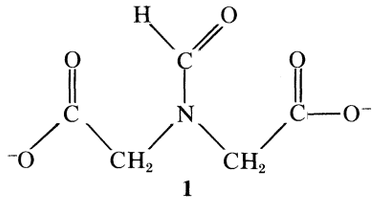
periment. This constancy, plus the resemblance of the proton NMR spectrum of the reaction products to that of dimethylformamide, suggested that the five unknown signals might be associated with a single species, possibly the *syn* and *anti* forms of the *N*-formyl derivative of iminodiacetic acid (IDA), a new compound (**1**) not previously reported in the literature:



1

In an effort to demonstrate whether this was the case, we prepared an authentic sample of *N*-formyliminodiacetic acid {FIDA or [*N*-(formyl)-*N*-(carboxymethyl)-glycine]} by an alternate route. We refluxed the disodium salt of IDA (Aldrich Chemical; 20g, containing 2 percent by weight of NTA as an impurity) with 100 ml of 90 percent HCOOH at 150°C for 10 minutes (> 95 percent conversion) (*6*). The solution was distilled under a vacuum overnight to remove the bulk of the HCOOH, and the residual solution was brought to *p*H 9 with 10 percent NaOH. The synthesized compound was positively identified as FIDA on the basis of elemental and spectral analysis (*7*). The ^{13}C NMR spectrum of the synthesized FIDA was coincident with that of the five unknown signals in the solution containing the products of the NTA-chlorine reaction.

The product yields from the NTA-chlorine reaction at *p*H 11.0 were estimated from the NMR peak heights to be as follows: FIDA, 1.46 mmole (40 percent yield); trioxymethylene, 0.36 mmole; HCO_3^-, 0.73 mmole; and $HCOO^-$, 0.13 mmole. The amount of unreacted NTA was 0.52 mmole or 15 percent of the starting material. We have also found that FIDA is produced at *p*H 5.0 from NTA and hypochlorous acid under otherwise identical conditions to those at *p*H 11.0, although apparently in somewhat lower yields (~ 10 percent). Formaldehyde (trioxymethylene) was also present, and so was IDA when the reaction solution contained phosphate buffer for *p*H stabilization (*8*).

To provide a preliminary assessment of the potential hazard of FIDA, we evaluated the synthesized material in the Ames assay using five strains of *Salmonella typhimurium* (TA98, TA100, TA1535, TA1537, and TA1538) with and without metabolic activation (*9*). We tested the FIDA twice in water at six dosages ranging from 10 to 5000 μg per plate. No mutagenic response was observed at any dosage, an indication that FIDA is not a potent mutagen.

This report is a further step toward addressing the concerns originally voiced by the Woods Committee over a need for information on NTA derivatives that might be formed from reactions in the environment and on their mutagenic potential. The lack of mutagenicity of FIDA in the Ames test is reassuring but, because this compound is new, other toxicological and environmental data should be sought as part of the continuing program to ensure the safety of NTA input into widespread use (*10*). Research on the reactivity of NTA with other oxidants in the environment is also warranted for the same reason.

RONALD J. SPANGGORD
CHARLES A. TYSON*
Life Sciences Division, SRI International, Menlo Park, California 94025

References and Notes

1. L. A. Woods, J. Cornfield, R. A. Goyer, A. E. Martell, T. S. Miya, W. E. Ribelin, W. R. Richter, T. H. Shepherd, W. E. Sutton, "Final report: Assessment of the potential of nitrilotriacetate (NTA) to compromise human health" (submitted by the Assistant Secretary for Health and Scientific Affairs' Ad Hoc Group on NTA, Department of Health, Education, and Welfare, Washington, D.C., 19 April 1972).
2. H. A. Mottola, *Toxicol. Environ. Chem. Rev.* **2**, 99 (1974).
3. W. W. Morganthaler and R. P. Langguth, *Chem. Spec. Manuf. Assoc. Proc. Mid-Year Meet.* **57**, 68 (1970).
4. C. B. Warren, "Chlorine oxidation of nitrilotriacetate (NTA) and related imino and amino acids," internal report (Monsanto Industrial Chemicals Company, St. Louis, Mo., December 1976).
5. Relative to a dioxane standard appearing at 67.4 ppm.
6. A. I. Vogel, *A Textbook of Practical Organic Chemistry* (Wiley, New York, ed. 3, 1956), p. 655.
7. The proton spectrum showed two methylene singlets at δ4.21 and 4.29 and a formyl singlet at δ8.32 corresponding to the monosodium salt of *N*-formyliminodiacetic acid in D_2O (δ4.89). The infrared spectrum showed the following absorbances (KBr pellet): 3400 (OH), 2950 (C–H), 2380 (N^+–H), 1660 (C=O), 1450, 1430, 1380, 1300, 1000, 875, 750, and 705 cm^{-1}. The ultraviolet spectrum (H_2O): maximum wavelength λ_{max} = 196 nm, molar absorptivity = 6200. Elemental data were consistent for a mixture containing 80 percent of the monosodium salt hydrate and 20 percent of the disodium salt hydrate. Calculated percentages: C, 29.26; H, 3.58; N, 6.82; Na, 13.27; values found: C, 29.67; H, 3.75; N, 6.74; Na, 13.49.
8. Not all signals were identified at *p*H 5.0, and minor amounts of other products may be formed at both experimental *p*H values studied (*p*H 5 and 11). A fragment, identified as diformylglycine [*N*,*N*-bis(formyl)-glycine], was observed in the gas chromatographic-mass spectrometric spectrum of the reaction solution, but neither this compound nor the FIDA was quantifiable because of the lability of the formyl group in strong acid during derivatization. When NTA, IDA, or glycine ($2.3 \times 10^{-3}M$) and excess NaOCl (0.02 to 0.07*M*) were reacted at 55°C and *p*H 10, the ratios of OCl^- consumed to the initial amine concentration were 10.5, 7.5, and 4.5, respectively, signifying complete decomposition to CO_2, N_2, H_2O, and Cl^-. The rate of bleach loss, monitored iodometrically, was at least five times as fast with IDA and glycine as with NTA.
9. B. N. Ames, F. D. Lee, W. E. Durston, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 782 (1973); B. N. Ames, W. E. Durston, Y. Yamasaki, F. D. Lee, *ibid.*, p. 2281.
10. Great Lakes Research Advisory Board, Annual Reports to the International Joint Commission, Windsor, Ontario, July 1977 and 1978. After reviewing all of the then available literature and Canadian experience, the board found (pp. 26 and 27 of the 1978 report) "no reasonable cause for restricting the use of NTA as a replacement for phosphate in laundry detergents in the Great Lakes Basin." The reports concluded that NTA was environmentally safe to use as a means of reducing phosphorus inputs into the basin in order to slow the rate of eutrophication in those waters, a process of serious concern to both Canadian and U.S. governments. In coming to this conclusion, the board accepted the recommendations of its task force on ecological effects to conduct certain additional studies and monitoring programs on NTA in the environment.

* To whom all correspondence should be addressed.

27 February 1979

DNA Organization of *Methanobacterium thermoautotrophicum*

Abstract. *The organism* Methanobacterium thermoautotrophicum, *an archaebacterium, is evolutionarily very distant from both traditional prokaryotes and eukaryotes. Its genome (DNA) has physical characteristics typical of most prokaryotes except that it is quite small (about 10^9 daltons, less than half the size of the genome of* Escherichia coli*) and contains a significant amount (6 percent) of DNA which renatures extremely rapidly.*

Studies by Woese and Fox (*1*) indicate that the traditional classification of life into two kingdoms, prokaryotes and eukaryotes, should be replaced by three categories or lines of descent, which Woese and Fox call archaebacteria, eubacteria, and urkaryotes. They propose these divisions because the comparison of two-dimensional nucleotide patterns (fingerprints) of 16*S* and 18*S* RNA's indicates that these three groups are evolutionarily equidistant from each other. The eubacteria and urkaryotes correspond roughly to the conventional categories of prokaryotes and eukaryotes, respectively. For example, the eubacteria comprise the typical bacteria (such as *Escherichia, Eubacterium, Micrococcus,* and *Spirochaeta*) as well as blue-green algae and chloroplasts. The urkaryotes comprise the typical eukaryotes such as animals, plants, fungi, and slime molds; or, more precisely, urkaryotes could be said to comprise the

0036-8075/79/0608-1082$00.50/0

cell components of the traditional eukaryotes that are not of endosymbiont origin. The eukaryotic cell is probably of composite lineage, some parts being prokaryotic (such as chloroplasts and mitochondria) and other parts urkaryotic (cytoplasm and nucleus).

In the third category, the archaebacteria, are found several groups of unusual organisms such as the methanogens, halophilic bacteria, and thermophilic bacteria. These apparently dissimilar groups prove, on the basis of their two-dimensional nucleotide patterns, to be evolutionarily rather closely related. Accordingly, though they are prokaryotes in the traditional sense of not having their genetic material localized in a separate nucleus, they are nevertheless placed by Woese and Fox in a kingdom of their own because of their vast evolutionary distance from all other prokaryotes (the eubacteria) and, equally, from all eukaryotes (the urkaryotes). This classification is supported by studies (*2*) of the cell walls, lipids, and modification patterns of bases in ribosomal and transfer RNA's (rRNA's and tRNA's).

Perhaps the most surprising element of Woese and Fox's new classification is not that the archaebacteria are so different from the eubacteria but that they are in fact *as close* to the urkaryotes as they are to the eubacteria. It is then of interest to examine whether the DNA organization of the methanogen *Methanobacterium thermoautotrophicum* resembles more closely that of the eubacteria or the urkaryotes with respect to complexity of the DNA, and whether there is a significant amount of repeated DNA.

To address these questions, we isolated DNA from frozen cells of *M. thermoautotrophicum* (strain ΔH), characterized the DNA with respect to various physical parameters (see below), and then performed renaturation kinetic experiments (below). The DNA was purified by two different methods: the Marmur procedure (*3*) and equilibrium banding on a potassium iodide (KI) gradient (*4*). The KI-isolated DNA was purer. Its hyperchromicity (increase of absorbance at 260 nm when thermally denatured to single strands) was 41.1 percent compared to 34.9 percent for the DNA isolated by the Marmur method.

The thermal denaturation profile of *M. thermoautotrophicum* DNA isolated by both methods is very sharp: it required 6.0°C to go from 10 to 90 percent denatured, while *Escherichia coli* required 6.3°C. This indicates homogeneity in base composition throughout the genome. The T_m (the temperature at which half the final hyperchromic shift is reached) is 89.4 ± 0.4°C relative to a T_m of 90.0°C for *E. coli* DNA. This corresponds to a G · C content of 49.4 percent (*5*) for the methanogen.

The buoyant density in cesium chloride at neutral *p*H was measured in the Spinco Model E ultracentrifuge. For the Marmur-isolated DNA, the banding pattern seen in ultraviolet scanner optics revealed a somewhat asymmetrical peak with a density of 1.704 ± 0.003 g/cm³ relative to a marker density of 1.724 for *Micrococcus lysodeikticus* DNA. To see whether the asymmetry was real or spurious, we looked at the purer, KI-isolated DNA. We found a density of 1.707 ± 0.003 g/cm³ relative to the same bacterial marker density; and now the peak was symmetrical, with no satellites being visible. The profile was very similar to that of the marker DNA, both peaks having a half-width at half maximum of 0.45 ± 0.005 mm. The content (mole percent) of G · C is 54.7 percent for the KI-isolated DNA and 51.0 percent for the Marmur-isolated DNA relative to *M. lysodeikticus* at 72 percent, based on the formula of Schildkraut *et al* (*6*).

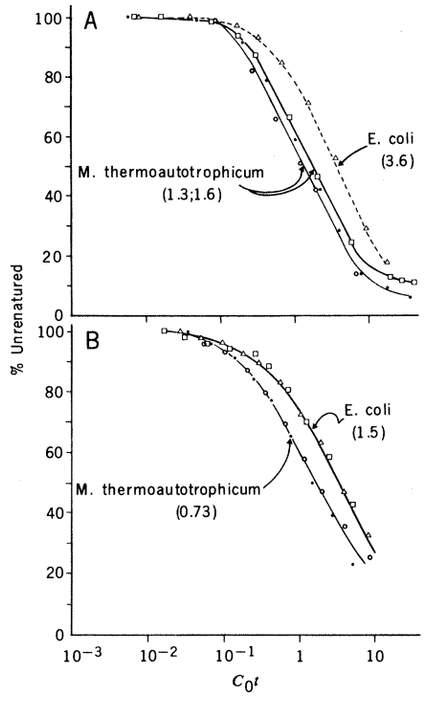


Fig. 1. Renaturation kinetic experiments on the DNA of *M. thermoautotrophicum* and *E. coli*. (A) Renaturation followed by HAP with ³H-labeled DNA (□) and the optical HAP method (○ and ●, for two different experiments). The fraction of DNA bound at zero time has been subtracted out. The average fragment length, L, was 366 base pairs for the ³H-labeled methanogen DNA, 280 and 292 base pairs for the methanogen DNA in the two optical HAP experiments, and 250 base pairs for *E. coli*. The data were adjusted to 400 base pairs in the figure with the use of the single-strand length dependence (*16*) of $L^{0.45}$. The C_0t is measured in (moles nucleotides) × (second)/liter. (B) Renaturation was followed by optical absorbance. The original data were normalized by dividing by the absorbance at zero time (actually, at about 1 minute), finding the percentage change in absorbance at 5 minutes on this basis, and correcting the value at 5 minutes to 100 percent in order to account for the single-strand structure formation that occurs during the first 5 minutes or so (*11, 14*). The C_0t is measured as (moles base pairs) × (second)/liter.

The purpose of the DNA renaturation kinetic experiments was to determine genome size (complexity), number of classes of DNA, and the amount of highly repeated DNA or foldback DNA. These may be determined most readily by "melting" the DNA to single strands, following the renaturation to double-stranded form, and plotting the unrenatured fraction as a function of time by the C_0t method (*7*) where C_0 is the initial concentration of nucleotides (moles per liter) and t is time (seconds). The genome size (complexity) is then proportional to $C_0t_{1/2}$, the value of C_0t when renaturation is half complete; the shape of the curve tells whether there is only one class of (nonrepeated) DNA or whether there is also one or more classes of intermediately repeated DNA; and the "zero-time" renaturation (fraction which forms double helices faster than can be measured) determines the amount of foldback or highly repeated DNA.

The DNA's used in the renaturation experiments were sheared to ~ 300 base pairs and sized by gel electrophoresis with restriction enzyme (Hae III) digest of plasmid pBR722 as a marker. The renaturation was studied by two different methods; binding to hydroxyapatite (*8*) (HAP) and ultraviolet absorption at 260 nm (Fig. 1). In the HAP experiments, the DNA was denatured at 100°C for 10 minutes and was allowed to renature at T_m −25°. The time course of the reaction was followed by measuring at intervals the portion (double-stranded) of DNA bound to the HAP at 60°C after which the single-stranded DNA was eluted at 100°C and independently measured. We used both tritium-labeled DNA, prepared by the method of nick translation (*9*), and nonradioactive DNA in concentrations large enough for samples and washings to be measured optically. For the labeled DNA, the curve (Fig. 1A) shows a $C_0t_{1/2}$ of about 1.6, while the curve for unlabeled DNA yields a $C_0t_{1/2}$ of about 1.3 (Fig. 1A). For

both these experiments, $C_0t_{1/2} = 3.6$ for *E. coli*, yielding a complexity for the methanogen 2.3 and 2.8 times smaller than that of *E. coli* (Fig. 1A).

The ultraviolet absorption (optical) method for studying renaturation kinetics consists of denaturing the DNA at 100°C and then following the formation of double-stranded DNA at $T_m - 25$°C by monitoring the loss of optical hyperchromicity at 260 nm. In this case, the time course of the reaction follows a somewhat different formula from the single second-order renaturation reaction, since ultraviolet absorption measures the fraction of *nucleotides* in double-stranded form while HAP measures the fraction of *strands* that have formed double-helical structures. The two are different because the shearing of the DNA (to 300 base pairs) occurs at random: a typical "+" strand will then have only a two-thirds overlap on the average with a complementary "−" strand (*10*), so that 33 percent of the nucleotides will be in single-stranded tails.

The result is that the time of 50 percent renaturation ($C_0t_{1/2}$) measured on HAP corresponds to one-third renaturation (33 percent) measured optically. We illustrate this briefly [for more detail see (*11*)]: let $f(t)$ be the fraction of strands remaining single at time t; this is what we determine with HAP. For a single class of DNA (for example, nonrepeated DNA) under nucleation rate-limiting conditions (*11*), $f(t)$ will follow standard second-order kinetics: $f(t) = 1/(1 + K_2S_0t)$ where K_2 is the nucleation rate constant and S_0 is the initial total concentration of single strands. Let $g(t)$, on the other hand, be the fraction of nucleotides in single-stranded form at time t; this is clearly greater than $f(t)$, since we must now count not only the nucleotides occurring in free single strands but also those in single-stranded tails of partial double helices. If we take the average double-stranded length to be two-thirds (so that one-third of the nucleotides are single-stranded) and note that the fraction of strands forming double-helical regions is $1 - f(t)$, we see that

$$g(t) = f(t) + 1/3\,[1 - f(t)]$$

or

$$g(t) = 2/3\,f(t) + 1/3$$

Thus, while $f(t)$ goes from unity (100 percent) at $t = 0$ to zero at $t = \infty$, $g(t)$ goes from 100 percent at $t = 0$ to 33 percent at $t = \infty$ (*12*). At 50 percent HAP renaturation $f(t) = 1/2$, while $g(t) = 2/3$, or 67 percent. Using this value, we see from Fig. 1B that the $C_0t_{1/2}$ (comparable to HAP) is 0.73 for the methanogen and 1.5 for *E. coli*, suggesting a complexity for the methanogen 2.1 times smaller than that of *E. coli*.

The three methods agree in that the complexity of the methanogen is less than that of *E. coli*. The most reliable, we feel, is that of the optical HAP. The nick-translation process, used in the isotopically labeled HAP, is well known for creating artifacts (*13*). (However, our labeled HAP and optical HAP data agree very well.) The optical method is uncertain because of (i) the complicating presence of random single-strand base-pairing during the first few minutes of the experiments (*14*), (ii) inaccuracy in measuring hyperchromicity, and (iii) uncertainty as to the average amount of single-stranded DNA in the tails of partly double-helical molecules (*11*). On the basis of the HAP data alone, we calculate a complexity for the methanogen about 2.5 ± 0.3 times smaller than that of *E. coli*. If we assume that *E. coli* is 2.7×10^9 daltons (*14*), then the complexity of the methanogen is $1.1 \pm 0.15 \times 10^9$ daltons.

The renaturation reaction of methanogen DNA parallels that of *E. coli* and takes place over two orders of magnitude of C_0t, indicating that only one class of DNA is present (Fig. 1, A and B). However, there was some zero-time binding—about 6 ± 2 percent of the methanogen's DNA in the HAP renaturation experiments while the *E. coli* has only about 0.5 percent. In Fig. 1, A and B, this zero-time-reacted DNA has been subtracted out. Possible sources of zero-time binding are: highly repeated DNA, so-called foldback DNA (*15*), and cross-linked DNA (from free radicals occurring during shearing, or from nick translation).

In some respects, the DNA of the methanogen *M. thermoautotrophicum* resembles the DNA of a typical bacterium such as *E. coli*: melting profiles show that the methanogen DNA is as homogeneous as *E. coli*'s, CsCl banding indicates that no satellites are present, and the renaturation studies indicate that the DNA is of one class. The differences from typical bacteria are: (i) lesser complexity and (ii) 6 percent zero-time-reacted DNA in the methanogen, the nature of which remains ambiguous at this time. It is impressive that, even with a small genome size, *M. thermoautotrophicum* is still a free-living organism capable of using H_2 alone as an energy source and CO_2 alone as a carbon source.

Woese and Fox's 16*S* and 18*S* RNA sequence data (*1*) place the urkaryotes, eubacteria, and archaebacteria at equal distances from each other evolutionarily. However, the question still exists as to the branching pattern of the evolutionary tree containing the three groups. If the rates of evolution of the RNA sequences of rRNA are equal for the three groups, the branches corresponding to each must diverge from a common point and must be of the same length. However, if we allow for unequal rates of evolution of the sequences of one or more groups, other branching patterns are possible—for example, the archaebacteria and a common ancestor of eubacteria and urkaryotes may first separate, and this ancestor may then further branch into the eubacteria and urkaryotes themselves. Given the presumed primitive nature of methanogens, either of these patterns seems attractive.

RALPH M. MITCHELL
Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138
LAUREL A. LOEBLICH
Marine Science Program, University of Houston, Galveston, Texas 77550
LYNN C. KLOTZ
Department of Biochemistry and Molecular Biology, Harvard University
ALFRED R. LOEBLICH III
Marine Science Program, University of Houston, Galveston

References and Notes

1. C. R. Woese and G. E. Fox, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5088 (1977).
2. C. R. Woese, L. J. Magrum, G. E. Fox, *J. Mol. Evol.* **11**, 245 (1978); O. Kandler and H. Konig, *Arch. Microbiol.* **118**, 141 (1978); T. G. Tornabene and T. A. Langworthy, *Science* **203**, 51 (1979); L. J. Magrum, K. R. Luehrsen, C. R. Woese, *J. Mol. Evol.* **11**, 1 (1978).
3. J. Marmur, *J. Mol. Biol.* **3**, 208 (1961).
4. N. Blin, A. V. Gabain, H. Bujard, *FEBS Lett.* **53**, 84 (1975).
5. J. Marmur and P. Doty, *J. Mol. Biol.* **5**, 109 (1962).
6. C. L. Schildkraut, J. Marmur, P. Doty, *J. Mol. Biol.* **4**, 430 (1962).
7. R. J. Britten and D. E. Kohne, *Science* **161**, 529 (1968).
8. R. J. Britten, D. E. Graham, B. R. Neufeld, *Methods Enzymol.* **29**, 363 (1974).
9. T. Maniatis, A. Jeffrey, D. G. Kleid, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1184 (1975).
10. J. G. Wetmur, thesis, California Institute of Technology (1967).
11. D. C. Rau and L. C. Klotz, *Biophys. Chem.* **8**, 41 (1978).
12. In practice, this form is not observed for optical renaturation beyond 40 percent or so because of reaction of the single-stranded tails to form multimers. Thus, the "infinity point" of 1/3 for $g(t)$ slowly decreases, typically to a value of ~ 0.1 (*11*).
13. T. Roberts, personal communication.
14. G. D. Lauer, T. M. Roberts, L. C. Klotz, *J. Mol. Biol.* **114**, 507 (1977).
15. D. A. Wilson and C. A. Thomas, *ibid.* **84**, 115 (1974).
16. A. G. Hinnebusch, V. E. Clark, L. C. Klotz, *Biochemistry* **17**, 1521 (1978).
17. We thank Professor R. S. Wolfe for a supply of methanogen frozen cells, C. Woese for useful discussion and his interest in our work as well as advance copies of his manuscripts, and A. Hinnebusch and N. Komar for useful discussions. Supported by grants from NIH and NSF to L.C.K. and A.R.L. III.

21 December 1978; revised 20 March 1979