American Petroleum Institute

1220 L Street, Northwest
Washington, DC 20005-4070
Tel 202-682-8344
Fax 202-682-8270
E-mail twerdokl@api.org

Lorraine E. Twerdok, Ph. D., DABT
Manager, Health Sciences

June 28, 2004

Suzanne Rudzinski
Director, Transportation and Regional Programs Division
U.S. EPA Headquarters
6406J
Ariel Rios Building
1200 Pennsylvania Avenue NW
Washington, D.C. 20460

RE: Section 211(b) Fuels and Fuel Additives Health Effects Testing Regulation: Peer Reviewer Comments on Draft Baseline Gasoline + MTBE Vapor Condensate Two-Generation Reproductive Toxicity Study Report

Dear Ms. Rudzinski:

The American Petroleum Institute (API), on behalf of the Section 211(b) Research Group (RG), is submitting independent peer reviewer comments and the RG response to those comments as required in the Alternative Tier 2 provisions of the testing regulation.

In accordance with the Alternative Tier 2 provisions, the RG has submitted the draft final report entitled *Baseline Gasoline MTBE Vapor Condensate: A Two-Generation Whole-Body Inhalation Reproductive Toxicity Study In Rats* to EPA for review and comment. This transmittal of independent reviewer comments and the RG response completes the information package for the EPA review of the MTBE draft reproductive report. Once the RG has received EPA's comments, the draft report will be finalized with incorporation of Independent Reviewer comments as indicated, as well as EPA's comments.

When the final report of the MTBE Reproductive Study is complete, it will be submitted to you at the EPA Office of Transportation and Air Quality, Transportation and Regional Programs Division, as part of the requirements of Clean Air Act Section 211(b)(2) and 211(e) (Docket No. A-90-07). If you require further information, please contact Lorraine Twerdok at 202-682-8344, or by mail at this address.

Regards,

Lorraine E. Twerdok

Lorraine Twerdok, Ph.D., DABT
Administrator, 211(b) Research Group

Encl (3): Comments from Dr. Thomas Goldsworthy
Comments from Dr. Richard Schlesinger
RG response to comments

Cc: Joe Sopata, EPA (via email)
Monica Alvarez, EPA (via email)
Mike Davis, EPA (via email)
Tom Goldsworthy (via email)
Rich Schlesinger (via email)
211b RG Oversight and Technical Committees

Section 211 (b) Research Group Response to Independent Peer reviewers Comments on:

Baseline Gasoline + MTBE Vapor Condensate: A Two- Generation Whole-Body Inhalation Reproductive Toxicity Study In Rats

Provisions of the Clean Air Act, Section 211(b) Testing Program call for review of all draft final reports by two independent peer reviewers. These reviewers are Dr. Thomas Goldsworthy, Consultant, and Dr. Richard Schlesinger, Pace University. The present submission by the Research Group (RG) is the second set of comments on a series of reproductive toxicity studies on baseline gasoline and gasoline blended with oxygenates. The reproductive studies included in the 211 (b) program consist of two-generation studies on vapor condensates of baseline gasoline and gasoline containing MTBE as well as one-generation studies on gasoline containing TAME, ETBE, DIPE, ethanol and TBA.

In the present two-generation study on baseline gasoline + MTBE, rats were exposed by inhalation through a ten-week pre-mating period, during mating and gestation, and through lactation and maturation of the F1 and F2 generation offspring. Standard measures of reproductive function were made. In addition, selected F1 pups (Day 28) were sacrificed for neuropathological evaluations and determination of brain GFAP levels. The results indicated that, in general, there was no effect of baseline gasoline + MTBE vapor condensate on reproductive performance and no evidence of neurotoxicity. Discussion of other specific findings is contained in the report.

Overall, the reviewers concurred that the study was conducted in a "scientifically sound manner" and agreed with the report's overall conclusions. Each reviewer provided additional specific comments, which are summarized below, along with the 211 (b) Research Group's response to those comments and proposed follow up course of action, where appropriate.

Dr. Schlesinger:

- Pointed out inconsistency in numbering of exposure groups in different sections of the report – Roman numerals in some tables, 1 2, 3, 4 in others.

  RG Response: This inconsistency will be brought to the attention of the study director and, where practical, adjustments can be made. However, in computer generated tables and compilation of raw data the number format is predetermined. This observation does not affect the validity of the data.

- Statistical Analysis: [p. 37. 2.14.1]
  A number of questions were raised about specific statistical tests and procedures that were used. These included the following: a) which tests were used when significant differences among the means were found, b) what tests were conducted at two levels of significance, and provide justification for use of two levels, c) statistical evaluations were not performed when the SD for the control group was 0, other tests could have been performed.

  RG Response:
  a) Only 1 test (the Dunnetts t test) was performed to compare the group means from the control mean. The report will be thus clarified.
  b) Only 1 test (the Dunnetts t test) was conducted at both the 5% and 1% significance levels. Although not stated in the protocol, these levels of significance were preselected since they are inherent in the laboratory's computer statistical package. Comparison at 2 levels of significance has been common practice for this laboratory.
  c) The standard deviation for an endpoint would be 0 if the n=1, or all of the individual values were the same. This is a standard statement for the statistical methods section. These events occur infrequently, usually in macro/micro pathology observations, and it was not necessary to consider alternative statistics.

- Statistical Analysis [last paragraph]
  The last sentence of this section seems out of place in a discussion of the statistics of sperm and ovary data.

  RG response:

The last sentence to which Dr. Schlesinger refers, deals with macro/microscopy of the testes. The RG agrees that this sentence would be better placed elsewhere, specifically in section 2.13.5 Sperm counts, motility & morphology. The study director will be requested to move this sentence.

- Protocol deviations: [p. 40. 2.17]
  Concern was expressed about the number of "technician-related errors" as noted in Section 2.17 (Protocol Deviations) of the report. It was suggested that this needed to be addressed with the study sponsor. Despite this concern, however, Dr. Schlesinger also concluded, "there were no significant deviations from these protocols that would have affected the outcome of the study."

  RG Response:
  Nearly all of the protocol deviations were related to isolated errors in data collection. None significantly impacted the overall results or conclusions of the study. While the absolute number of protocol deviations may seem excessive, a Two-Generation Reproduction study is a considerable undertaking and involves a very large number of animals. The fact that these protocol deviations were noted and documented reflects positively on the Quality Assurance (QA) practices that have been put in place by API, and at the laboratory. Nonetheless, the RG will bring this issue to the attention of the laboratory and will strive to minimize future protocol deviations.

- Chamber Monitoring: [p. 42. 3.1]
  Note was made that the report referred to a slight inaccuracy in the calibration of the IR monitor (pg 42), and a view was expressed by Dr. Schlesinger that this would be an unacceptable source of experimental error.

  In addition, comment was made about the presence of particles in all of the chambers. Questions were raised about the source of these particles and why filters were not fitted to the chambers to remove these particles.

  RG Response:
  To be precise, the report stated that the ratio of the measured to nominal concentration was not 1:1, and speculated that the cause **MAY** have been due calibration error. In any event, the inaccuracies are believed to be slight. The RG feels confident that the reported chamber concentrations accurately reflect the atmospheres to which the animals were exposed.

  With regard to the presence of particles, the RG notes that the mean particle concentrations were comparable between the control and treatment group chambers. The primary purpose of making these measurements in a study of this nature (high vapor concentrations) is to make certain that vapor condensation is not occurring and to preclude the possibility of aerosol formation. The results provided in the report confirm that gasoline exposure was to vapors and not aerosols. Some particles are invariably present in chamber atmospheres, representing background levels. In fact, careful analysis of the data in the table on page 45 indicates that the particle concentrations in all of the chambers ranged from about 3 to 4 micrograms per cubic meter of air. For comparison, it has been reported in EPA's PM 10 Criteria document that annual background PM 10 levels in eastern cities range from 5 to 11 micrograms / cubic meter of air. Thus, the RG believes that the particles in the chambers reflect ambient background values.

Results

- Feed Consumption [p.45. 3.2.4, Table 11]
  Dr. Schlesinger noted that differences in feed consumption in F1 animals during the first 7 weeks premating "appear to be just as evident in females as in males, although the comment in the text states otherwise".

  RG Response:
  The report states that exposure-related decreases in feed consumption during the initial 7 weeks premating occurred in F1 animals of both sexes but was more evident in males. In comparing these data, although absolute consumption in treated animals of both sexes was similar, the female controls consumed more than male controls, reducing the statistically significant differences with treated groups. While 10000 and 20000mg/m^3 F1 male feed consumption was significantly decreased during all 7 weeks, statistically significant decreases was seen in females

exposed to 20000mg/m^3 only at wks 1, 3 but more frequently at 10000mg/m^3 in wk 1, 3-6. . Given the variability of these data, the RG accepts the study director's evaluation of this parameter.

- F1 Maternal Lactation Body Weight [p. 46. 3.2.7, Table 9]
  It was pointed out that the statistically significant decrease in maternal lactation body weight at 20000mg/m^3 on days 21 and 28 ($p<0.05$) as seen in Table 9 was not addressed in the report text.

  RG Response:
  Dr. Schlesinger is correct. The RG will inquire why this information was not included in the report and request the study director to report these data and address the significance in the text.

- Macroscopic Postmortem evaluations [p.48. 3.4.1, Table 20]
  It was noted that discolored foci were observed in the lungs of P0 and F1 animals of both sexes exposed to gasoline+MTBE vapor, and suggested that no statistical tests were performed because incidence in control animals was zero.

  RG Response:
  Control (P0 and F1) and 20000mg/m^3 P0 males had 0 incidence of discolored foci. The RG accepts the discussion of the discolored foci in treated animals, based on observed incidence, as sufficient to address the occurrence of these lesions without employing statistical analysis. Histological identification of alveolar macrophage aggregates in P0 and F1 rats correlated with occurrence of discolored foci. The increased frequency of this background lesion was plausibly related to test substance exposure but was considered of minor clinical significance by the attending pathologist.

- Organ weights: [p. 49. 3.4.2, Table 22
  Dr. Schlesinger disagrees with the study director's conclusion that decreased lung weight in F1 males "did not occur in a treatment-related pattern" and suggested that since, 'the microscopic findings (cited above) would have resulted in increased, rather than decreased, lung weights", an explanation addressing the actual effect should be considered.

  RG Response:
  As indicated by Dr. Schlesinger, the absolute lung wt in F1 males did decrease in a statistically significant treatment-related manner. However, when lung wts were compared relative to final body wt, only the 10,000mg/m^3 relative wt was statistically significantly lower compared to controls, indicating that the organ wt changes related to body wt did not occur in a treatment –related manner. Furthermore decreased lung wt were not seen in P0 animals of either sex or in F1 females. The incidence of alveolar macrophage aggregates seen in lungs of F1 males was generally lower than for other groups. There does not seem to be a substantive linkage between decreased lung weights in F1 males and discolored foci or alveolar microphage aggregates, leaving the clinical significance of these lower weights indefinable. The RG agrees with the study conclusions but will discuss appropriate expansion to the presentation of these data with the study director.

- Sperm [p. 49-50. 3.4.3]
  Dr. Schlesinger disagrees with the study conclusions that the increase in #sperm/g epididymis and increased % abnormal sperm in the P0 generation were not biologically significant, although statistically significant compared to concurrent controls, because the values fall within the range of historical controls. In his opinion, only concurrent controls are appropriate in evaluating the significance of data.

  RG Response:
  The biological significance of increased sperm/g epididymis and increased % abnormal sperm can be better evaluated when actual increases are presented in report text. Although the increased sperm count in P0 males is 15% higher than controls (891.2 vs. 753.6 million sperm/g epididymis), this value is similar to control rats in F1 generation (887.5) and the historical control average at Pathology Associates Inc (PAI – 869.9). The statistical significance is produced by the low value for control sperm in the P0 generation as is demonstrated in the study itself and not by historical control data alone. The increased % abnormal sperm at 20000mg/m^3 P0 rats is only 0.7% greater than controls (2% vs. 1.3%) and although statistically significant, falls within the historical range at PAI (0-6.7% in control animals).

In cases of statistically significant small changes at single doses or changes which appear statistically significant because control values are unusually high or low, it is appropriate to consider these events in relation to historical data to determine biological/clinical significance. Only a small sub-set of a population of animals is used in an individual study, and where changes are minimal, that data must be evaluated in a broader context. This approach is similar to evaluating results from medical tests in humans where the individual's values are considered in relation to the population range for that endpoint as well as the patient's baseline value if available from earlier tests. Interestingly, Dr. Goldworthy also addressed these data and suggested that the historical control values be included in the text (see comments below).
The RG accepts the study director's evaluation of these male reproductive endpoints but will direct that the control data be included in the report text as well as in Appendix CC.

- GFAP Assay
- Dr. Schlesinger noted that only one level of significance was used in this study compared to the main study where two levels were used and inquired why a two way ANOVA that included gender, as a factor was not used in place of a one-way ANOVA.

  RG Response:
  The GFAP study was performed by Dr. O'Callaghan, a different investigator, from the main study and selection of statistical methods conformed to those used in his laboratory. The RG will request Dr. O'Callahan to clarify the statistical choices.

  Suggests that in Table 5 & 6, the notation * for footnote should not be used since it is easily confused with the notation for statistical differences.

  RG Response
  This recommendation will be relayed to Dr. O'Callaghan.

Dr. Goldsworthy:

- Temperature and Humidity
  Noted that there was wide variation in temperature (up to 31.3^0C) and humidity (2-98% range) and recommended that additional comments address these variations from reported average values of 24^0C and 44% relative humidity [p. 42].

  RG Response:
  The RG will discuss appropriate addition to the text with the study director

- Organ weights: [p. 49. 3.4.2, Table 22]
  Dr. Goldsworthy suggested that kidney weights were increased in both male and female rats. He acknowledged that the increase in male kidney wt occurred with light hydrocarbon nephropathy but considered the text vague in presenting microscopic findings for female rats and requested clearer statements concerning the presence or absence of kidney lesions in females.

  RG Response:
  First, in Table 22, female absolute kidney wts were not statistically significantly higher than controls for P0 or F1 parental animals. Only the organ/body wt ratio for F1 parental females exposed to 20000mg/m^3 was significantly increased at $p< 0.05$.
  Secondly, Tables on p. 51 (P0 rats) and p. 52 (F1 parental rats) show that no histological findings were observed in female kidneys of either generation at 20000mg/m^3 level of exposure.
  No adjustments to the text are needed.

- Sperm [p. 49-50. 3.4.3]
  Dr. Goldworthy did not express concern about the use of historical control data in assessing the biological significance of sperm changes in treated rats but did recommend that the report text should include PAI's historical control values.

RG Response
Historical control values from PAI are supplied in Appendix CC. The RG will instruct the study director to include these values in the report text.

- GFAP Data
  Dr. Goldsworthy agreed that exposures to baseline gasoline + MTBE vapor condensate did not increase GFAP levels in either males or females. He also added that editorial comments made about the GFAP section from previous subchronic toxicity reports apply here as well. (Previous comments included the need for more information on statistical methodology, and the observation that the report format was unusual).

  RG Response:
  Clarification of statistical methods as requested for previous reports will be incorporated. In earlier comments about the subchronic study on baseline gasoline, EPA agreed that there were no major technical problems with the GFAP section of the report (including format), but that a request for additional statistical information was reasonable.

- Text / Editorial
  GFAP section in Main study [p. 37. 2.13.7]
  Request definition of the "spirit of GLP" and how it is assessed.

  RG Response:
  Dr. O'Callaghan is one of the most experienced and well recognized investigators performing the GFAP assay but his laboratory is research oriented and is working to fully comply with all GLP record-keeping requirements as stated in EPA guidelines. The spirit of GLP is defined as collection and maintenance of accurate data as is expected of any careful investigator with appropriate review and documentation. Inspections by the internal CDC QA as well as the RG QA expert have been performed and dates of inspections and data review will be provided in the final report.

  Conclusion [p.54]
  Dr Goldsworthy pointed out that a sentence summarizing the reproductive findings should be added prior to the final statement of the reproductive NOEL value.

  RG Response.
  The RG agrees with the need for a summary statement indicating that "No effect on reproductive performance was observed for animals exposed to gasoline+MTBE vapor condensate in either generation", to be followed by the final sentence that the reproductive NOEL is 20000mg/m^3.