10-AQ do not reflect sound and objective scientific practices

The CPC RFC asserts that the 9,10-AQ PPRTV does not reflect "sound and objective scientific practices" because the assessment relies on TR 494 in deriving toxicity values. The CPC RFC specifically requests that the 9,10-AQ PPRTV be withdrawn and that 9,10-AQ be removed from the RSL Tables. The information for 9,10-AQ in the RSL Tables was taken directly from the 9,10-AQ PPRTV. As such, these two issues are interrelated and will be discussed together below.

1. The 9,10-AQ PPRTV should be withdrawn

2. 9,10-AQ should be removed from the RSL Tables

The subchronic and chronic provisional screening Reference Doses (p-RfDs) derived in the 9,10-AQ PPRTV are based on noncancer adverse effects, not carcinogenic (or mutagenic) effects. The issues raised in the CPC RFC concerning the mutagenicity/carcinogenicity findings in TR 494 are not relevant to

the derivation of noncancer toxicity values in the 9,10-AQ PPRTV. There is no new information provided in the CPC RFC that is relevant to the noncancer subchronic or chronic screening p-RfDs derived in the 9,10-AQ PPRTV.

The 9,10-AQ PPRTV was developed following all applicable EPA guidelines. Following public release of the 9,10-AQ PPRTV, the 9,10-AQ toxicity values derived in the PPRTV were included in the RSL Tables.

EPA acknowledged the issue of 9-nitroanthracene (9-NA) contamination of the 9,10-AQ test article utilized in TR 494 within the PPRTV and summarized the NTP Board of Scientific Counselors Technical Review Subcommittee findings regarding the issue (see pp 8, 25 & 26 of the 9,10-AQ PPRTV). This information was reviewed by independent scientific experts (external peer reviewers) following all applicable EPA peer review guidelines.

In response to this RFC, EPA conducted an additional evaluation of TR 494 using standardized study quality evaluation criteria (see Attachment 2). This evaluation resulted in a determination of "High Confidence" for TR 494. Based on this study quality evaluation, EPA again concludes that there is no evidence of noncompliance with EPA's IQGs for TR 494.

EPA concludes that there is no evidence of noncompliance with EPA's IQGs for the 9,10-AQ PPRTV or the 9,10-AQ toxicity values in the RSL Tables.

Conclusion

EPA, after careful review of the RFC submitted by CPC, has concluded that the underlying information and conclusions presented in *Provisional Peer Reviewed Toxicity Values for 9,10-Anthraquinone (CASRN 84-65-1)* and in the RSL Tables are consistent with EPA's IQGs.

Attachment 2

Study Quality Evaluation of *NTP Technical Report on the Toxicology and Carcinogenesis Studies of Anthraquinone (CAS No. 84-65-1) in R344/N Rats and B6C3F1 Mice (Feed Studies)* ("TR 494")

September 2018

National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency

Study Quality Evaluation of TR 494

The Chemical Products Corporation (CPC) Request for Correction (RFC) asserts that *NTP Technical Report on the Toxicology and Carcinogenesis Studies of Anthraquinone (CAS No. 84-65-1) in R344/N Rats and B6C3F1 Mice (Feed Studies)* (herein referred to as "TR 494") does not represent "sound science." In response to this RFC, EPA conducted a study quality evaluation of TR 494 to assess risk of bias and sensitivity. The results of the study quality evaluation of TR 494 are shown in Figure 2-1.

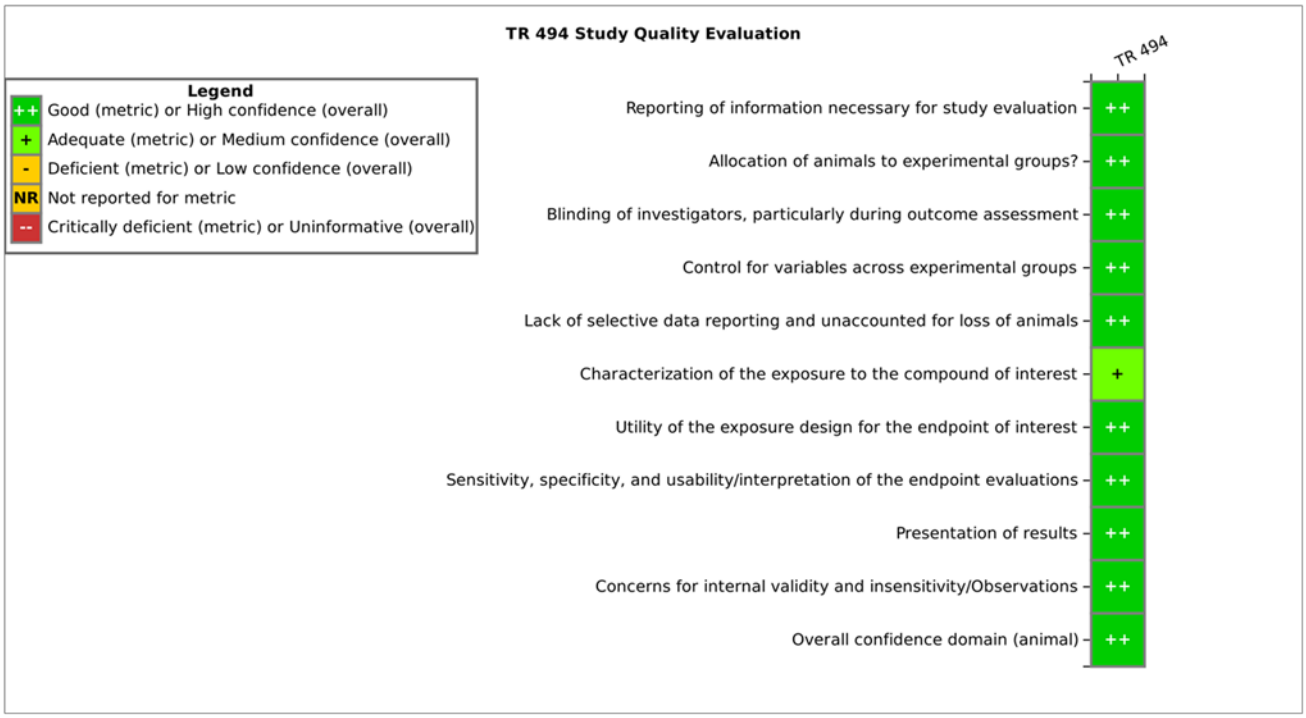
Key issues during this evaluation were potential bias (factors that affect the magnitude or direction of an effect) and insensitivity (factors that limit the ability of a study to detect a true effect). The study quality evaluation of TR 494 was conducted for the following study domains: reporting quality, selection or performance bias, confounding/variable control, reporting or attrition bias, exposure methods sensitivity, and outcome measures and results display (see Table 2-1).

All study domains were judged to be *Good* with the exception of "Characterization of the exposure to the compound of interest" which was judged to be *Adequate*. This domain was judged to be *Adequate* due to uncertainty related to the impact, if any, of the impurities in the TR 494 test compound. The percent purity of the test compound was well documented, the impurities detected were reported and the issue of contamination was discussed within TR 494 and as part of the TR 494 peer review process. EPA agrees with the conclusion of NTP (Dr. Irwin) (TR 494; p18) that: "The low exposure levels, bioavailability, and relative mutagenicity make it unlikely that 9-nitroanthracene contributed significantly to the results of the carcinogenicity studies."

The study quality evaluation of TR 494 results in an overall study quality classification of "High Confidence" (see "Study Quality Evaluation Methodology" below).

EPA concludes that TR 494 is an appropriate study for use in the derivation of toxicity values for 9,10-AQ.

Figure 2-1. Study quality evaluation results for TR 494



Study Quality Evaluation Methodology

The study quality evaluation of TR 494 was conducted on the following domains: reporting quality, selection or performance bias, confounding/variable control, reporting or attrition bias, exposure methods sensitivity, and outcome measures and results display (see Table 2-1).

For each study domain, a judgment of *Good*, *Adequate*, *Deficient*, *Not Reported* or *Critically deficient* was made. These five categories were applied to each evaluation domain as follows:

- *Good* represents a judgment that the study was conducted appropriately in relation to the evaluation domain, and any minor deficiencies that are noted would not be expected to influence the study results.

- *Adequate* indicates a judgment that there are methodological limitations relating to the evaluation domain, but that those limitations are not likely to be severe or to have a notable impact on the results.

• *Deficient* denotes identified biases or deficiencies that are interpreted as likely to have had a notable impact on the results or that prevent reliable interpretation of the study findings.

• *Not reported* indicates that the information necessary to evaluate the domain question was not available in the study. Generally, this term carries the same functional interpretation as *Deficient* for the purposes of the study confidence classification (described below).

• *Critically deficient* reflects a judgment that the study conduct relating to the evaluation domain question introduced a serious flaw that is the primary driver of any observed effect(s) or makes the study uninterpretable.

Table 2-1. Domains of study quality evaluation for TR 494

Evaluation type		Domain – Core question	Prompting questions	Basic Considerations
Reporting Quality		**Reporting Quality** – Does the study report information for evaluating the design and conduct of the study for the endpoint(s)/outcome(s) of interest? *Notes:* *This domain is limited to reporting. Other aspects of the exposure methods, experimental design, and endpoint evaluation methods are evaluated using the domains related to risk of bias and study sensitivity.*	Does the study report the following? • **Critical information** necessary to perform study evaluation: o Species; test article name; levels and duration of exposure; route (e.g., oral; inhalation); qualitative or quantitative results for at least one endpoint of interest • **Important information** for evaluating the study methods: o Test animal: strain, sex, source, and general husbandry procedures o Exposure methods: source, purity, method of administration o Experimental design: frequency of exposure, animal age and lifestage during exposure and at endpoint/outcome evaluation o Endpoint evaluation methods: assays or procedures used to measure the endpoints/outcomes of interest	• *Good*: All **critical and important information** is reported or inferable for the endpoints/outcomes of interest. • *Adequate*: All **critical information** is reported but some **important information** is missing. However, the missing information is not expected to significantly impact the study evaluation. • *Deficient*: All **critical information** is reported but **important information** is missing that is expected to significantly reduce the ability to evaluate the study. • *Critically Deficient*: Study report is missing any pieces of **critical information**.
Risk of Bias	**Selection and performance bias**	**Allocation** – Were animals assigned to experimental groups using a method that minimizes selection bias?	For each study: • Did each animal or litter have an equal chance of being assigned to any experimental group (i.e., random allocation)? • Is the allocation method described? • Aside from randomization, were any steps taken to	A judgment and rationale for this domain should be given for each cohort or experiment in the study. • *Good*: Experimental groups were randomized and any specific randomization procedure was described or inferable (e.g., computer-generated scheme). [Note that normalization is not the same as

Evaluation type		Domain – Core question	Prompting questions	Basic Considerations
			balance variables across experimental groups during allocation?	randomization (see response for '*Adequate*').] • *Adequate*: Authors report that groups were randomized but do not describe the specific procedure used (e.g., "animals were randomized"). Alternatively, authors used a non-random method to control for important modifying factors across experimental groups (e.g., body weight normalization). • *Not Reported* (interpreted as *Deficient*): No indication of randomization of groups or other methods (e.g., normalization) to control for important modifying factors across experimental groups. • *Critically Deficient*: Bias in the animal allocations was reported or inferable.
		Observational bias/Blinding – Did the study implement measures to reduce observational bias?	For each endpoint/outcome or grouping of endpoints/outcomes in a study: • Does the study report blinding or other methods/procedures for reducing observational bias? • If not, did the study use a design or approach for which such procedures can be inferred? • What is the expected impact of failure to implement (or report implementation) of these methods/procedures on results?	A judgment and rationale for this domain should be given for each endpoint/outcome or group of endpoints/outcomes investigated in the study. • *Good*: Measures to reduce observational bias were described (e.g. blinding to conceal treatment groups during endpoint evaluation; consensus-based evaluations of histopathology lesions[1]). • *Adequate*: Methods for reducing observational bias (e.g., blinding) can be inferred or were reported but described incompletely. • *Not Reported*: Measures to reduce observational bias were not described. o (interpreted as *Adequate*) The potential concern for bias was

[1] For non-targeted or screening-level histopathology outcomes often used in guideline studies, blinding during the initial evaluation of tissues is generally not recommended as masked evaluation can make "the task of separating treatment-related changes from normal variation more difficult" and "there is concern that masked review during the initial evaluation may result in missing subtle lesions." Generally, blinded evaluations are recommended for targeted secondary review of specific tissues or in instances when there is a pre-defined set of outcomes that is known or predicted to occur (Crissman et al., 2004). Crissman et al. (2004). Best practices guideline: toxicologic histopathology. Toxicol Pathol. Jan-Feb;32(1):126-31.

Evaluation type	Domain – Core question	Prompting questions	Basic Considerations
			mitigated based on use of automated/computer driven systems, standard laboratory kits, relatively simple, objective measures (e.g., body or tissue weight), or screening-level evaluations of histopathology. o (interpreted as *Deficient*) The potential impact on the results is major (e.g., outcome measures are highly subjective). • *Critically Deficient*: Strong evidence for observational bias that could have impacted results
Confounding/ variable control	**Confounding –** Are variables with the potential to confound or modify results controlled for and consistent across all experimental groups?	For each study: • Are there differences across the treatment groups (e.g., co-exposures, vehicle, diet, palatability, husbandry, health status, etc.) that could bias the results? • If differences are identified, to what extent are they expected to impact the results?	A judgment and rationale for this domain should be given for each cohort or experiment in the study, noting when the potential for confounding is restricted to specific endpoints/outcomes. • *Good*: Outside of the exposure of interest, variables that are likely to confound or modify results appear to be controlled for and consistent across experimental groups. • *Adequate*: Some concern that variables that were likely to confound or modify results were uncontrolled or inconsistent across groups, but are expected to have a minimal impact on the results. • *Deficient*: Notable concern that potentially confounding variables were uncontrolled or inconsistent across groups, and are expected to substantially impact the results. • *Critically deficient*: Confounding variables were presumed to be uncontrolled or inconsistent across groups, and are expected to be a primary driver of the results.
Reporting and attrition bias	**Selective reporting and attrition –** Did the study report results for all	For each study:	A judgment and rationale for this domain should be given for each cohort or experiment in the study. • *Good*: Quantitative or qualitative results were reported for all prespecified outcomes (explicitly

Evaluation type		Domain – Core question	Prompting questions	Basic Considerations
		prespecified outcomes and tested animals? *Note:* *This domain does **not** consider the appropriateness of the analysis/results presentation. This aspect of study quality is evaluated in another domain.*	*Selective reporting bias:* • Are all results presented for endpoints/outcomes described in the methods (see note)? *Attrition bias:* • Are all animals accounted for in the results? • If there are discrepancies, do authors provide an explanation (e.g., death or unscheduled sacrifice during the study)? • If unexplained results omissions and/or attrition are identified, what is the expected impact on the interpretation of the results?	stated or inferred), exposure groups and evaluation timepoints. Data not reported in the primary article is available from supplemental material. If results omissions or animal attrition are identified, the authors provide an explanation and these are not expected to impact the interpretation of the results. • *Adequate*: Quantitative or qualitative results are reported for most prespecified outcomes (explicitly stated or inferred), exposure groups and evaluation timepoints. Omissions and/or attrition are not explained, but are not expected to significantly impact the interpretation of the results. • *Deficient*: Quantitative or qualitative results are missing for many prespecified outcomes (explicitly stated or inferred), exposure groups and evaluation timepoints and/or high animal attrition; omissions and/or attrition are not explained and may significantly impact the interpretation of the results. • *Critically Deficient*: Extensive results omission and/or animal attrition are identified and prevents comparisons of results across treatment groups.
Sensitivity	**Exposure methods sensitivity**	**Chemical administration and characterization –** Did the study adequately characterize exposure to the chemical of interest and the exposure administration methods?	For each study: • Does the study report the source and purity and/or composition (e.g., identity and percent distribution of different isomers) of the chemical? If not, can the purity and/or composition be obtained from the supplier (e.g., as reported on the website) • Was independent analytical verification of the test article	A judgment and rationale for this domain should be given for each cohort or experiment in the study. • *Good*: Chemical administration and characterization is complete (i.e., source, purity, and analytical verification of the test article are provided). There are no concerns about the composition, stability, or purity of the administered chemical, or the specific methods of administration. For inhalation studies, chemical concentrations in the exposure chambers are verified using reliable analytical methods. • *Adequate*: Some uncertainties in the chemical administration and characterization are identified but

Evaluation type		Domain – Core question	Prompting questions	Basic Considerations

			purity and composition performed?
			• Did the authors take steps to ensure the reported exposure levels were accurate?
			  o For inhalation studies: were target concentrations confirmed using reliable analytical measurements in chamber air?
			  o For oral studies: if necessary based on consideration of chemical-specific knowledge (e.g., instability in solution; volatility) and/or exposure design (e.g., the frequency and duration of exposure), were chemical concentrations in the dosing solutions or diet analytically confirmed?
			• Are there concerns about the methods used to administer the chemical (e.g., inhalation chamber type, gavage volume, etc.)?
				these are expected to have minimal impact on interpretation of the results (e.g., source and vendor- reported purity are presented, but not independently verified; purity of the test article is sub-optimal but not concerning; For inhalation studies, actual exposure concentrations are missing or verified with less reliable methods).
				• *Deficient*: Uncertainties in the exposure characterization are identified and expected to substantially impact the results (e.g., source of the test article is not reported; levels of impurities are substantial or concerning; deficient administration methods, such as use of static inhalation chambers or a gavage volume considered too large for the species and/or lifestage at exposure).
				• *Critically Deficient*: Uncertainties in the exposure characterization are identified and there is reasonable certainty that the results are largely attributable to factors other than exposure to the chemical of interest (e.g., identified impurities are expected to be a primary driver of the results).

Evaluation type		Domain – Core question	Prompting questions	Basic Considerations
		Exposure timing, frequency and duration – Was the was the timing, frequency, and duration of exposure sensitive for the endpoint(s)/outcome(s) of interest?	For each endpoint/outcome or grouping of endpoints/outcomes in a study: • Does the exposure period include the critical window of sensitivity? • Was the duration and frequency of exposure sensitive for detecting the endpoint of interest?	A judgment and rationale for this domain should be given for each endpoint/outcome or group of endpoints/outcomes investigated in the study. • *Good*: The duration and frequency of the exposure was sensitive and the exposure included the critical window of sensitivity (if known). • *Adequate*: The duration and frequency of the exposure was sensitive and the exposure covered most of the critical window of sensitivity (if known). • *Deficient*: The duration and/or frequency of the exposure is not sensitive and did not include the majority of the critical window of sensitivity (if known). These limitations are expected to bias the results towards the null. • *Critically deficient*: The exposure design was not sensitive and is expected to strongly bias the results towards the null. The rationale should indicate the specific concern(s).
	Outcome measures and results display	**Endpoint sensitivity and specificity –** Are the procedures sensitive and specific for evaluating the endpoint(s)/outcome(s) of interest? *Note:* *Sample size alone is not a reason to conclude an individual study is critically deficient.*	For each endpoint/outcome or grouping of endpoints/outcomes in a study: • Are there concerns regarding the specificity and validity of the protocols? • Are there serious concerns regarding the sample size (see note)? • Are there concerns regarding the timing of the endpoint assessment?	A judgment and rationale for this domain should be given for each endpoint/outcome or group of endpoints/outcomes investigated in the study. Examples of potential concerns include: • Selection of protocols that are insensitive or non-specific for the endpoint of interest • Use of unreliable methods to assess the outcome • Assessment of endpoints at inappropriate or insensitive ages, or without addressing known endpoint variation (e.g., due to circadian rhythms, estrous cyclicity, etc.). • Decreased specificity or sensitivity of the response due to the timing of endpoint evaluation, as compared to exposure (e.g., short-acting depressant or irritant effects of chemicals; insensitivity due to

Evaluation type	Domain – Core question	Prompting questions	Basic Considerations
			prolonged period of non-exposure prior to testing).
	Results Presentation – Are the results presented in a way that makes the data usable and transparent?	For each endpoint/outcome or grouping of endpoints/outcomes in a study: • Does the level of detail allow for an informed interpretation of the results? • Are the data analyzed, compared, or presented in a way that is inappropriate or misleading?	A judgment and rationale for this domain should be given for each endpoint/outcome or group of endpoints/outcomes investigated in the study. Examples of potential concerns include: • Non-preferred presentation, such as developmental toxicity data averaged across pups in a treatment group, when litter responses are more appropriate • Failing to present quantitative results • Pooling data when responses are known or expected to differ substantially (e.g., across sexes or ages) • Failing to report on or address overt toxicity when exposure levels are known or expected to be highly toxic • Lack of full presentation of the data (e.g., presentation of mean without variance data; concurrent control data are not presented)

Once the evaluation domains were considered, an overall study confidence classification for TR 494 was made. This classification was based on the judgments across the evaluation domains and included consideration of the likely impact of the noted deficiencies in bias and sensitivity, or inadequate reporting, on the results. The overall study confidence classifications are defined as follows:

• *High confidence*: No notable deficiencies or concerns were identified; the potential for bias is unlikely or minimal, and the study used sensitive methodologies. In general, although classifications are not decided by "scoring," *High confidence* studies would reflect judgments of *Good* across all or most evaluation domains.

• *Medium confidence*: Possible deficiencies or concerns were noted, but the limitations are unlikely to be of a notable degree. Generally, *Medium confidence* studies will include *Adequate* or *Good* judgments across most domains, with the impact of any identified limitation not being

judged as severe.

• *Low confidence*: Deficiencies or concerns were noted, and the potential for substantive bias or inadequate sensitivity could have a significant impact on the study results or their interpretation. Typically*, Low confidence* studies would have a *Deficient* evaluation for one or more domains (unless the impact of the limitations on the results is judged as unlikely to be severe).

• *Uninformative*: Serious flaw(s) make the study results unusable for informing hazard identification. Studies with *Critically deficient* judgements in any evaluation domain will almost always be classified as *Uninformative* (see explanation above). Studies with multiple *Deficient* judgments across domains may also be considered *Uninformative*, particularly when there is a robust database of studies on the outcome(s) of interest or when the impact of the limitations is viewed as severe.