Clinical Chemistry 47:2
292–300 (2001)

Drug Monitoring and Toxicology

Performance Characteristics of a Carbon Isotope Ratio Method for Detecting Doping with Testosterone Based on Urine Diols: Controls and Athletes with Elevated Testosterone/Epitestosterone Ratios

Rodrigo Aguilera,[1*] Thomas E. Chapman,[1] Borislav Starcevic,[1] Caroline K. Hatton,[1] and Don H. Catlin[1,2]

Background: **Carbon isotope ratio methods are used in doping control to determine whether urinary steroids are endogenous or pharmaceutical.**

Methods: **Gas chromatography-combustion-isotope ratio mass spectrometry (GC-C-IRMS) was used to determine the δ^{13}C values for 5β-androstane-3α,17β-diyl diacetate (5βA), 5α-androstane-3α,17β-diyl diacetate (5αA), and 5β-pregnane-3α,20α-diyl diacetate (5βP) in a control group of 73 healthy males and 6 athletes with testosterone/epitestosterone ratios (T/E) >6.**

Results: **The within-assay precision SDs for 5βA, 5αA, and 5βP were ± 0.27‰, ± 0.38‰, and ± 0.28‰, respectively. The between-assay precision SDs ranged from ± 0.40‰ to ± 0.52‰. The system suitability and batch acceptance scheme is based on SDs. For the control group, the mean δ^{13}C (SD) values were −25.69‰ (± 0.92‰), −26.35‰ (± 0.68‰), and −24.26‰ (± 0.70‰), for 5βA, 5αA, and 5βP, respectively. 5βP was greater than 5βA and 5αA (P <0.01), and 5βA was greater than 5αA (P <0.01). The means − 3 SD were −28.46‰, −28.39‰, and −26.37‰ for 5βA, 5αA, and 5βP, respectively. The maximum difference between 5βP and 5βA was 3.2‰, and the maximum 5βA/5βP was 1.13. Three athletes with chronically elevated T/Es had δ^{13}C values consistent with testosterone administration and three did not.**

Conclusions: **This GC-C-IRMS assay of urine diols has low within- and between-assay SDs; therefore, analysis of one urine sample suffices for doping control. The means, SDs, ±3 SDs, and ranges of δ^{13}C values in a control group are established. In comparison, testosterone users have low 5βA and 5αA, large differences between 5βA or 5αA and 5βP, and high 5βA/5βP and 5αA/5βP ratios.**

A critical issue in doping control is establishing the origin of testosterone and other steroids typically found in human urine. If the origin can be shown to be pharmaceutical, as opposed to endogenous, a doping offense has occurred. Androgens for pharmaceutical use are not synthesized de novo; they are obtained by semisynthesis from starting materials such as diosgenin and stigmasterol, which are derived from plants *(1)*. Furthermore, these compounds have less ^{13}C than their endogenous homologs *(2, 3)*; therefore, urinary steroids with a low $^{13}C/^{12}C$ ratio are likely to have originated from pharmaceutical sources. This thesis has led to the use of gas chromatography-combustion-isotope ratio mass spectrometry (GC-C-IRMS)[3] in doping control because the method determines the δ^{13}C value of steroids extracted from urine, based on the equation:

[1] UCLA Olympic Analytical Laboratory, Department of Molecular and Medical Pharmacology and [2] Department of Medicine, University of California at Los Angeles, Los Angeles, CA 90025.

*Address correspondence to this author at: UCLA Olympic Analytical Laboratory, Department of Molecular and Medical Pharmacology, University of California at Los Angeles, 2122 Granville Ave., Los Angeles, CA 90025-6106. Fax 310-206-9077; e-mail rodrigoa@ucla.edu.

Received June 7, 2000; accepted November 22, 2000.

[3] Nonstandard abbreviations: GC-C-IRMS, gas chromatography-combustion-isotope ratio mass spectrometry; T/E, testosterone/epitestosterone ratio; 5βA, δ^{13}C value for 5β-androstane-3α,17β-diyl diacetate; 5αA, δ^{13}C value for 5α-androstane-3α,17β-diyl diacetate; 5βP, δ^{13}C value for 5β-pregnane-3α,20α-diyl diacetate; and QC, quality control.

$$\delta^{13}C[‰] = \frac{(^{13}C/^{12}C)_{sample} - (^{13}C/^{12}C)_{standard}}{(^{13}C/^{12}C)_{standard}} \times 1000$$

where $^{13}C/^{12}C$ refers to the isotope ratio in the sample or an international standard (a reference gas calibrated relative to Pee Dee Belemnite). Accordingly, for several years, GC-C-IRMS has been applied to detect the use of exogenous steroids such as testosterone *(3, 4)*, dihydrotestosterone *(5, 6)*, and dehydroepiandrosterone *(6, 7)* by athletes.

Testosterone administration increases the ratio of urine testosterone to epitestosterone (T/E). This finding led to the use of T/E as a screening test for testosterone administration *(8)*. If the ratio exceeds 6, laboratories accredited by the International Olympic Committee report the result to the sport authority *(9)*. It is known, however, that factors other than testosterone administration may also increase the T/E ratios *(10–12)*. Furthermore, testosterone administration does not always lead to an elevated[4] T/E ratio *(3, 13)*. These factors complicate the interpretation of elevated T/E ratios and limit the significance of T/E ratios <6. Determining the T/E time profile of an individual is useful, but it requires consideration of the results of past tests or the collection of additional samples. Our previous work demonstrated that the measurement of $\delta^{13}C$ values of testosterone and its metabolites can detect testosterone use *(4, 14, 15)*. Others have reported that $\delta^{13}C$ values may be abnormally low even in samples with T/E ratios <6 *(3)*.

With the introduction of new methods, it is essential to fully characterize the assay and to establish the values in the control group. This is particularly true with $\delta^{13}C$ measurements when applied to drug control because the method is inherently difficult, the analytical results are likely to be litigated, and it has been suggested *(3, 5)* that diet or ethnicity may influence the $^{13}C/^{12}C$ ratio of urine steroids. Accordingly, herein we characterize the $\delta^{13}C$ measurements of the diacetates of 5β- and 5α-androstane-3α,17β-diol (5βA and 5αA, respectively), and 5β-pregnane-3α,20α-diol (5βP) with respect to precision, linearity, system suitability, and batch acceptance. In addition, we determined $\delta^{13}C$ values for these steroids in a control group of 73 healthy male subjects from four different ethnic groups. Finally, data gathered from the control group were used to interpret $\delta^{13}C$ values obtained on athletes with elevated T/E ratios and to discuss diagnostic criteria for doping.

Materials and Methods

STEROID CALIBRATION SOLUTION

A steroid calibration solution was prepared that contained 25 mg/L 5β-androstane-3α,17β-diol, 5α-androstane-3α,17β-diol, and 5β-pregnane-3α,20α-diol, all as diacetates (Steraloids, Inc.).

CONTROL GROUP

The control group consisted of 74 male medical students (age range, 19–29 years) from the University of California at Los Angeles. Body weights were 45–110 kg. The students had the option to give a medical and drug history and to describe their ethnicity. No student declared taking steroids or reported any chronic disease. The ethnicities were "Asian" (n = 16), "African American" (n = 9), "Caucasian" (n = 25), "Hispanic" (n = 13), "other" (n = 7), and "unknown" (n = 4). The students collected 24-h urine samples, and 50-mL aliquots were stored at 4 °C. The study was conducted under the guidance of the Human Subject Protection Committee of the University of California at Los Angeles.

ATHLETES WITH ELEVATED T/E RATIOS

Of the large number of anonymous urine samples that our laboratory analyzes for sport drug control programs in the US, some have a urine T/E ratio >6. In such cases, the sport organization typically collects additional samples to track the urine T/E over weeks or months. We selected some of these urine samples for GC-C-IRMS analysis based on the availability of at least two samples per person, sufficient volume to perform the analysis, and permission from the sport organization.

QUALITY-CONTROL URINE SAMPLES

Quality-control urines (QC-H and QC-L) were obtained from two subjects known not to be using testosterone or any substance likely to alter the $\delta^{13}C$ values of 5β-androstane-3α,17β-diol, 5α-androstane-3α,17β-diol, and 5β-pregnane-3α,20α-diol (measured as diacetates). Each subject donated one 8-h urine, which was aliquoted into 10-mL cryogenic tubes and stored at −20 °C until analysis.

URINARY STEROID CONCENTRATIONS AND SAMPLE PREPARATION FOR GC-C-IRMS

The T/E ratio was estimated by GC-MS as described previously *(16)*. The concentrations of 5β- and 5α-androstane-3α,17β-diol were estimated from a QC sample prepared by adding to steroid-free urine 5β- and 5α-androstane-3α,17β-diol. [16,16,17-^{2}H]-testosterone (CDN Isotopes) was used as an internal standard. The 5β-pregnane-3α,20α-diol was not quantified. The sample preparation methods for GC-C-IRMS analysis and the instrument conditions have been described *(15)*. The volume of urine extracted for GC-C-IRMS was 10 mL.

GC-C-IRMS DETERMINATION OF 5αA, 5βA, AND 5βP

The GC-C-IRMS analyses were conducted on a Finnigan Delta Plus Isotope Ratio Mass Spectrometer. The IRMS was connected to a Hewlett Packard Model 6890 gas chromatograph via a Finnigan Combustion III interface. The GC was equipped with an HP-50+ capillary column [30 m × 0.25 mm (i.d.); 0.15 μm film thickness] and a Finnigan A200S autosampler. The combustion oven was

[4] The word "elevated" is used to refer to a result that is above the administrative cutoff.

oxidized by back-flushing with oxygen for 1 h every 30 samples. The oxidation reactor in the combustion oven was replaced every 500 samples. The injection volume was 1 μL. The recoveries for 5β- and 5α-androstane-3α,17β-diol were 85% and 86%, respectively.

GC-C-IRMS response linearity

The linearity of the GC-C-IRMS response was determined by preparing seven solutions containing all three analytes at 2.5, 5, 10, 25, 50, 100, and 150 mg/L, respectively. One microliter of each solution was injected three times on 1 day, and the three $\delta^{13}C$ values for each compound were averaged.

precision studies

Instrument precision for 5βP, 5βA, and 5αA was determined by extracting one aliquot of each QC urine and injecting each extract four times in succession. The within-day precision was determined by extracting 20 aliquots of QC-H in the same batch and injecting each one once. The between-assay precision was determined by extracting one aliquot of QC-H and QC-L per day for 16 days, spanning 15 months, and injecting each aliquot once.

daily system suitability test

To establish a tolerance range, the steroid calibrator was injected five times on each of 5 different days over 2 weeks (n = 25), and the mean $\delta^{13}C$ values and SDs were calculated for each steroid. Each day, before urine sample analysis, the system suitability was assessed by injecting the calibrator three times and calculating the mean $\delta^{13}C$ values. The system was considered suitable for batch analyses if the mean $\delta^{13}C$ values of at least two of the three steroids were within ± 2 SD of the means described above. If the system suitability test failed, maintenance was performed on the gas chromatograph injection port, the GC column and/or the oxidation reactor were replaced, and the calibrator was re-injected. Data acceptance criteria included absence of peak tailing, retention times of the steroids within ± 1% of established values, and a minimum 0.8 V response at *m/z* 44 for each steroid.

criteria for batch acceptance

To establish a tolerance range for the QC urines for batch data acceptance, two aliquots of each QC urine were extracted and each injected twice per day for 5 days spanning 2 weeks (n = 20). ANOVA was performed using the factors day (5), injection (2), duplicate (2), and QC (2). The mean $\delta^{13}C$ values and SDs were calculated for each of the three steroids. For batch analysis, each unknown urine sample was extracted and injected once. One aliquot of the two QC urines and one aliquot of the steroid calibrator were analyzed with each batch of samples. These three controls were injected at the beginning, in the middle, and at the end of the batch. The batch was accepted if at least six of the nine means for the three steroids were within ± 2 SD of the previously established means. The acceptance criteria were as noted above except that an instrument response of 0.3 V at *m/z* 44 was accepted.

statistical analyses

The Smirnov-Grubbs method was used to test the control group for outliers; otherwise, all statistical tests utilized statistical software (SAS). The linearity of the GC-C-IRMS response was assessed by least-squares linear regression. The response was considered linear if the slope was zero at $P = 0.01$. The method of Koch and Peters *(17)* was used to determine the SDs of duplicates. The normality of the distributions of 5αA, 5βA, and 5βP, differences, and ratios was determined by the Anderson-Darling test. The values for 5αA, 5βA, and 5βP were correlated with Pearson's correlation coefficient, *r*. Two-sided paired *t*-tests were used to compare the means of 5αA, 5βA, and 5βP in the control group. The general linear model procedure was used to assess differences between the mean $\delta^{13}C$ values of the ethnic groups, and the power of the analysis was assessed by one-way ANOVA.

Results

linearity of GC-C-IRMS response

The IRMS response was linear for 5β- and 5α-androstane-3α,17β-diyl diacetate from 2.5 to 150 mg/L (Fig. 1). However, for 5β-pregnane-3α,20α-diyl diacetate, the response was linear only between 5 and 150 mg/L. The concentrations of 5β- and 5α-androstane-3α,17β-diol in the control group were 23–430 and 25–204 μg/L, respectively. Therefore, assuming 85% recovery, the amounts of 5β- and 5α-androstane-3α,17β-diol extracted from the 10-mL urine and reconstituted in 25 μL of cyclohexane were analyzed in the linear range of the IRMS.

precision of the GC-C-IRMS instrument and assay

Instrument precision (SD) for 5βA in QC-H and QC-L was 0.32‰ and 0.41‰, respectively. For 5αA, it was 0.08‰

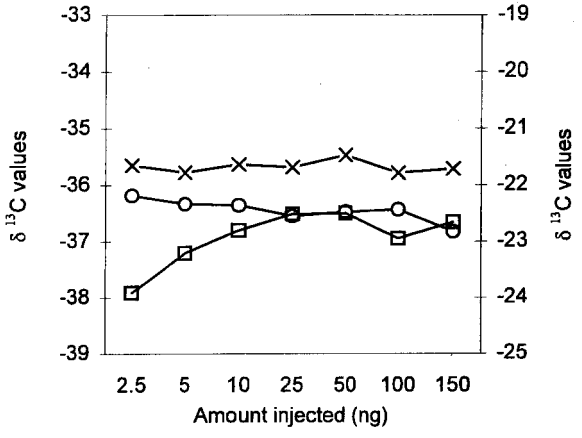

Fig. 1. Relationships between concentration of 5β-androstane-3α,17β-diyl diacetate (×), 5α-androstane-3α,17β-diyl diacetate (○), 5β-pregnane-3α,20α-diyl diacetate (□; *right axis*) and $\delta^{13}C$ values.

and 0.59‰, respectively, and for 5β-P, it was 0.27‰ and 0.16‰, respectively. The descriptive statistics for the within-assay experiment on QC-H are shown in Table 1. The SDs were 0.27‰, 0.38‰, and 0.28‰ for 5β-A, 5α-A, and 5β-P, respectively, and the CVs were ≤1.4%. The range of values was 0.9‰ for 5β-A, 1.2‰ for 5β-P, and 1.8‰ for 5α-A. The between-assay SDs for QC-H were 0.40‰, 0.42‰, and 0.44‰ for 5β-A, 5α-A, and 5β-P, respectively, and the CVs were ≤1.8% (Table 2). For QC-L, the values were slightly higher. The mean $\delta^{13}C$ values for the three steroids in QC-L were significantly lower than the means for QC-H.

System suitability and batch acceptance

In the system suitability test, the SDs of 5β-A, 5α-A, and 5β-P were 0.31‰, 0.59‰, and 0.54‰, respectively. The tolerance range for the batch acceptance data provided SDs for 5β-A, 5α-A, and 5β-P of 0.68‰, 0.67‰, and 0.65‰ for QC-H, and 0.66‰, 0.56‰, and 0.39‰ for QC-L, respectively. Because these data consisted of duplicate analyses, we calculated the SDs of duplicates and performed an ANOVA. The SDs of duplicates were 0.24‰, 0.32‰, and 0.37‰ for 5β-A, 5α-A, and 5β-P, respectively. These values were approximately one-half the SDs obtained for the factor injection. ANOVA revealed that QC-H and QC-L were different, and there was no difference for the factors injection, duplicate, or day.

Control subjects: descriptive statistics for the concentrations and $\delta^{13}C$ of urinary steroids

The geometric means for urine testosterone, epitestosterone, and 5β- and 5α-androstane-3α,17β-diol concentrations were 20, 29, 98, and 58 μg/L, respectively. The geometric mean for urine T/E was 0.7 (arithmetic range, 0.2–5.3). The concentration of urine 5β-pregnane-3α,20α-diol was not determined. Table 3 and Fig. 2 show the descriptive statistics and the histograms for 5β-A, 5α-A, and 5β-P. The outlier test was applied, and no data points were excluded. The distributions of 5α-A and 5β-P were gaussian ($A^2 = 0.28$, $P > 0.25$; and $A^2 = 0.21$, $P > 0.25$, respectively). The distribution of 5β-A was gaussian at $P = 0.01$, but not gaussian at $P = 0.05$, $A^2 = 0.85$.

There were significant ($P < 0.0001$) correlations between 5β-A and 5α-A ($r = 0.70$), between 5β-A and 5β-P ($r = 0.67$), and between 5β-P and 5α-A ($r = 0.58$). The CVs of the $\delta^{13}C$ values for the diacetates were 2.6–3.6%. The means − 3 SD for 5β-A, 5α-A, and 5β-P were −28.46‰, −28.39‰, and −26.37‰, respectively. The mean 5α-A was lower (−26.35‰) than the mean 5β-A (−25.69‰; $P < 0.001$). The Wilcoxon nonparametric two-sample test confirmed that 5α-A was lower than 5β-A. The means for 5β-A and 5α-A were lower than the mean for 5β-P (both $P < 0.001$).

Table 3 also summarizes the statistics of the differences between the means of 5β-P and 5β-A (5β-P − 5β-A) and 5β-P and 5α-A (5β-P − 5α-A), and the ratios 5β-A/5β-P and 5α-A/5β-P. The distributions of 5β-P − 5β-A and 5β-P − 5α-A were gaussian ($A^2 = 0.23$ and 0.19, respectively). The 5β-P − 5β-A differences ranged from −0.08‰ to 3.17‰, and the mean + 3 SD was 3.47‰. The 5β-P − 5α-A differences ranged from 0.16‰ to 3.72‰, and the mean + 3 SD was 3.99‰. The distributions of 5β-A/5β-P and 5α-A/5β-P were gaussian ($A^2 = 0.19$ and 0.21, respectively). The means of 5β-A/5β-P and 5α-A/5β-P were 1.06 and 1.09, respectively, and their corresponding means + 3 SD were 1.14 and 1.17. There were no differences between ethnic groups for the means of 5α-A, 5β-A, or 5β-P (Table 4).

Table 1. Within-assay precision for 5β-A, 5α-A, and 5β-P.

	QC-H		
	5β-A	5α-A	5β-P
Mean, ‰ (n = 20)	−24.46	−26.83	−23.13
SD, ‰	0.27	0.38	0.28
CV, %	1.1	1.4	1.2
Minimum, ‰	−24.90	−27.81	−23.62
Maximum, ‰	−23.96	−26.05	−22.47
Range, ‰	0.9	1.8	1.2

$\delta^{13}C$ values in athletes with elevated T/E ratios

The urine concentrations of testosterone and epitestosterone, the T/E ratios, and the $\delta^{13}C$ values for the six male athletes with urine T/E ratios >6 are summarized in Table 5. Of the six, athletes 1 and 2 were known to be taking testosterone for medical reasons. According to the sport organizations ordering the tests, athletes 3–6 denied taking any substance known to influence the T/E ratio. The number of samples per athlete ranged from three to eight. The T/E ratio of all samples was >6 for four of the six athletes. Athlete 4 had one T/E ratio of 5.0, and athlete 5 had one T/E ratio of 5.1; otherwise, all T/E ratios were >6.

Table 5 reveals that for athletes 1–3, all 5β-A and 5α-A values were lower than the mean − 3 SD values of the controls in Table 3, whereas the 5β-P values were all close to the mean 5β-P value of the controls (−24.26‰). The cells in Table 5 that are outside the ± 3 SD values of Table 4 are outlined or shaded to facilitate comparisons. The combination of 5β-P values that are near the mean and low 5β-A and 5α-A values produce large differences in 5β-P −

Table 2. Between-assay precision for 5β-A, 5α-A, and 5β-P.

	QC-H			QC-L		
	5β-A	5α-A	5β-P	5β-A	5α-A	5β-P
Mean, ‰	−24.72	−26.11	−24.00	−26.99	−27.57	−25.97
SD, ‰	0.40	0.42	0.44	0.50	0.47	0.52
CV, %	1.6	1.6	1.8	1.8	1.7	2.0
Minimum, ‰	−25.43	−27.19	−24.47	−27.65	−28.20	−26.86
Maximum, ‰	−24.06	−25.44	−22.89	−25.69	−26.55	−24.85
Range, ‰	1.4	1.8	1.6	2.0	1.7	2.0
n	15	14	16	15	14	16

Table 3. Control group of 73 subjects: Descriptive statistics for $\delta^{13}C$, differences between the means, and ratios of the means.

	$\delta^{13}C$, ‰			Difference, ‰		Ratio	
	5βA	5αA[a]	5βP[b]	5βP − 5βA	5βP − 5αA	5βA/5βP	5αA/5βP
Mean	−25.69	−26.35	−24.26	1.43	2.09	1.06	1.09
SD	0.92	0.68	0.70	0.68	0.63	0.028	0.027
CV, %	3.6	2.6	2.9			2.7	2.5
Mean + 3 SD	−22.92	−24.31	−22.15	3.47	3.99	1.14	1.17
Mean − 3 SD	−28.46	−28.39	−26.37	−0.62	0.18	0.97	1.00
Maximum	−23.90	−24.55	−22.92	3.17	3.72	1.13	1.16
Minimum	−27.82	−27.89	−25.49	−0.08	0.16	1.00	1.01
Max − Min	3.9	3.3	2.6				

[a] Mean significantly different from 5βA.
[b] Mean significantly different from 5βA and 5αA.

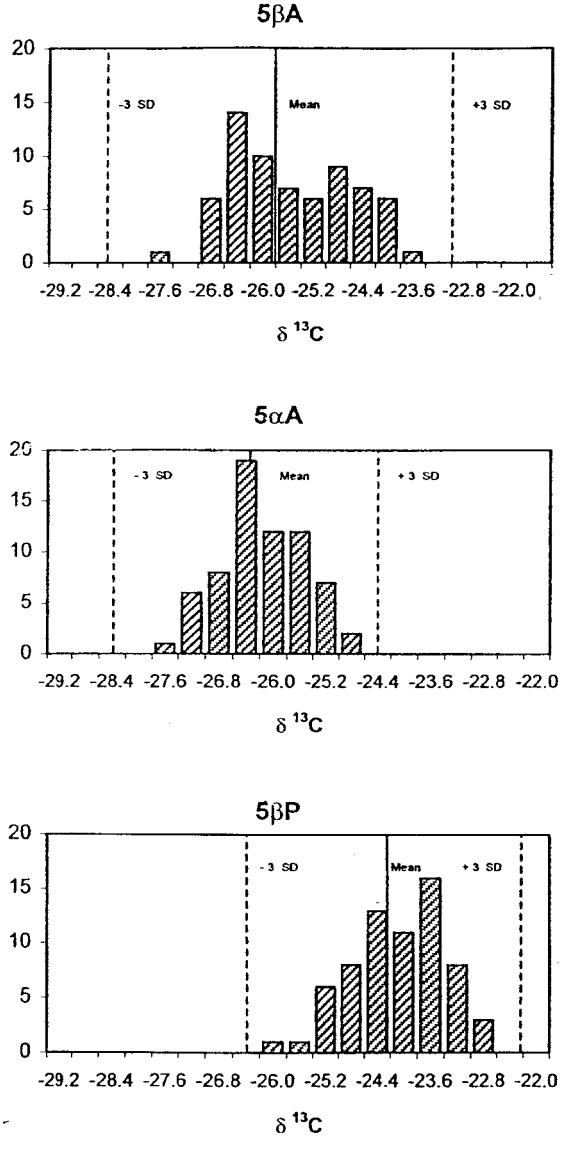

Fig. 2. Histograms of $\delta^{13}C$ values for 5βA (*top*), 5αA (*middle*), and 5βP (*bottom*) for a control group of 73 healthy males.

5βA and 5βP − 5αA, and large 5βA/5βP and 5αA/5βP ratios. For athletes 1–3, all differences and ratios were outside the + 3 SD range. The means of athletes 1–3 and 4–6 and their *z*-scores are presented at the bottom of Table 5. Another aspect of athletes 1–3 was that all 5αA values were lower than the 5βA values for the same urine: the means were −29.84‰ for 5βA and −31.83‰ for 5αA ($P = 0.04$). The means of 5βP − 5βA (5.9‰; *z*-score = 6.5) and of 5βP − 5αA (7.9‰; *z*-score = 9.1) were substantially greater than the + 3 SD values determined for the control group. Similarly, the mean 5βA/5βP ratio was 1.25 (*z*-score = 6.6), and the mean 5αA/5βP ratio was 1.33 (*z*-score = 8.9).

To compare the testosterone concentrations in Table 5 to the general population of athletes that we routinely test, we determined the log mean (3.22 μg/L) and log SD (1.19) of the distribution of the latest 11 938 male urine samples tested in our laboratory. The urine testosterone concentrations in subjects 1 and 2, the two permitted testosterone users, ranged from a low of 2.3 μg/L on the first urine ever collected to 186 μg/L (*z*-score = 1.5, in log

Table 4. Control group: Descriptive statistics for 5βA, 5αA, and 5βP by ethnicity.

Steroid	Ethnicity	n	Mean, ‰	SD, ‰	CV, %
5βA	Cauc.[a]	25	−25.74	0.95	3.7
	Asian	15	−26.17	0.66	2.5
	Hispanic	13	−25.02	1.03	4.1
	A.A.	9	−25.56	0.82	3.2
5αA	Cauc.	25	−26.43	0.79	3.0
	Asian	15	−26.60	0.49	1.8
	Hispanic	13	−26.06	0.73	2.8
	A.A.	9	−26.25	0.79	3.0
5βP	Cauc.	25	−24.34	0.70	2.9
	Asian	15	−24.19	0.68	2.8
	Hispanic	13	−24.14	0.92	3.8
	A.A.	9	−24.22	0.70	2.9

[a] Cauc., Caucasian; A.A., African American.

Table 5. Concentrations of testosterone and epitestosterone; T/E ratios; $\delta^{13}C$ values for 5βA, 5αA, and 5βP; difference scores between 5βP and testosterone metabolites (5βA and 5αA); and ratios of metabolites to 5βP (5βA/5βP and 5αA/5βP) for multiple urine samples obtained from two athletes permitted to take testosterone for medical purposes and four athletes whose drug-taking behavior was unknown.

					$\delta^{13}C$[c]			Difference,[c] ‰		Ratios[c]	
Athlete	Day[a]	[T],[b] μg/L	[E], μg/L	T/E	5βA	5αA	5βP	5βP − 5βA	5βP − 5αA	5βA/5βP	5αA/5βP
1	0	2.3	~0.1	20[d]	NA[e]	NA[e]	NA[e]				—
	1	40	1.0	40	−30.42	−31.96	−25.67	4.8	6.3	1.19	1.25
	4	186	2.6	72	−31.10	−31.76	−24.57	6.5	7.2	1.27	1.29
2	0	27	0.9	29	−31.43	−34.57	−26.14	5.3	8.4	1.20	1.32
	28	100	5.1	20	−28.56	−31.04	−24.00	4.6	7.0	1.19	1.29
	63	56	5.7	9.8	−28.66	−30.76	−22.97	5.7	7.8	1.25	1.34
3	0	343	4.3	80	−28.76	−31.25	−23.06	5.7	8.2	1.25	1.36
	56	690	8.9	77	−28.64	−31.10	−22.60	6.0	8.5	1.27	1.38
	70	598	13.7	44	−29.54	−32.30	−22.30	7.2	10.0	1.32	1.45
	156	820	14.6	56	−29.84	−31.33	−21.67	8.2	9.7	1.38	1.45
4	0	312	29.9	10.4							
	273	128	0.5	10.9	−25.32	−25.76	−24.54	0.8	1.2	1.03	1.05
	357	197	20.4	9.7	−24.96	−26.07	−23.43	1.5	2.6	1.07	1.11
	393	166	24.6	6.8							
	428	394	41.3	9.5	−23.55	−26.14	−22.47	1.1	3.7	1.05	1.16
	461	81	16.0	5.0							
	592	383	43.0	8.9							
	656	211	24.8	8.5							
5	0	108	8.6	12.6							
	64	33	3.5	9.4	−24.82	−25.47	−23.49	1.3	2.0	1.06	1.08
	345	130	15.0	8.6	−24.69	−25.52	−23.27	1.4	2.3	1.06	1.10
	412	36	3.7	9.8	−26.91	−24.50	−24.85	2.1	−0.4	1.08	0.99
	436	35	3.8	9.1							
	798	59	11.4	5.1	−26.66	−27.29	−24.77	1.9	2.5	1.08	1.10
6	0	49	6.3	7.8							
	38	46	6.5	7.1	−24.62	−26.04	−23.36	1.3	2.7	1.05	1.11
	73	11	1.3	8.4							
	103	92	10.3	9.0	−25.12	−25.04	−23.52	1.6	1.5	1.07	1.06
Means, athletes 1–3					−29.84	−31.83	−23.97	5.9	7.9	1.25	1.33
Means, athletes 4–6					−25.08	−25.74	−23.67	1.4	2.1	1.06	1.09
z-Scores, athletes 1–3					4.5	8.0		6.5	9.1	6.6	8.9
z-Scores, athletes 4–6					0.7	0.9		−0.02	−0.02	0.30	0.40

[a] Day of urine collection relative to the first sample (day = 0).
[b] [T] and [E], testosterone and epitestosterone.
[c] Cells >3 SD, shaded; cells <3 SD, solid outline.
[d] Value is an approximation because of large error in measurement of epitestosterone.
[e] Unable to measure because of low response (<0.3 V).

units). For subject 3, the urine testosterone concentrations (mean, 613 μg/L; *z*-score = 2.6) and T/E ratios (mean = 64) were very high. In addition, like athletes 1 and 2, athlete 3 was characterized by low 5βA and 5αA values, and large differences and ratios outside the ± 3 SD range.

In contrast to athletes 1–3, for athletes 4–6, the values for 5βA, 5αA, and 5βP and their differences and ratios were similar to the averages for the control group, and none of the values were remarkable except for one 5βP − 5αA value of −0.4‰ and one 5αA/5βP ratio of 0.99, which were slightly lower than the mean − 3 SD. For subjects 4–6, 5βP − 5βA and 5βP − 5αA were less than 2.1‰ and 3.7‰, respectively, i.e., within ± 3 SD of the control group values. Similarly, all but one of the ratios to 5βP were inside the ± 3 SD range. In addition, the urine testosterone concentrations for subjects 5 and 6 were within + 1.5 SD of the log mean for 11 938 athletes. Thus, compared with the control group, athletes 5 and 6 were characterized by elevated T/E ratios; whereas the values for 5βA, 5αA, and 5βP, their differences, and their ratios for athletes 5 and 6 were, with two minor exceptions, within ± 3 SD of the means of the controls. Athlete 4 had a relatively high mean urine testosterone (234 μg/L; *z*-score = 1.9).

The use of testosterone or testosterone precursors is indicated by low 5αA and 5βA values, large 5βP − 5αA

and 5βP − 5βA differences, and large 5αA/5βP and 5βA/5βP ratios. Of these six variables (two $\delta^{13}C$ values, two differences, and two ratios), for athletes 1–3, the ratio 5βA/5βP (*z*-score = 8.9) and the difference 5βP − 5βA (*z*-score = 9.1) were the most sensitive indicators of testosterone administration as judged by *z*-scores. The value for 5αA was also a good indicator (*z*-score = 8.0). For differences and ratios that utilized 5βA (5βP − 5βA and 5βA/5βP), the *z*-scores were 6.5 and 4.5, respectively. The least discriminating variable was 5βA (*z*-score = 4.5). In contrast, for athletes 4–6 in Table 5, the *z*-scores were all low (range, −0.02 to 0.9).

Discussion

CHARACTERISTICS OF THE GC-C-IRMS ASSAY

GC-C-IRMS measurement of urinary steroids has become a potent analytical method in doping control, although scrupulous attention to detail is necessary. Successful implementation of the method requires strict adherence to QC, system suitability, and batch acceptance criteria such as those described herein. In the present study, after the tolerance ranges for the system and batches were established, the system suitability test never failed and the batch acceptance criteria were always met over the 14 months of the study.

In the instrument precision study, the mean SDs were less than ± 0.33‰; thus, when the variability attributable to extraction and derivatization was eliminated, all of the SDs were low. In the within-assay study, the SDs were also low: ± 0.27‰, ± 0.38‰, and ± 0.28‰ for 5βA, 5αA, and 5βP, respectively. The between-assay SDs ranged from ± 0.40‰ to ± 0.52‰, which is slightly higher than the within-assay values, but they indicate the excellent repeatability and reproducibility of the entire assay.

We did not find comparable precision studies for urinary steroids in the literature; however, Shackleton et al. *(3)* provided evidence that $\delta^{13}C$ values for diols are reproducible based on duplicate analysis of 13 samples from one subject. We used the data in Table 1 from Shackleton et al. *(5)* to calculate the SD of duplicates and found values of ± 0.36‰, ± 0.15‰, and ± 0.28‰ for 5βA, 5αA, and 5βP, respectively. These data show that the SD for duplicates in another laboratory was similar to the SD that we obtained for duplicates in our batch acceptance studies. The concentrations of 5β-androstane-3α,17β-diol and 5α-androstane-3α,17β-diol in the control urines were 23–430 μg/L; thus, given the 85% recovery of steroids from the 10-mL sample volume used in the analysis, the concentrations of 5β-androstane-3α,17β-diol and 5α-androstane-3α,17β-diol in the 25 μL of the injection solution were 8–146 mg/L, and the samples were analyzed in the linear range of the system (Fig. 1).

HEALTHY CONTROL MALE SUBJECTS

The $\delta^{13}C$ values in Table 3 represent the largest group of healthy males living in the US measured to date. The ranges for the T/E ratios (0.1–3.6) and urine testosterone concentrations (2–116 μg/L) are consistent with the fact that no student declared taking testosterone or a steroid supplement. In addition, none of the students declared any endocrine disorders. Of the 74 urines, we were able to measure 5βA, 5αA, and 5βP values for all but 1. Sixty-eight of the samples provided satisfactory data on the first extraction of 10 mL of urine. For the other six, three provided an adequate signal on repeat analysis. Two samples required a 20-mL sample volume to obtain an adequate signal. In one sample with low diol concentrations (<25 μg/L), the voltage criteria were not met even with a 20-mL sample. Thus, our signal criteria were met on the first analysis with 10 mL of urine in 94% of the 74 cases, and all but 1 of the remaining samples provided satisfactory data on repeat analysis with 10 or 20 mL of urine. One person with one instrument can perform approximately three assays per week with a batch size of 20 samples plus 5 controls. The software to process the samples is somewhat cumbersome; thus, ~2 days per week are needed for data processing.

The mean 5αA and 5βA values for our control group were −26.35‰ and −25.69‰, respectively. The mean δ values − 3 SD for 5βA and 5αA were approximately −28.4‰, which is similar to the δ values reported for synthetic testosterone *(6, 7)*; however, this comparison should not be made. This is because the diols measured herein were acetylated, which makes the δ value more negative by approximately two units, whereas the synthetic testosterone *(6, 7)* was underivatized.

There are no comparable control group values in the literature; however, the baseline values in Fig. 6 of Shackleton et al. *(3)*, which was based on analysis of 20 samples from individuals with mixed nationalities, are very similar to ours. Their lowest value for a diol was −28.2‰, whereas our lowest diol value was −27.89‰. Ueki and Okano *(7)* also measured 5αA and 5βA; however, their data are not directly comparable to ours because they used a correction factor to convert all measured $\delta^{13}C$ values of the acetylated compounds to give values for underivatized steroids, which we could not estimate. The correction factor would be the same for 5αA and 5βA; thus, their difference of 2.6‰ between the means of 5αA (−16.4‰) and 5βA (−19.0‰) for 10 Japanese male samples can be compared to our difference of 0.66‰ (Table 3). In addition, the value for 5αA was lower than 5βA in the study by Ueki and Okano *(7)*, which is the opposite of what we found. We can also compare the range of values in our control subjects to the range reported by Ueki and Okano *(7)* in their Table 2 for 20 healthy Japanese males. Our ranges for 5αA and 5βA were 2.9‰ and 3.9 ‰ (Table 4), whereas they found ranges of 10.8‰ and 7.2‰, respectively. The striking difference between the ranges found in the two laboratories could reflect differences in method, calibration, ethnicity, or diet. The explanation that Japanese subjects have different steroid biochemistry and metabolism was deemed unlikely because equally large ranges of 11.0‰ (5βA) and 14.1‰ (5αA) were

reported for urine samples from >350 Olympic athletes, and Olympic athletes are multiracial. In addition, Shackleton et al. *(3)* found similar values for Chinese and other nationalities, and we found no difference among four ethnic groups in our control group. That dietary factors could explain the differences was also considered, but diet is an unlikely explanation for the large differences in Olympic athletes because it is likely to take several weeks on a local diet before a measurable change in urinary steroid $\delta^{13}C$ develops. Therefore, it is likely that our analytical method differs from that of Ueki and Okano *(7)*.

The diversity of the control group was an advantage in that four major ethnic groups were present; however, it was a disadvantage because the n for each ethnic group was relatively small and therefore the statistical power of the GLM procedure was relatively low. Thus, we could not attribute any differences to ethnicity. Similarly, Shackleton et al. *(5)*, who compared $\delta^{13}C$ values for 20 individuals from 12 nationalities, did not attribute any of the differences to ethnicity or diet.

In the present study, the mean $\delta^{13}C$ values for 5αA (−26.35‰) and 5βA (−25.69‰) in the control group were different ($P < 0.001$), whereas in our previous study with only 10 subjects, we did not observe such a difference *(15)*. In addition, it appears from the data in Table 5 of the present study and from Table 1 in our previous study *(15)* that testosterone administration may decrease 5αA more than it decreases 5βA. This difference occurs despite the fact that the percentage of conversion of a tracer dose of [^{14}C]testosterone to 5β-androstane-3α,17β-diol (3.2%) is greater than the percentage of conversion to 5α-androstane-3α,17β-diol (1.2%) *(18)*. Urinary 5α-androstane-3α,17β-diol is known to arise from both hepatic and peripheral metabolism, whereas urinary 5β-androstane-3α,17β-diol is considered to arise only from hepatic metabolism *(19)*. After pharmaceutical doses of testosterone, 5αA may decrease more than 5βA because peripheral paths to 5α-androstane-3α,17β-diol are favored. However, there is also a path from dehydroepiandrosterone sulfate to 5β-androstane-3α,17β-diol that does not pass through testosterone *(20)*. Utilization of this path could explain why 5βA is higher than 5αA. Finally, the observed differences may be an analytical artifact. Similarly, the higher mean $\delta^{13}C$ value for 5βP (−24.26‰) might be related to metabolism or artifacts.

athletes with elevated T/E ratios

To clarify the status of an athlete with an elevated T/E ratio, the International Olympic Committee recommends that additional samples be obtained *(9)*. The relevant sport authority typically charts the course of the urine T/E values over time and interprets the data pattern to decide whether to penalize the athlete. Unless there is a record of three or more past samples, this involves collecting additional samples. There are no guidelines stating how many samples are needed, how often they should be collected, and for how long, and there are no published criteria for determining whether the elevated T/E ratio is natural or is attributable to taking an exogenous substance. The data in Tables 1–3 and 5 support the argument that by measuring $\delta^{13}C$, the decision to penalize the athlete can be made on a single sample. The six athletes described in Table 5 were undergoing such testing, although at the time of the analysis the laboratory was not aware that subjects 1 and 2 were permitted testosterone users. No additional information was available on the status of athletes 3–6; however, inspection of the data for subject 3 indicates that it is reasonable to classify him as a testosterone or testosterone precursor user.

The six subjects in Table 5 with increased T/E ratios may be characterized by the concentration of testosterone in their urines and by their 5βA and 5αA values and the associated differences and ratios to 5βP. A classification based on $\delta^{13}C$ values, differences, and ratios revealed that athletes 1–3 were outside the control range ± 3 SD except for 5βP. For athletes 1 and 2, this fits with our understanding of the basis of the carbon isotope ratio method, specifically, that after testosterone administration only the $\delta^{13}C$ values of testosterone and testosterone metabolites will decrease and that those of other urinary steroids will not. Because 5β-pregnane-3α,20α-diol is not a metabolite of testosterone, its $\delta^{13}C$ values are not expected to change with testosterone administration *(3, 4, 15)*.

Applying the means ± 3 SD criteria to the $\delta^{13}C$ data for athlete 3 requires classifying him as a testosterone or testosterone precursor user. This possibility is further supported by the very high urine T/E ratios and testosterone concentrations. The *z*-scores for his urine testosterone concentrations range between 2.2 and 2.9, which is a further indication of testosterone or testosterone precursor administration. Once the GC-C-IRMS method is validated to the point of full acceptance by the sport and legal community, we expect that athlete 3 could be classified as a user on one urine showing the pattern of a high T/E ratio, high testosterone concentration, and $\delta^{13}C$ values less than − 3 SD.

Athletes with naturally elevated urine T/E ratios would be expected to have persistently elevated T/E ratios, unremarkable urine testosterone concentrations, and values for 5βA and 5αA that are within ± 3 SD of the mean of the control group. The data for athletes 5 and 6 fit this pattern. Athlete 4 also fits, but some of his urine testosterone concentrations were relatively high. The *z*-score for the values 383 and 394 μg/L was ~2.3, and all but one of the others were >1.3. Until additional data are available, athlete 4 has been classified as having a naturally elevated T/E ratio, although we cannot completely exclude the possibility that athlete 4 is a testosterone or testosterone precursor user with normal $\delta^{13}C$ values.

The data on mean *z*-scores in Table 5 reveal that both the differences and ratios are better indices of testosterone use than the absolute $\delta^{13}C$ values, 5βA and 5αA. This is

not unexpected given that 5βP correlates with both 5βA and 5αA. Furthermore, the difference 5βP − 5βA and the ratio 5βA/5βP are more robust indicators of testosterone use than the corresponding differences 5βP − 5αA and the ratio 5αA/5βP. This follows from the finding (Tables 3 and 5) that the $\delta^{13}C$ values for 5αA were lower than the values for 5βA.

In conclusion, a fundamental issue in doping control is whether GC-C-IRMS techniques will resolve ambiguities in the interpretation of the T/E test. The methods presented herein provide compelling evidence that the test has excellent precision and that when it is coupled with strict system suitability and batch acceptance criteria, it can be used in a routine fashion to obtain valid $\delta^{13}C$ data. By comparing the $\delta^{13}C$ values of a control group to data obtained from athletes, it is possible to make informed decisions regarding the origin of urinary 5β-androstane-3α,17β-diol, 5α-androstane-3α,17β-diol, and 5β-pregnane-3α,20α-diol.

We are grateful for financial support from the National Collegiate Athletic Association, the National Football League, and the United States Olympic Committee. We thank K. Schramm for technical assistance.

References

1. Coppen JJW. Steroids: from plants to pills—the changing picture. Trop Sci 1979;21:125–41.
2. Smith BN, Epstein S. Two categories of $^{13}C/^{12}C$ ratios for higher plants. Plant Physiol 1971;47:380–4.
3. Shackleton CHL, Phillips A, Chang T, Li Y. Confirming testosterone administration by isotope ratio mass spectrometric analysis of urinary androstanediols. Steroids 1997;62:379–87.
4. Aguilera R, Becchi M, Casabianca H, Hatton CK, Catlin DH, Starcevic B, Pope HG. Improved method of detection of testosterone abuse by gas chromatography/combustion/isotope ratio mass spectrometry analysis of urinary steroids. J Mass Spectrom 1996;31:169–76.
5. Shackleton CH, Roitman E, Phillips A, Chang T. Androstanediol and 5-androstenediol profiling for detecting exogenously administered dihydrotestosterone, epitestosterone, and dehydroepiandrosterone: potential use in gas chromatography isotope ratio mass spectrometry. Steroids 1997;62:665–73.
6. Horning S, Geyer H, Flenker U, Schänzer W. Detection of exogenous steroids by $^{13}C/^{12}C$ analysis. In: Schänzer W, Geyer H, Gotzmann A, Mareck-Engelke U, eds. Recent advances in doping analysis, Vol. 5. Köln: Sport und Buch Strauss, 1997:135–48.
7. Ueki M, Okano M. Analysis of exogenous dehydroepiandrosterone excretion in urine by gas chromatography/combustion/isotope ratio mass spectrometry. Rapid Commun Mass Spectrom 1999; 13:2237–43.
8. Donike M, Bärwald KR, Klostermann K, Schänzer W, Zimmermann J. Nachweis von exogenem Testosteron. In: Heck H, Hollmann W, Liesen H, Rost R, eds. Sport: Leistung und Gesundheit. (Kongressbd. Deutsche. Sportarztekongress) Köln: Deutsche Ärzte-Verlag, 1983:293–8.
9. International Olympic Committee. List of doping classes and methods. Lausanne, Switzerland: IOC, 1992.
10. Falk O, Palonek E, Björkhem I. Effect of ethanol on the ratio between testosterone and epitestosterone in urine. Clin Chem 1988;34:1462–4.
11. Oftebro H. Evaluating an abnormal urinary steroid profile. Lancet 1992;339:941–2.
12. Catlin DH, Hatton CK, Starcevic S. Issues in detecting abuse of xenobiotic anabolic steroids and testosterone by analysis of athletes' urine. Clin Chem 1997;43:1280–8.
13. Dehennin L, Matsumoto AM. Long-term administration of testosterone enanthate to normal men: alterations of the urinary profile of androgen metabolites potentially useful for detection of testosterone misuse in sport. J Steroid Biochem Mol Biol 1993;44: 179–89.
14. Becchi M, Aguilera R, Farizon Y, Flament M-M, Casabianca H, James P. Gas chromatography/combustion/isotope-ratio mass spectrometry analysis of urinary steroids to detect misuse of testosterone in sport. Rapid Commun Mass Spectrom 1994;8: 304–8.
15. Aguilera R, Catlin DH, Becchi M, Phillips A, Wang C, Swerdloff RS, et al. Screening urine for exogenous testosterone by isotope ratio mass spectrometric analysis of one pregnanediol and two androstanediols. J Chromatogr B 1999;727:95–105.
16. Catlin DH, Kammerer RC, Hatton CK, Sekera MH, Merdink JM. Analytical chemistry at the Games of the XXIIIrd Olympiad in Los Angeles, 1984. Clin Chem 1987;33:319–27.
17. Koch DD, Peters TP. Selection and validation of statistical methods. In: Burtis CA, Ashwood ER, eds. Fundamentals of clinical chemistry, 2nd ed. Philadelphia: WB Saunders, 1996:170–81.
18. Baulieu EE, Robel P. Catabolism of testosterone and androstenedione. In: Eik-Nes KB, ed. The androgens of the testis. New York: Marcel Dekker, 1970:49–71.
19. Mauvais-Jarvis P, Floch H, Jung I, Robel P, Baulieu EE. Studies on testosterone metabolism. VI. Precursors of urinary androstanediols. Steroids 1968;11:207–24.
20. Mauvais-Jarvis P, Baulieu EE. Studies on testosterone metabolism. IV. Urinary 5α- and 5β-androstanediols and testosterone glucuronide from testosterone and dehydroisoandrosterone sulfate in normal people and hirsute women. J Clin Endocrinol Metab 1965;25:1167–78.