ELSEVIER

Journal of Chromatography B, 778 (2002) 375–381

JOURNAL OF CHROMATOGRAPHY B

www.elsevier.com/locate/chromb

Precision and sensitivity of aminobiphenyl hemoglobin adduct assays in a long-term population study

Paul L. Skipper*

Massachusetts Institute of Technology, Division of Bioengineering and Environmental Health, Building 56, Room 747, Cambridge, MA 02139, USA

Abstract

Exploratory statistical analysis of aminobiphenyl hemoglobin adduct data obtained in a large-scale population study was performed to assess precision and sensitivity over the 7 years required to conduct the analyses. A time-dependent trend toward higher values was observed that may be attributable to aging of the internal standard used throughout the study. A several-fold improvement in sensitivity from the beginning to end of the study was also noted. Repeated analysis of duplicate blood specimens provided a worst-case estimate of the coefficient of variation to be 0.31, attributable almost entirely to sample preparation rather than instrumental analysis. Substantial variability in calibration curves for the deuterated internal standard (standard deviation was ±15% of the mean) was observed. The results obtained here will be used in support of further analyses of the data with respect to factors of epidemiological interest.

Keywords: Aminobiphyl hemoglobin adduct

1. Introduction

Hemoglobin adducts of aromatic amines, specifically those with the sulfinamide structure that results from reaction of the oxidized amine with sulfhydryl residues [1], have been developed very successfully as biomarkers for this class of toxic and carcinogenic compounds. A considerable number of studies in human populations have now been reported. These include several occupational exposure studies [2–5] and numerous demonstrations of increased exposure to aromatic amines as the result of exposure to tobacco smoke [6–9]. Investigations of gene-environment interaction [2,7,9–13] have also been greatly facilitated by the availability of hemoglobin adduct assays, since several metabolic pathways are either known or suspected to mediate the formation of the adducts.

Quantitative analysis of aromatic amine hemoglobin adducts requires hydrolysis of the relatively labile sulfinamide, isolation and derivatization of the amine thus produced, and, finally, GC–MS quantitation. An internal standard is added very early in the process to ensure precision, generally just prior to the hydrolysis step. Different approaches to the individual steps have been adopted by different laboratories, but the overall process is undertaken essentially as outlined. No attempt will be made here to review the various reported methods; instead, it will be the focus of this paper to address the issue of

*Corresponding author. Tel.: +1-617-253-0983; fax: +1-617-252-1787.

E-mail address: skipper@mit.edu (P.L. Skipper).

1570-0232/02/$ – see front matter
PII: S0378-4347(01)00463-7

assay performance for one specific assay [14] that has been used in this laboratory for many years.

In 1992, this laboratory joined an ongoing case-control study of urinary bladder cancer in Los Angeles County, which has recently been completed, to add aromatic amine hemoglobin adduct assays to the study protocol. One focus of this laboratory component was on adducts of the carcinogen 4-aminobiphenyl (4-ABP) and, over the course of the study, assays of 1536 blood specimens were performed. Assays of adducts of other aromatic amines, including 3-aminobiphenyl, were also included. Analysis of the results of aminobiphenyl adduct assays forms the body of work discussed here. By virtue of its size, this data set offers an exceptional opportunity to evaluate a variety of assay parameters with statistical power not available from smaller studies.

2. Experimental

Assays were performed essentially as described previously [14]. Briefly, solutions of hemoglobin were prepared by lysis of red cells with distilled H_2O and toluene, centrifugation, and dialysis of the supernatant against distilled H_2O. After addition of an aliquot of internal standard, adducts were hydrolyzed by addition of NaOH to 0.1 *M*, followed by incubation at room temperature for 1 h. Amines were extracted with hexane, which was subsequently dried over $MgSO_4$, and derivatized with $(C_2F_5CO)_2O$ and $(CH_3)_3N$. Hexane was removed using a rotary evaporator and samples were redissolved in 20 μl of hexance for GC–MS analysis. Exceptions to the procedures described and details specific to this study are indicated below.

Red blood cells were prepared at the University of Southern California, where the case-control study was conducted and were stored frozen prior to shipment to MIT where they continued to be stored frozen at −30°C. Shipments were made when warranted by collection of sufficient numbers of specimens. Duplicate blood specimens were obtained from each subject, except in a few cases where only one specimen was donated.

Blood specimens were generally prepared for analysis in batches of either 16 — in the early part of the work — or, later, in batches of 24. Each batch comprised an equal number of cases and their matched controls. Specimens were identified with a code number only and the analysts were blinded to the coding.

In the early phase of the study, hexane was redistilled as described [14]. Baker N168-08 hexane was later found to be suitable for use as supplied.

A Hewlett-Packard 5989A MS-engine was used for all GC–MS analyses, operated in the negative ion chemical ionization mode with CH_4 as the moderating bath gas. Several different GC columns were used throughout the study, but all were essentially equivalent, having a thin carbowax-type stationary phase and dimensions of 0.2–0.25 mm I.D. by 12–15 m length. An autoinjector (HP7673) was used for sample introduction. GC oven temperature programs typically began with a 1-min isothermal phase at 60°C followed by a rapid (20°C/min) ramp to a final temperature of 240–250°C, which was held for several minutes. The aminobiphenyl derivatives eluted during the final phase. Selected ion monitoring at m/z 295 and 304, which correspond to $[M-20]^{-\cdot}$, was used to detect analytes and internal standard, respectively. All chromatograms were manually integrated.

The internal standard was prepared by reaction of *N*-hydroxy-4-aminobiphenyl-d_9 with normal human hemoglobin isolated from frozen, washed erythrocytes. The hydroxylamine was prepared from 4-nitrobiphenyl-d_9 by reduction with ammonium sulfide (Fisher A-705) using a published method [15]. The nitrobiphenyl was prepared by nitration of biphenyl-d_{10} (Cambridge Isotope Laboratories, Andover, MA, USA) by published methods [16]. The concentration of adduct was initially estimated by spiking an aliquot of the adducted hemoglobin solution with 4′-F-ABP, hydrolyzing the adduct and isolating the aromatic amines by globin precipitation with acidic acetone, followed by GC–MS analysis of the amines after derivatization. Based on the initial estimate of 154 ng of 4-aminobiphenyl-d_9/100 μl, a portion of the concentrated adduct solution was diluted to 1:1000 for use in future assays. Aliquots of the diluted solution were prepared and stored at −30°C. This one preparation of the internal standard was used throughout the study. A fresh aliqout was

used for each batch of specimens. The amount of standard added to each specimen was 100 μl.

Prior to the commencement of this study, the internal standard described above was used in six other studies in combination with the previously described 4′-F-ABP internal standard.

Results from these analyses agreed well with the initial determination of the concentration of 4-aminobiphenyl-d_9. Consequently, a value of 150 was assumed as the factor to be used in the present study to convert the 4-ABP/4-ABP-d_9 peak area ratio to picograms of 4-ABP: 150 = 154 · (mol. wt. of 4-ABP/mol. wt. of 4-ABP-d_9), rounded off to two significant figures. The amount of amine adduct detected was normalized to the amount of hemoglobin in the sample so that results are presented as pg/g hemoglobin.

3. Results and discussion

3.1. Long-term stability of the assay

Analyses for this study were conducted over a period of 7 years. Consequently, it was important to determine whether consistent results were obtained throughout the period and to characterize any systematic variation that might have occurred. Assay reproducibility is frequently determined by inclusion of quality control (QC) samples along with samples undergoing analysis. QC samples were not included as part of this protocol for a variety of reasons. A pooled blood sample might be considered the ideal QC sample, but this option was rejected because it has not been established that the adduct level in such a sample would itself be stable over the lifetime of the study. QC options requiring more than a single pooled sample were also rejected as being unnecessarily labor and cost intensive given the historical experience of the relative stability of this assay. Long-term assay stability can, nevertheless, be evaluated without QC data by examination of the study data because of the large number of analyses conducted.

Fig. 1 shows all the adduct values as a function of the time elapsed since the first analysis was conducted. The numbers have been log transformed here and elsewhere in this report to normalize the skewed distributions that characterize the data. Linear regression analysis revealed a highly significant increase in

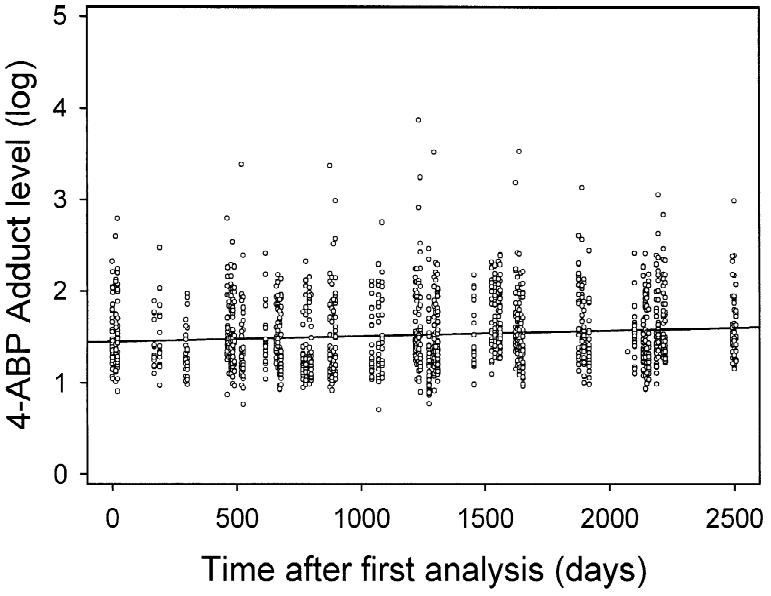

Fig. 1. Linear regression analysis of all adduct values. The independent variable is the interval between the date of sample preparation and Nov. 15, 1993, the first day on which the samples were prepared. The equation of the fitted model is $Y = 1.447 + 0.0000625 \cdot X$. Adduct levels are expressed as pg/g hemoglobin.

adduct level as a function of time (slope = 6.25×10^{-5}, $P < 0.0001$). Adduct levels calculated from the equation of the fitted model range from 28.0 at the beginning of the study to 40.1 at the end, an increase of 43%.

It appears from inspection of Fig. 1 and the regression analysis described that there was a gradual increase in the adduct levels detected over the course of the study. It cannot be ruled out that the increase represents a true change in the adduct levels in the subjects. It can be shown, though, that the increase is not attributable to a demographic change related to the smokers in the sample population. First, the number of smokers per each 100 subjects, which ranged from 16 to 28, did not change significantly from beginning to end. Second, regression analysis of the nonsmoker data produced a result very similar to that obtained with the entire data set: a slope of 5.46×10^{-5} ($P < 0.0001$) and an increase of 37% in the calculated values from 21.8 at the beginning of the study to 29.9 at the end.

A probable explanation for the increase in measured adduct levels, if it does not represent a true change in the adduct levels in the subjects, is a gradual degradation of the internal standard, leading to increased detected values from the same actual values. This explanation is not testable after the fact, but it is reasonable to assume that the standard is not indefinitely stable. In any event, a steady increase in values, on the order of 40% over the course of the study, even if it results from a change in the internal standard concentration, is not expected to materially influence the conclusions of the study. While it may reduce statistical power to some degree by increasing dispersion of the data, it should not introduce any systematic bias because the number of cases and controls was evenly distributed throughout the course of the study by design and other factors of interest such as gender, genotype, medication use, occupation, etc., are expected to be randomly distributed unless very small in number.

3.2. Precision

At one point in the study, several batches of specimens were worked up with a modified extraction technique that resulted in very low recoveries, not more than 5–10% of the usual recovery. GC–MS analysis yielded positive 4-ABP adduct values from only 67 out of 96 samples. Since it was desired that there be as few missing data as possible and since very low recovery was expected to produce less reliable results, a second specimen of red cells from each of the 96 subjects involved was worked up and analyzed. Values from the first round of sample preparation and analysis are compared with values from the second round in Fig. 2. The correlation ($r = 0.944$) between the two sets of data provides a measure of the overall precision of the assay, since it compares results obtained with identical blood specimens carried through the entire procedure from blood collection to mass spectrometric analysis. Assessment of precision from this data set undoubtedly produces an underestimate of precision obtainable under optimum conditions since half the results are compromised by very low recoveries. This ''worst case'' analysis is nevertheless useful in setting a lower limit to what may be expected. In the present instance, the coefficient of variation was 0.31.

Nearly all the variability can be attributed to factors other than instrumental analysis. Fifty-one of the 67 low recovery samples were analyzed twice by GC–MS. Statistical analysis of the results yielded a

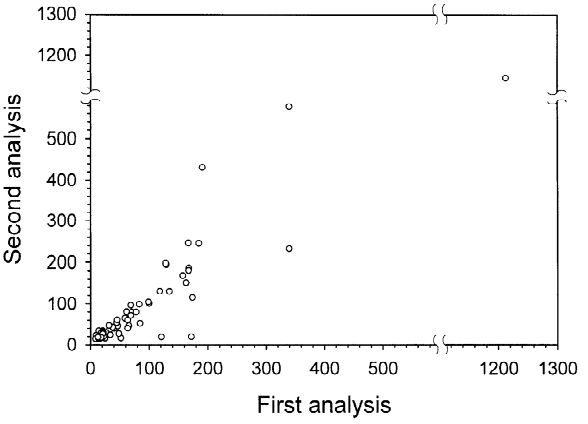

Fig. 2. Comparison of results obtained by repeated analysis of duplicate blood specimens. The first set of data were obtained under conditions of low recovery; these values are thus expected to be less precise than those in the second set. The relative differences between each pair of assays were apparently normally distributed, but, because the data set is relatively small, fitting a normal distribution to the data did not quite achieve statistical significance.

correlation coefficient of 0.999, and there were no outliers such as those that are apparent in Fig. 2. Manual integration of chromatograms was used throughout the study, and, while not always required, was clearly necessary to obtain good data from these samples. The very low C.V. obtained (0.034) demonstrates that manual integration can be very consistent, even with quite noisy chromatograms.

3.3. Sensitivity

As with all previous studies of aminobiphenyl hemoglobin adducts, every sample in this study contained readily detectable 4-ABP. Signal-to-noise ratios were typically greater than 100 and frequently over 1000. Adducts of 3-ABP, in contrast, were often undetectable. Over the course of the study, the limit of detection improved, so that the chromatographic peak corresponding to 3-ABP became apparent much more frequently. Fig. 3 shows a chromatogram produced by analysis of a nonsmoker's hemoglobin toward the end of the study that is representative of the lowest detectable values. The 3-ABP adduct level determined from this analysis is 0.2 pg/g Hb, and, taking into account recovery and injection volume, it is estimated that no more than 50 amol of derivatized 3-ABP was injected into the GC–MS for this analysis.

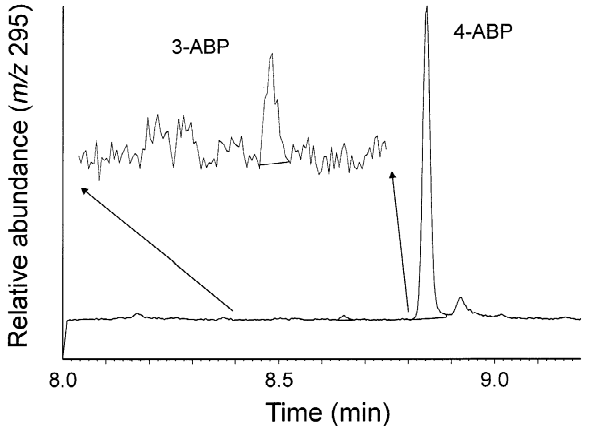

Fig. 3. GC–MS analysis representative of the lowest detectable 3-aminobiphenyl adduct levels. Values calculated from this chromatogram are 0.2 and 23.1 pg/g Hb for 3-ABP and 4-ABP, respectively. The 4-ABP adduct level in this subject is 4.5 times greater than the lowest observed value.

It is reasonable to question whether such GC–MS peaks represent true analyte concentration or result from other causes. Because the available data set is large, one approach to testing the validity of the lowest values is to analyze the data with respect to a known associated factor. Smoking status, for instance, has been clearly established as a determinant of 3-ABP adduct levels [6]. While smokers typically exhibit 4-ABP adduct levels that are 3–4 times higher than those of nonsmokers, the difference is even greater — about ten-fold — for 3-ABP adducts [6,7]. Thus, considerable support for the validity of the lowest detectable 3-ABP adduct values would be obtained if they also showed a smoking status-related difference.

The set of detectable 3-ABP values less than 1 pg/g Hb contains enough smoker data to allow statistical testing. This set comprises 11 smoker (5.4%) and 193 nonsmoker (94.6%) values. In itself, the distribution between smokers and nonsmokers is a strong indication of the validity of the values because 22.0% of the entire set of subjects were current smokers. One-way ANOVA confirms that the smoker values differ from the nonsmoker values. The mean values for smokers and nonsmokers are 0.69 and 0.46 pg/g Hb, respectively, and the difference is highly significant ($P = 0.0048$).

3.4. Mass spectrometer calibration

In the detailed description of the methods used for GC–MS analysis given in Ref. [14], it is stated that there is no significant difference in mass spectrometer response to deuterated amines as compared to the isotopically normal amines. As a result of the extensive experience gained from the present study, it has become apparent that there can indeed be significant differences in response. The average slope of all the calibration curves comparing d_0:d_9 peak area ratios with actual composition was 1.30, rather than 1. More importantly, the response ratio was not always the same. Calibration curves run on different days exhibited slopes that varied with a standard deviation of $\pm 15\%$ of the mean and ranged as low as 1.00.

It is not apparent why such substantial variation occurs. An explanation may lie in the fact that the ion monitored is a fragment ion rather than a

molecular ion and that conditions in the ion source affect the fragment yield, but the experiments required to test the hypothesis have not been conducted. An interlaboratory comparison, though, confirmed that the experience is not unique.

When the same set of calibration standards were analyzed by two other laboratories, different and variable results were observed. One laboratory reported a response ratio closer to but still greater than one, while the other observed a value of 0.45. The second laboratory has had considerable experience with 4-ABP assays and characterizes their instrument response variability as greater than that reported here.

Clearly, variability in response to deuterated standards as compared to the non-deuterated analytes is a significant factor affecting precision in aromatic amine assays. The extent to which it is a factor appears to differ between laboratories, but it may be that it has been underestimated previously.

4. Conclusion

Several factors affecting precision of the aminobiphenyl hemoglobin adduct assay have been identified by exploratory statistical analysis of the data obtained in one extended population study. They include systematic variation such as a time-dependent increase in values, as well as random variation such as that noted from repeated analysis of duplicate blood specimens.

Awareness of these factors will be important for the subsequent statistical analysis of the data with respect to epidemiological variables.

Acknowledgements

Financial support was provided by a NIH grant ES05622. The author also wishes to acknowledge the invaluable assistance of Joseph Glogowski in the preparation of samples for analysis.

References

[1] H.-G. Neumann, in: J.D. Groopman, P.L. Skipper (Eds.), Molecular Dosimetry and Human Cancer, CRC Press, Boca Raton, 1991, p. 363.
[2] M. Riffelmann, G. Muller, W. Schmieding, W. Popp, K. Norpoth, Int. Arch. Occup. Environ. Health 68 (1995) 36.
[3] D. Schütze, O. Sepai, J. Lewalter, L. Miksche, D. Henschler, G. Sabbioni, Carcinogenesis 16 (1995) 573.
[4] E.M. Ward, G. Sabbioni, D.G. DeBord, A.W. Teass, K.K. Brown, G.G. Talaska, D.R. Roberts, A.M. Ruder, R.P. Streicher, J. Natl. Cancer Inst. 88 (1996) 1046.
[5] G. Sabbioni, J. Wei, Y.Y. Liu, J. Chromatogr. B 682 (1996) 243.
[6] M.S. Bryant, P. Vineis, P.L. Skipper, S.R. Tannenbaum, Proc. Natl. Acad. Sci. USA 85 (1988) 9788.
[7] M.C. Yu, P.L. Skipper, K. Taghizadeh, S.R. Tannenbaum, K.K. Chan, B.E. Henderson, R.K. Ross, J. Natl. Cancer Inst. 86 (1994) 712.
[8] B. Branner, C. Kutzer, W. Zwickenpflug, G. Scherer, W.-D. Heller, E. Richter, Biomarkers 3 (1998) 35.
[9] N.M. Probst-Hensch, D.A. Bell, M.A. Watson, P.L. Skipper, S.R. Tannenbaum, K.K. Chan, R.K. Ross, M.C. Yu, Cancer Epidemiol. Biomark. Prev. 9 (2000) 619.
[10] J. Lewalter, U. Korallus, Int. Arch. Occup. Environ. Health 56 (1985) 179.
[11] P. Vineis, N. Caporaso, S.R. Tannenbaum, P.L. Skipper, J. Glogowski, H. Bartsch, M. Coda, G. Talaska, F. Kadlubar, Cancer Res. 50 (1990) 3002.
[12] M.C. Yu, R.K. Ross, K.K. Chan, B.E. Henderson, P.L. Skipper, S.R. Tannenbaum, G.A. Coetzee, Cancer Epidemiol. Biomark. Prev. 4 (1995) 861.
[13] M.T. Landi, C. Zocchetti, I. Bernucci, F.F. Kadlubar, S. Tannenbaum, P. Skipper, H. Bartsch, C. Malaveille, P. Shields, N.E. Caporaso, P. Vineis, Cancer Epidemiol. Biomark. Prev. 5 (1996) 693.
[14] P.L. Skipper, W.G. Stillwell, in: J. Everse, K.D. Vandegriff, R.M. Winslow (Eds.), Methods in Enzymology, Vol. 231, Academic Press, San Diego, 1994, p. 643.
[15] M.R. Thissen, R.W. Roth, W.P. Duncan, Org. Prep. Proc. Int. 12 (1980) 337.
[16] F. Bell, J. Kenyon, P.H. Robinson, J. Chem. Soc. (1926) 1239.

Paul L. Skipper was awarded the A.B. degree in chemistry from Washington University (St. Louis) in 1972. He continued studying chemistry in the doctoral program at Brandeis University, completing his studies in 1975, and receiving his Ph.D. in 1976. At Brandeis, he specialized in organic chemistry with a focus on new synthetic methods and reagents under the direction of Prof. James. B. Hendrickson. He then pursued further research on organic synthetic methods and reagents as a postdoctoral fellow at the Université de Louvain, Belgium, with Prof. Heinz G. Vienhe. Since 1976 he has been a staff scientist at the Massachusetts Institute of Technology, beginning with an initial postdoctoral appointment in the laboratory of Prof. Steven R. Tannenbaum. His present appointment is as Principal Research Scientist, a non-academic position equivalent in rank to Associate Professor. At MIT, the focus of his research shifted from synthetic organic chemistry to the organic chemistry of biological systems. His current research interests include development of methods for quantitative analysis of protein and DNA adducts formed by chemical carcinogens and application of such methods in epidemiological studies.