DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
Food and Drug Administration

Memorandum

Date: August 27, 2007

From: David Hattan, Ph.D.
Senior Toxicologist, OFAS

Subject: Current Aspartame Safety Assessment Based On Evaluation of Recent Toxicological Data

To: Laura Tarantino, Ph.D.
Director, Office of Food Additive Safety

This memorandum is written to provide an overview of the current findings and conclusions of the FDA Center for Food Safety and Applied Nutrition's Office of Food Additive Safety on the safety of the artificial sweetener, aspartame (ASP). This reassessment of ASP safety was elicited by receipt of a new rat carcinogenicity study. This overview includes a review of other more recently published toxicological data as well.

Background of the European Ramazzini Foundation (ERF) Study

In the summer of 2005, the European Ramazzini Foundation of Oncology and Environmental Sciences Cancer Research Center in Bologna, Italy, released the results from a lifetime feeding study that they believed showed that ASP induces lymphomas and leukemia in female rats. The results have been published in various publications, including the National Institute of Environmental Health's *Environmental Health Perspectives.* Subsequently (July 2005) the FDA requested the ERF to provide data from this study. The Foundation provided FDA a portion of the information requested in February 2006 at which time the Agency proceeded with its review. In June, 2006, FDA requested additional data from the Ramazzini Foundation which consisted of the data that were not provided but requested in the February, 2006 letter to the ERF. The Foundation responded to the letter but has not provided the requested additional data.

The European Food Safety Authority announced the results of their review of the Ramazzini study (May, 2006) concluding that the data did not support that ASP is a carcinogen and saw no need to change its safety decision on ASP.

Design of the ERF Lifetime Rat Study

This study involved the feeding of ASP in varying doses (100,000 ppm, 50,000 ppm, 10,000 ppm, 2,000 ppm, 400 ppm, and 80 ppm in the diet and equivalent to 5,000, 2,500, 500, 100, 20, 4 or 0 mg/kg b.w in humans) to an in-house colony of Sprague-Dawley rats for their natural lifetime. The 1800 rats used in the study were allocated to 7 treatment groups. One hundred males and 100 females were assigned to each of the three highest dose groups. One hundred fifty males and 150 females were assigned to each of the remaining dose and control groups.

Discussion of the ERF Study

A complete and acceptably rigorous review of this study could not be performed due to the lack of critical individual animal data (e.g., individual weights, clinical observations, etc.), and the use of mean values for other data without the inclusion of standard deviations or standard errors. A lack of current historical control data for the animal colony also makes the authors' clinical observations and pathologic findings difficult to verify and interpret.

The end-of-life study design creates problems in that the increase in background pathology over the total life of the animal can confound interpretation of changes that may be related to treatment (Federal Register, March 14, 1985, pp. 10371-10442; Office of Science and Technology Policy, Part II, Chemical Carcinogens; A review of the science and its associated principles). This increase in background pathology is one reason that most official guidelines recommend a testing duration for rat carcinogenicity studies of two years, which constitutes the major portion of their life span. Another issue with the design of this study is the unusually wide ASP dose range (six dose groups from 80 to 100,000 ppm, 4 to 5,000 mg/kg b.w./day), which is purportedly based on an "assumed" range of human daily intake (they do not reference a source for their numbers). Actual reported intakes by consumers are in the range of 3 – 5 mg/kg/day.

The study narrative discussed in-life variables such as weight gain, feed and water consumption and survival; however, there were no statistical analyses assessing intragroup changes in these variables. The observed decreased weight gain in the high dose ASP groups relative to the other treatment and the control groups may be attributable to taste aversion, which has been observed in high dose groups in other ASP studies. The survival data are of particular interest in that the two high dose groups for both sexes had a higher percentage of surviving animals at 105 weeks of age than did the lower dose groups and the control group. Whether this is related to the overall lower weight gain in these groups is difficult to determine. However, numerous studies have shown that dietary restriction in rodents can lead to increased longevity, as well as altered incidences of cancer over time (reference 5). There was no discussion of any of these potentially important changes in the report narrative.

Summary and Conclusions for ERF Study

General

There is insufficient information included in the review package to allow a complete and definitive review of the ERF end-of-life rat ASP study.

Issues such as a rarely observed type of dose-response for the wide range of doses, potential for increased background pathology due to extended duration of study, possible presence of epizootic infection among the test animals (see attachment 1), apparent taste aversion in high dose groups with concomitant effects on feed consumption and weight gain (and possibly on longevity), disparities in feed consumption and body weight change, and housing issues (i.e., small cage size) that may have increased stress on test animals, as well as unrecognized confounding factors in the study design (see attachment 2), raise questions concerning how much these study design shortcomings and uncontrolled variables may have adversely affected the outcomes of the study and the authors' objective interpretation of their results.

Pathology (see attachment 1 for complete discussion)

The ASP-related findings proposed by the study's authors are just not evident from the data presented. The study design, the data presentation and the use of diagnostic criteria are not consistent with current recommendations for the conduct of carcinogenicity studies. Specifically, the study duration over the lifetime of the test animals, the high incidence of inflammatory lesions and combining of incidences of unrelated changes in the summary tables make the reported study results highly questionable.

Based on review of the ERF aspartame study, the pathologic changes were incidental and appeared spontaneously in these rats, which lived up to a year longer than rats in routine carcinogenicity studies. None of the histopathologic changes are related to treatment with ASP.

With regard to evaluation of selected ERF slides by the National Toxicology Program's Pathology Working Group (PWG):

- The PWG agreed with many of the diagnoses, however, there were also substantial differences regarding the classification of some lesions.
- It appeared that there were many slides with autolytic changes and these changes impaired the PWG's ability to classify certain lesions in more detail.
- In the case of kidney lesions, the PWG reported that there were inflammatory lesions associated with the proliferative changes. This certainly raises questions about the diagnostic accuracy and the significance of the reported renal changes.

Statistics (see attachment 2 for complete discussion)

There is no mention of blinding in necropsies and histological evaluations. To prevent conscious or unconscious bias in these evaluations, examiners should always be blinded to dose group membership when possible. Blinding is critical in ensuring the validity of experimental results.

The rats had a very high incidence of bronchopneumonia. More importantly, there were substantial differences in the incidence of bronchopneumonia across dose groups. The dose-response for bronchopneumonia in each sex was that the lowest dose of aspartame had a substantially lower incidence rate than the control group, with the incidence increasing monotonically with dose until the incidence for the highest dose was nearly as high as the control group. This unusual dose-response pattern suggests that there may be unrecognized confounding factors in the design or conduct of the study that resulted in treatment groups that were not comparable.

A large number of statistical tests were performed in the analysis of the study. When conducting many of these tests, one would expect numerous statistically significant results by chance alone. The authors should have described how they dealt with this statistical testing multiplicity issue in the interpretation of the study.

Studies with Transgenically Modified Mouse Models

As science advances new toxicological test methods are developed. Recently advanced techniques in cell and embryo cultures have been used to design transgenic mouse models. DNA with appropriate single gene changes is microinjected into the cell nucleus of fertilized eggs. This technique results in the change of only one genetic characteristic within the resulting mice which can then be used to determine whether this single genetic change results in modification of the toxic response.

One of the first applications of this technique was to modify genes implicated in the development of cancer. When certain gene alterations were made, a high incidence of certain kinds of tumors occurred in the mouse model. Thus, a number of mouse transgenic strains have been developed specifically for the purpose of studying aspects of tumor growth, such as differentiation and/or cell proliferation. In addition to simply modifying a gene characteristic, it is also possible to "knockout" or remove certain genes. The three genetically modified mouse models discussed below were used to see if ASP administered to the mice would result in increased numbers of tumors being expressed.

Suitability of the mice models

The three genes altered in the genetically modified mice models used in this NTP study are v-Ha-*ras* oncogene, *p53* tumor suppressor gene and *Cdkn2a* tumor suppressor gene. Mutated, i.e., activated, v-Ha-*ras* oncogene alters signal transduction and cellular growth control. In Tg.AC mice models, the skin behaves as if genetically initiated and exposure to tumor promoters results in the development of papillomas without the need for prior initiation (Jacobson-Kram et al. *Toxicol. Pathol.* 32, Suppl. 1: 49-52, 2004). The U.S. FDA considers this model useful for dermally applied products; however, some data from products intended for systemic administration but assayed using the dermal route in the Tg.AC model, have been reviewed (MacDonald et al. *Toxicol. Sci.* 77: 188-194, 2004).

The *p53* gene is critical to a cell's response to environmental stress, including cell cycle control and DNA repair. Its reduced expression, due to inactivation of one or both alleles, has been associated with tumors of many organs (Vogelstein, B. and Kinzler, K. W. *Cell* 70: 523-526, 1992; http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=191170). Heterozygous (only one copy of the active gene) knockout mice ($p53^{+/-}$) have a very low occurrence of spontaneous tumors for the first 36 weeks of life, but after about 80 weeks, half the animals develop spontaneous tumors including lymphomas, osteosarcomas and hemangiosarcomas. Nullizygous (both copies of the affected active gene are removed) mice ($p53^{-/-}$) develop spontaneous tumors very quickly (Jacobson-Kram et al. *ibid*). The U.S. FDA considers this as an appropriate alternative model when dealing with compounds that are clearly genotoxic (MacDonald et al., *ibid*).

Cyclin-dependent kinase inhibitor-2A (*Cdkn2a*) is also a tumor suppressor gene which is frequently deleted or mutated in tumor cells. Lukas et al. (*Nature* 375: 503-506, 1995) showed that wildtype p16 arrests normal diploid cells in late G1, whereas a tumor-associated mutant of p16 does not (http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=600160). However, this model has not been as well studied as the other two models. Thus, its suitability for 9-month carcinogenicity studies has been neither well established nor disputed.

Overall Conclusion from the NTP Transgenic Mouse Studies

The data from the above studies do not provide any evidence that the occurrence of the observed primary neoplasms in these mice models was related to aspartame treatment; rather both the non-neoplastic and neoplastic lesions observed in the aspartame-treated groups in all three genetically modified mice models were incidental, random and not dose-related. Therefore, these effects are considered unrelated to aspartame treatment. Additionally, there was no evidence that aspartame could be a mutagen. The absence of a strong and clear-cut evidence of carcinogenicity and mutagenicity even in the highest-dose aspartame treatment group in these three genetically initiated/primed mice models strongly argues against aspartame's ability and potential to cause carcinogenesis through genotoxic mechanisms.

Brief Overview of the Prospective Epidemiology Study

This new epidemiology study from the National Cancer Institute focused on aspartame-containing beverages and aspartame added to cups of coffee or hot tea captured in the questionnaire. The study participants were men and women who ranged in age from 50-71 years, were enrolled in the National Institutes of Health (NIH)-American Association of Retired Persons (NIH-AARP) Diet and Health Study cohort, and resided in 8 study areas (6 states and 2 large metropolitan areas, Atlanta and Detroit). They had no prior history of cancer at the study's baseline. The investigators followed the NIH-AARP Diet and Health Study cohort for 5 years to ascertain newly diagnosed hematopoietic cancers (n=1,888) and malignant glioma (n=315).

This study prospectively evaluated whether or not consumption of aspartame-containing beverages is associated with the occurrence of hematopoietic or brain malignancies. At baseline in 1995-1996, the investigators used a version of the National Cancer Institutes' instrument, the Diet History Questionnaire, answered by 473,984 study participants to capture dietary intake of over 100 food items, most of which were collected to be analyzed for other research purposes.

Summary and Conclusions for Prospective Epidemiology Study

This large prospective epidemiological study did not detect an increased risk of hematopoietic (including lymphomas and leukemias) or brain malignancies associated with consumption of aspartame-containing beverages. A few of the important characteristics as well as the strengths and weaknesses of the study are summarized below followed by concluding remarks.

The large number of study participants and the prospective design of the study are major study strengths. The study size and the older age of the participants (50-71 years old) allowed for greater accrual of hematopoietic and brain malignancies over the 5 year follow-up period than would be expected with a smaller study enrolling younger participants. The prospective study design provided ascertainment of aspartame intake before diagnosis of hematopoietic or brain malignancy. Disease status, therefore, could not influence the study participant's recall of consumption of aspartame-containing beverages, and the investigators could evaluate the temporal relationship between aspartame intake and the diagnosis of cancer.

OVERALL SUMMARY AND CONCLUSIONS

Based on a review and evaluation of the data provided by the ERF for their lifetime Sprague-Dawley rat study it was not possible to confirm their reported results of aspartame-induced increases in the occurrence of tumors. There are significant shortcomings in the protocol design, conduct, reporting and interpretation of this study.

The evaluation and interpretation of this new ERF experimental evidence on aspartame by the FDA was problematic. The agency did not receive important portions of the study data which were reported to support the principal investigators' conclusions. As a result of this lack of information, FDA was not able to verify the conclusions of this study, indeed there is significant consensus among those who reviewed these study results that they are not subject to acceptable evaluation and verification until key pieces of information are provided. For example, individual animal data commonly available on food consumption, water intake, weight gain, clinical observations (such as histopathology slides) as well as other parameters.

One of the most important components of experimental findings missing from the ERF report of results was an adequate sample of histopathology slides from this study. Additional insight and enhanced definition and needed support of the findings of this study would have been provided by an internationally-sponsored expert pathology working group's examination of a statistically representative sample of histopathology slides from this study.

The limited nature and extent of the data submissions from this study raises a more generic issue that is worthy of consideration by all stakeholders in this type of data evaluation/interpretation process (the regulated, other interested contributors and regulators). What is the appropriate standard for the submission of experimental evidence to be used to support the safety of a FDA-regulated product? Is it only the sponsors of a regulated product that must provide comprehensive, in-depth data/evidence or is it the responsibility of any party who would submit data/evidence with the possibility of significantly affecting the regulatory fate of a compound under consideration? It is a well accepted concept in science that those who seek to dramatically alter widely accepted experimental findings bear a significant burden to appropriately present and defend their new findings and conclusions.

Considering the results from a broad array of studies, including five previously conducted negative chronic bioassays with aspartame, additional negative carcinogenic findings from a series of three transgenic mouse assays conducted by NTP and finally, a recently reported large NCI epidemiology study that reported no associations between the use of aspartame-based beverages and the occurrence of hematopoietic (including lymphomas and leukemias) or brain tumors, the FDA concludes that the present regulations governing the use of aspartame as a food additive are supported by all of the appropriately presented scientific evidence. The ERF study data does not provide sufficient evidence to alter the current FDA position that there is reasonable certainty of no harm with the food additive use of aspartame.

Another strength of the study is the use of a validated food frequency questionnaire to capture information on the study participants' intake of beverages. Generally, this type of questionnaire obtains information about the frequency of using foods over the preceding year with the idea that the food consumption pattern in a single year will tend to correlate with the individual's dietary intake over the past several years (Willett 1998, p.81). If the preceding year is indicative of several years, then there is a longer lag time than is indicated in this study (5.2 or fewer years) between initial exposure to aspartame and the diagnosis of cancer. An exposure that is truly causative of cancer, whether initiating or promoting cancer, must precede a cancer diagnosis by a period of time that allows for the proliferation and manifestation of the cancer.

The timing for the baseline questionnaire in 1995-1996 was important because it occurred prior to FDA's approval of two sugar substitutes, sucrolose and neotame, for use in products. Since the questionnaire did not ask specifically about aspartame in diet drinks, this timing provided the investigators a rationale for assigning aspartame to these products. As discussed earlier, saccharin and acesulfame-K were also on the market at this time but not nearly to the same extent as aspartame. One weakness of the study is some of the imprecision around the estimation of aspartame intake and possible misclassification of exposure for some study participants. The effect of this imprecision and misclassification on the association between aspartame intake and cancer is dependent on its extent and whether the person developed cancer or not. The estimation of the relative risk for hematopoietic or brain malignancies could either increase or be reduced with changes in aspartame assignment and measurement.

Because of the large number of persons enrolled in the study, there was accrual of sufficient number of relatively rare subtypes of hematopoietic cancers and malignant gliomas for analysis. Investigators indicated that they had the power (80%) to detect a moderate association of 600 mg/d or more of aspartame with several of the cancers, for example, overall hematopoietic cancer (RR>1.24) and gliomas (RR>1.52). The investigators adjusted for a number of dietary, lifestyle or medical history variables in the multivariate analyses to limit confounding of the aspartame – cancer association.

In conclusion, while there are issues in the study regarding estimation of aspartame intake and the sufficiency of time from intake of aspartame to cancer diagnosis that may have bearing on the study's findings, several strengths of the study lend support for the results. These strengths are the large sample size of the study that allowed for the accrual of sufficient numbers of hematopoietic and glioma malignancies for analysis, the study design (cohort) that allowed for temporal associations, the timing of the baseline questionnaire prior to FDA approval of sucrolose and neotame, the validated food frequency questionnaire used to obtain beverage consumption information, the careful evaluation of cancer ascertainment, and the consistency of the relative risk estimations across most dose categories. Epidemiological study results are best supported if there are multiple epidemiological studies in different populations that demonstrate the same findings. In this way, chance or other factors such as diet, lifestyle, medical/family history and other exposures are ruled out as explanations for the associations. One earlier epidemiological study, a case-control study conducted with 56 cases and 94 controls, evaluated aspartame consumption and brain tumor risk in children (Gurney et al. 1997) and did not find an increased risk of brain tumor with aspartame intake. No other epidemiological study has evaluated hematopoietic cancer and aspartame.