Structure determination of a new fluorescent tricyclic nucleoside from archaeobacterial tRNA

James A. McCloskey, Pamela F. Crain, Charles G. Edmonds, Ramesh Gupta¹, Takeshi Hashizume, Douglas W. Phillipson and Karl O. Stetter²

Departments of Medicinal Chemistry and Biochemistry, University of Utah, Salt Lake City, UT 84112, ¹Department of Genetics and Development, University of Illinois, Urbana, IL 61801 and Department of Medical Biochemistry, Southern Illinois University, School of Medicine, Carbondale, IL 62901, USA and ²Lehrstuhl für Microbiologie, Universität Regensburg, Regensburg, FRG

Received December 22, 1986; Accepted January 2, 1987

ABSTRACT

A highly fluorescent nucleoside was detected in enzymatic digests of the extremely thermophilic archaeobacterium *Sulfolobus solfataricus* by combined liquid chromatography-mass spectrometry (LC/MS). Following isolation, the structure was determined primarily by mass spectrometry, to be 3-(β -D-ribofuranosyl)-4,9-dihydro-4,6,7-trimethyl-9-oxoimidazo[1,2-a]purine (mimG), a new derivative of the Y (wye) nucleoside. The structural assignment was verified by comparison of the base released by acid hydrolysis with the corresponding synthetic base, using mass spectrometry, chromatography, and UV absorption and fluorescence properties. Nucleoside mimG was also detected by LC/MS in hydrolysates of the thermophiles *Thermoproteus neutrophilus* and *Pyrodicticum occultum*. These results constitute the first finding of a member of the hypermodified Y family of nucleosides in archaeobacteria.

INTRODUCTION

Post-transcriptional processing of tRNA produces several structurally distinct families of hypermodified nucleosides which generally require multi-step enzymatic processing for their formation and lead to relatively complex base modifications. The five principal groups of such nucleosides are those related to Q (queuosine) and Y (wye nucleosides), 2-thio-5-substituted uridines (s^2x^5U), and N^6 -substituted adenosines with isopentenyl or threonyl side chains. The tRNA of archaeobacteria, as phylogenetically distinguished from the other two primary kingdoms, the eukaryotes and eubacteria, have to date been known to exhibit rather simple modification patterns (1), with no known occurrences of the Y, Q or s^2x^5U nucleosides. We report the structural characterization of a new and highly fluorescent member of the tricyclic Y nucleoside family, 3-(β -D-ribofuranosyl)-4,9-dihydro-4,6,7-trimethyl-9-oxoimidazo[1,2-a]purine (structure 1), found in unfractionated tRNA from three extremely thermophilic archaeobacteria, *Sulfolobus solfataricus*, *Thermoproteus neutrophilus* and *Pyrodicticum occultum*. Nucleoside nomenclature used follows the recommendations in ref. (2).

EXPERIMENTAL

Organisms and Isolation of tRNA

The isolation, culture conditions and microbiological characterization of *Sulfolobus solfataricus* (optimum growth temperature 87 °C) (3), *Thermoproteus neutrophilus* (88 °C) (1) and *Pyrodicticum occultum* (105 °C) (5) have been previously reported. tRNA was isolated as earlier reported (6).

Reversed-Phase HPLC of Nucleoside Mixtures from Digests of tRNA

Approximately 2 A₂₆₀ units were digested to nucleosides using nuclease P₁ and alkaline phosphatase as previously described (7). Typically, ~1 A₂₆₀ unit was used for HPLC (Beckman model 332 chromatograph with model 420 controller) analyses on 250 x 4.6 mm LC-18S columns (Supelco, Inc.; Bellefonte PA), maintained at 31 °C using a Waters TCM (Millipore Corp.; Milford MA) column oven. Gradient elution was performed using the system of Buck, *et al.* (8) (solvent system 1), which employs 0.25 M ammonium acetate, pH 6.0 (solvent A) and acetonitrile/H₂O, 40:60 (solvent B). UV absorbance was monitored at 254 and 280 nm using a Waters (Milford MA) model 440 detector. A Gilson (Middleton WI) Spectra/glo filter fluorometer was used for fluorescence detection of *S. solfataricus* hydrolysates (excitation at 380-450 nm; emission at 313 nm). The detectors were connected in series.

Directly-Combined Liquid Chromatography-Mass Spectrometry of tRNA Digests

The liquid chromatograph-mass spectrometer, and its application to the direct analysis of mixtures of nucleosides from enzymatic hydrolysates of tRNA has been described (7). The instrument was operated to accept a flow rate from the liquid chromatograph of 2 ml/min. Vaporizer power was adjusted according to thermospray ionization response to maintain vaporizer exit temperatures in the range of 220-240 °C and vapor temperature at the point of ion sampling of 280-285 °C. Ion source temperature was 340 °C. The HPLC column and elution conditions are described above.

Preparative HPLC of *S. solfataricus* tRNA and UV Absorbance of Unknown Nucleoside

Isolation of nucleoside fractions was carried out using a semi-

preparative 250 x 10 mm LC-18DB column (Supelco); gradient elution was accomplished using solvent system 1, adjusted for the higher flow rates required (9). A ~25 mg sample of tRNA from *S. solfataricus* was digested by scaling up the amount of enzymes used for tRNA digestion for analytical HPLC. Three separate runs were made, and peaks corresponding to an unknown nucleoside designated M349 were pooled and dried in a vacuum centrifuge. The UV spectrum of M349 in water showed a maximum at 239 nm. Fractions were stored at -20 °C until used.

Preparation of the Trimethylsilyl Derivative of Unknown Nucleoside M349

A portion of the fraction corresponding to nucleoside M349 (~5 ug) was dried in vacuo and treated with 25 ul of a mixture of Sylon BFT (N,O-bis(trimethylsilyl)trifluoroacetamide [BSTFA] + 1% trimethylchlorosilane; Supelco, Inc.; Bellefonte PA) and dry pyridine (2:1) for 1 hr at 100°C. Reagents were removed in the direct inlet vacuum port prior to sample insertion for acquisition of mass spectra.

Electron Ionization Mass Spectrometry of Trimethylsilylated M349

Mass spectra were acquired using a MAT 731 instrument operated at 70 eV ionizing energy and 8 kV accelerating potential. Ion source temperature was 270 °C, and the trimethylsilyl (Me₃Si) derivative of M349 vaporized from the sample probe at 200 °C. Low resolution spectra were acquired at resolution (R) = 1,000. Determination of the elemental composition of M349 was accomplished by high resolution (R = 1000) peak-matching of the molecular ion (m/z 565) against the m/z 555 ion from perfluorokerosene (PFK-H; SCM Specialty Chemicals Gainesville FL) used as the reference mass.

Fast Atom Bombardment (FAB) Mass Spectrometry of M349

Mass spectra were recorded using a MAT 731 instrument, with an Ion Tech 11N FAB ion source operated at 7 keV with Xe. Nucleoside M349 was dissolved in glycerol matrix at a concentration of approximately 1 ug/ul.

Formic Acid Hydrolysis of M349

M349 (1-5 ug) was dried under vacuum and hydrolyzed with 15 ul of 0.1M formic acid (pH = 2.5) for 15 hours at room temperature (Scheme 1). Formic acid was removed in vacuo, and the digest was used without further purification for HPLC and GC/MS. The heterocyclic base released by acid hydrolysis is referred to as M217 in further discussion.

HPLC Comparison of Natural and Synthetic Bases

A Beckman model 332 chromatograph with model 420 controller, and Waters model 440 and Perkin-Elmer (Norwalk CT) model 650-10S fluorescence spectrophotometers (connected in series) were used for chromatographic and spectro-

Scheme 1

scopic comparisons of M217 and synthetic mimGua. Ultraviolet absorbance was monitored at 254 and 280 nm; excitation and emission wavelengths were 320 and 430 nm, respectively, selected on the basis of Y fluorescent properties (10). Approximately 0.5 μ g of M217 and mimGua were analyzed using LC-18DB columns with isocratic elution using solvents A and B described for HPLC (above) in a combination of 40% B in A + B at a flow rate of 2 ml/min (solvent system 2).

Preparation of Trimethylsilyl Derivatives of Bases M217 and mimGua for GC/MS

All samples for GC/MS analysis were transferred to reaction tubes fashioned from 4 mm i.d. Pyrex tubing (11), and dried in a vacuum centrifuge. When dry, each sample was treated with 5 μ l of dry pyridine and re-dried. Ten μ l of derivatizing solution was prepared (by mixing 100 μ l of Sylon BFT and 10 μ l of dry pyridine) and added to each sample tube which was then flame-sealed and heated in a 100 $^{\circ}$ C oven for 1 hr. Samples treated by this procedure were a weighed sample (1.15 μ g) of synthetic mimGua, a quantity of unknown base M217 produced by acid hydrolysis of approximately 1 μ g of nucleoside M349, and a blank corresponding to the amount of formic acid used for hydrolysis.

Gas Chromatography-Mass Spectrometry for Comparison of Natural and Synthetic Bases

A VG Analytical 70SEQ mass spectrometer with 11-250J data system was used to acquire mass spectra of Me_3Si derivatives introduced from a directly-coupled Hewlett Packard Model 5890 gas chromatograph equipped with a 25 m OV-1 bonded-phase fused silica capillary column. The mass spectrometer was operated under the following conditions: 70 eV ionizing energy; 8 kV accelerating potential; ion source and inlet transfer line temperatures 290 $^{\circ}$ C; $R = 1,000$. The gas chromatograph was operated at a He flow rate of 2 ml/min with the injector at 270 $^{\circ}$ C. A splitless injection mode was used for introduction of approximately 25 nanograms of each sample. The GC column was pro-

grammed from 150 °C to 280 °C at a rate of 5 °/min. Data acquisition was begun 4.0 min after sample injection; mass spectra were acquired over the mass range 65-350 Da at a rate of 1.5 seconds/decade with 0.5 sec interscan delay.

RESULTS

Reversed-Phase HPLC of tRNA Digests

Nucleosides from nuclease P₁ / alkaline phosphatase digests of unfractionated tRNAs from *S. solfataricus*, *I. neutrophilus* and *P. occultum* were examined by HPLC using UV and fluorescence detection. A strongly-fluorescent and previously unreported constituent was observed to elute at ~31.1 minutes.

Liquid Chromatography-Mass Spectrometry of Nucleosides from tRNA Digests

tRNA digests were examined by LC/MS and the component eluting at 31.1 minutes produced a relatively simple thermospray mass spectrum consisting of m/z 350 (1% relative intensity) and m/z 218 (100%). These ions were assigned as the protonated molecule, MH^+ , and protonated base BH_2^+ , respectively; the 132 Da difference is characteristic (7) of ribonucleosides not methylated in the sugar. The value $M_r = 349$ was judged not to correspond to any previously-known nucleoside from RNA. This new nucleoside was designated M349.

Mass Spectrometry of the Unknown Nucleoside M349

The electron ionization mass spectrum of trimethylsilylated M349 is readily interpretable from fragmentation patterns characteristic of trimethylsilylated ribonucleosides (12). As shown in Fig. 1, the ion of highest mass occurs at m/z 565 and is assigned as the molecular ion, supported by presence of an accompanying ion at m/z 550 due to characteristic loss of a

Figure 1. Electron ionization mass spectrum of trimethylsilylated nucleoside M349.

Table 1: Diagnostic base-containing ions from the mass spectrum of trimethylsilylated nucleoside M349

<u>m/z</u>	<u>ion^a</u>	<u>composition</u>
202	B - 14	B - CH ₂
216	B	B
217	B + 1	B + H
274	B + 58	B + C ₂ H ₆ Si
289	B + 73	B + C ₃ H ₉ Si
316	B + 100	B + C ₄ H ₈ OSi
331	B + 115	B + C ₅ H ₁₁ OSi
332	B + 116	B + C ₅ H ₁₂ OSi
404	B + 188	B + C ₈ H ₂₀ OSi

^a Ions are designated as in ref. (12).

methyl radical. The difference in mass between the derivatized (565 Da) and free (349 Da) molecules is 216 Da, corresponding to the incorporation of three Me₃Si blocking groups into M349. Ions of m/z 348, 259, 245, and 243 arise from the presence of a fully-trimethylsilylated unmodified ribose moiety, which therefore accounts for all three Me₃Si groups in the molecule. Ions consisting of the heterocyclic base with associated portions of the ribose moiety are listed in Table 1.

The FAB mass spectrum of M349 showed two principal peaks: m/z 350 (52% relative intensity), assigned as the protonated molecule, MH⁺, and m/z 218 (100%), assigned as the protonated base.

Determination of the Elemental Composition of M349

The exact molecular mass of silylated M349 was measured as 565.2566. Using relatively broad limits for mass error (± 10 ppm), and elemental composition (C₂₅H₄₅N₇O₇SSi₃), two computer-derived (13) composition candidates were obtained, C₁₉H₄₃N₇O₇Si₃ and C₂₄H₄₃N₅O₅Si₃.

HPLC Comparison of Bases M217 and mimGua Using UV and Fluorescence Detection

Both M217 and mimGua have identical retention times of 7.5 minutes under the HPLC conditions used, as shown in Figure 2. The peak ratio of UV absorbance at 254 nm to that at 280 nm is 2.3 for M217 and 2.2 for mimGua. The ratio of fluorescence to absorbance at 254 nm is 10.8 for M217 and 11.0 for mimGua, measured under the same chromatographic conditions and instrumental conditions.

Figure 2. Reversed-phase chromatography (solvent system 2) with UV and fluorescence detection of: a) base M217; b) authentic mimGua. Sensitivity (full-scale absorbance units) of the 254 nm and 280 nm channels was 0.05 and 0.3 for the fluorescence channel.

Gas Chromatography-Mass Spectrometry of Bases M217 and mimGua

The heterocyclic base moiety released from M349 by formic acid digestion (Scheme 1) was compared to synthetic mimGua by GC/MS of their Me_3Si derivatives. The total ion current chromatograms are shown as Figs. 3a and 4a, and reveal that the components of interest are chromatographically indistinguishable, eluting within 0.06 minutes of each other at retention times of 28.2 minutes. The mass spectra of the major component of each sample are shown in Figs. 3b and 4b; the molecular ion is observed in each case at m/z 289, corresponding to the incorporation of a single Me_3Si group into the molecule (in lieu of ribose), as expected from the mass spectrum of the derivatized intact nucleoside. Other major ions correspond to the characteristic losses of a methyl radical to give m/z 274, and loss of a Me_3Si radical to give m/z 216 (14). The ion at m/z 73 corresponds to the Me_3Si ion; a set of minor ions is present to the same extent in both spectra. Gas chromatographic peaks in the chromatogram for M217 (Fig. 3a), eluting at approximately 12 minutes, are judged from their spectra (not shown) to

Figure 3 (upper panel). a) Gas chromatogram of trimethylsilylated base M217, reconstructed from total ion current; b) mass spectrum of the Me₃Si derivative of M217, which elutes at ~28.2 minutes.

Figure 4 (lower panel). a) Gas chromatogram of the Me₃Si derivative of authentic mimGua, reconstructed from total ion current; b) mass spectrum of the Me₃Si derivative of mimGua, which elutes at ~28.2 minutes.

represent the expected degradation products of the ribose moiety, and silyloxy-containing by-products. The peak eluting at ~11 minutes in the chromatogram of mimGua (Fig. 4a) is a minor silyloxy-containing impurity (mass spectrum not shown). The peak at ~23 minutes in both chromatograms (Figs. 3a and 4a) represents a component also present in the reagent blank (data not shown).

DISCUSSION

Examination of nuclease digests of unfractionated tRNA from the thermophilic archaeobacterium *S. solfataricus* by directly combined liquid chromatography-mass spectrometry revealed a fluorescent and relatively hydrophobic component eluting at 31.1 minutes which did not correspond to any known tRNA

nucleoside. The use of LC/MS in this fashion for analysis of unfractionated tRNA is highly advantageous because it permits detection of nanogram-levels of individual nucleosides, without reliance on UV detection, and in general, without chromatographic separation from other constituents (7). In addition, as in the present case, limited structural information such as molecular weight can be established in the same experiment. The nucleoside from *S. solfataricus* designated M349, was found from its thermospray mass spectrum to have a molecular weight of 349, a corresponding (free) base of $M_r = 217$, and unmethylated ribose. The absence of further fragmentation of the MH^+ or BH_2^+ ions strongly suggested absence of a (non-methyl) side chain.

Following isolation of M349 and trimethylsilylation, the electron ionization mass spectrum (Fig. 1) showed incorporation of three blocking groups in the sugar and none in the base, suggesting the absence of active hydrogen in the base moiety. In addition to the well-characterized series of base-containing ions (12) (Table 1), no evidence of a side chain was observed. The relative intensity pattern (Fig. 1) in which $M^+ > (M-CH_3)^+$ is highly characteristic of (although not restricted to) guanosine derivatives. The absence of methylation at N-7 of guanosine was specifically excluded by failure of the base to undergo oxidation at C-8 during silylation, a reaction characteristic of 7-methylpurine nucleosides (15). The correspondence in molecular weight values obtained by thermospray LC/MS, FAB and electron ionization mass spectrometry militates against degradation reactions occurring during LC/MS analysis or preparation of the silyl derivative.

From the exact molecular mass of the tri-Me₃Si derivative of M349, 565.2566 (corresponding to 349.1386 in the free nucleoside), two compositional candidates were obtained, of which C₂₄H₄₃N₅O₅Si₃ was concluded to be correct. The alternative, C₁₉H₄₃N₇O₇Si₃, was considered unlikely in terms of rational structures that could be written and the high heteroatom content of the base without sites for trimethylsilylation. The elemental composition thus derived by exact mass measurement, which corresponds to C₁₅H₁₉N₅O₅ in the unblocked nucleoside, requires a total of 9 rings and double bonds in the nucleoside, 8 in the base. The latter values, taken in conjunction with the probable equivalence of M349 with guanosine + C₅H₆, strongly suggest the presence of a third ring without heteroatoms or active hydrogen.

The foregoing data, taken together with UV and fluorescent properties of M349 and its relative hydrophobicity, lead to structure 1, 3-(β-D-ribofuranosyl)-4,9-dihydro-9-oxoimidazo[1,2a]purine, a Y nucleoside deriva-

tive previously unknown naturally or synthetically. The UV absorbance maximum for 1 in water is 239 nm, compared with 235 nm (100% methanol) for the Y base from yeast tRNA^{Phe} (16). Assignment of the glycosidic linkage as 3-β is based on: (a) the 3-β structure in the nucleoside of yeast tRNA^{Phe} Y base (16) which was shown by X-ray crystallography (17) to have the C-7 side chain "trans" to ribose, and (b) the conclusion that m¹G is a precursor to Y, as recently reviewed by Grosjean (18).

The base which corresponds to nucleoside was synthesized several years ago by Itaya (19), and was therefore used for a rigorous structural comparison against M217, the base released by acid hydrolysis from nucleoside M349 (Scheme 1). The comparison was based on: (a) HPLC retention times, (b) 254/280 nm UV absorbance ratio of HPLC peaks, (c) 254 nm absorbance/430 nm fluorescence emission ratio of HPLC peaks, (d) retention times of the volatile trimethylsilyl derivatives (Scheme 1) from capillary column gas chromatography, and (e) electron ionization mass spectra of the capillary column eluants.

From comparison of the HPLC retention times and UV and relative fluorescent properties of M217 and synthetic mimGua, shown in Fig. 2, they are judged to be the same. The 254/280 nm and 254/430 nm ratios of 2.3 vs. 2.2, and 10.8 vs. 11.0, are within experimental error. The comparison of capillary GC retention times and corresponding mass spectra are considered to be a particularly powerful and sensitive means of isomer differentiation. The difference in retention times (Figs. 3a, 4a) of less than 4 seconds at 28 minutes, and virtual identity of mass spectra (Figs. 3b, 4b) further conclusively establish that the base from *S. solfataricus* tRNA is 4,9-dihydro-4,6,7-trimethyl-9-oxoimidazo[1,2-a]purine, corresponding to nucleoside 1.

In addition to the present study, five other members of the Y family of nucleosides are presently known, summarized by Dirheimer (20); all are located at position 37 of tRNA^{Phe}. Functional aspects of the Y moiety have been recently reviewed by Bjork (21), and biosynthesis by Droogmans and Grosjean (18). The post-transcriptional conversion of G-37 to Y is heretofore known only in eukaryotes (22). The present finding of mimG in archaeobacteria adds further to the list of modified nucleosides in archaeobacterial tRNA which are otherwise usually found in eukaryotes (m⁵C, i⁶A, m¹I, m²G, m²G), and supports the observation that archaeobacterial modification patterns tend to be eukaryotic in nature (1).

ACKNOWLEDGEMENTS

This work was supported by N.I.H. grants GM 36042 to R.G., GM 21584 and GM29812 to J.A.M., N.A.S.A. grant NSG-7044 to C. R. Woese, and grants from the DFG and Fonds der Chemischen Industrie to K.O.S. The authors are indebted to Dr. T. Itaya, Kanazawa University, for an authentic sample of mimGua, and to the Center for Human Toxicology, University of Utah, for the use of a fluorescence detector.

REFERENCES

1. McCloskey, J. A. (1986) *System. Appl. Microbiol.* **7**, 246-252.
2. Blobstein, S.H., Grunberger, D., Weinstein, I. B. and Nakanishi, K. (1973) *Biochemistry* **12**, 188-193.
3. Zillig, W., Stetter, K. O., Wunderl, S., Schulz, W., Priess, H. and Scholz, J. (1980) *Arch. Microbiol.* **125**, 259-269.
4. Zillig, W., Stetter, K. O., Schafer, W., Jaecovic, D. Wunderl, S., Holz, I., and Palm, P. (1981) *Zbl. Bakt. Hyg., I, Abt. Orig. C.*, **2**, 205-227.
5. Stetter, K. O., (1982) *Nature* **300**, 258-260.
6. von Ehrenstein, G., (1967) *Meth. Enzymol.* **12A**, 588-596.
7. Edmonds, C. G., Vestal, M. L. and McCloskey, J. A. (1985) *Nucleic Acids Res.* **13**, 8197-8206.
8. Buck, M., Connick, M. and Ames, B. N. (1983) *Anal. Biochem.* **129**, 1-13.
9. Phillipson, D. W., Edmonds, C. G., Crain, P. F., Smith, D. L., Davis, D. R. and McCloskey, J. A. (1987) *J. Biol. Chem.* (in press).
10. Beardsley, K. and Cantor, C. R. (1970) *Proc. Natl. Acad. Sci USA* **65**, 39-46.
11. Crain, P. F. and McCloskey, J. A. (1983) *Anal. Biochem.* **132**, 124-131.
12. Pang, H. P., Schram, K. H., Smith, D. L., Gupta, S. P., Townsend, L. B. and McCloskey, J. A. (1982) *J. Org. Chem.* **47**, 3923-3932.
13. Crain, P. F., Yamamoto, H., McCloskey, J. A., Yamaizumi, Z., Nishimura, S., Limburg, K, Raba, M. and Gross, H. J. (1980) *Adv. Mass Spectrom.* **8**, 1135-1141.
14. White V, E., Krueger, P.M. and McCloskey, J.A. (1972) *J. Org. Chem.* **37**, 430-438.
15. von Minden, D. L., Stillwell, R.N., Koenig, W. A., Lyman, K. J. and McCloskey, J.A. (1972) *Anal. Biochem.* **50**, 110-121.
16. Nakanishi, K., Furutachi, N., Funamizu, M. (1970) *J. Am. Chem. Soc.* **92**, 7617-7619.
17. Quigley, G. J., Teeter, M. M. and Rich. A. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **75**, 64-68.
18. Droogmans, L. and Grosjean, H. (1987) *EMBO J.*, in press.
19. Itaya, T. and Mizutani, A. (1985) *Tetrahedron Lett.* **26**, 347-350.
20. Dirheimer, G. (1983) *Recent Results Cancer Res.* **84**, 1-46.
21. Bjork, G.R., Eriksson, J. U., Gustafsson, C.E.D., Hagerwall, T.G., Jonsson, Y.H. and Wikstrom, P. M. (1987) *Ann. Rev. Biochem.*, in press.
22. Sprinzl, M., Moll, J., Meissner, F. and Hartmann, T. (1985) *Nucleic Acids Res.* **11**, r1-r49.