Isolation and Structure Elucidation of Low Molecular Weight Components of Aurintricarboxylic Acid (ATA)

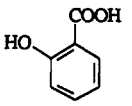
Pinglang Wang, John Kozlowski, and Mark Cushman*

Department of Medicinal Chemistry and Pharmacognosy, School of Pharmacy and Pharmacal Sciences, Purdue University, West Lafayette, Indiana 47907

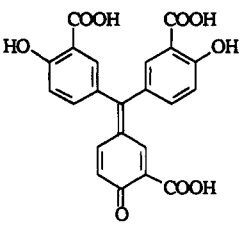
Received February 19, 1992

Nine low molecular weight components of aurintricarboxylic acid (ATA) have been isolated by silica gel column and thin-layer chromatography. The structures of these compounds have been elucidated by ^1H NMR and ^{13}C NMR spectroscopy in conjunction with COSY and HETCOR techniques for proton and carbon assignments. Four of the components afforded protection against the cytopathic effect of HIV-1 in CEM cell culture at relatively high concentrations that were near those resulting in cytotoxicity.

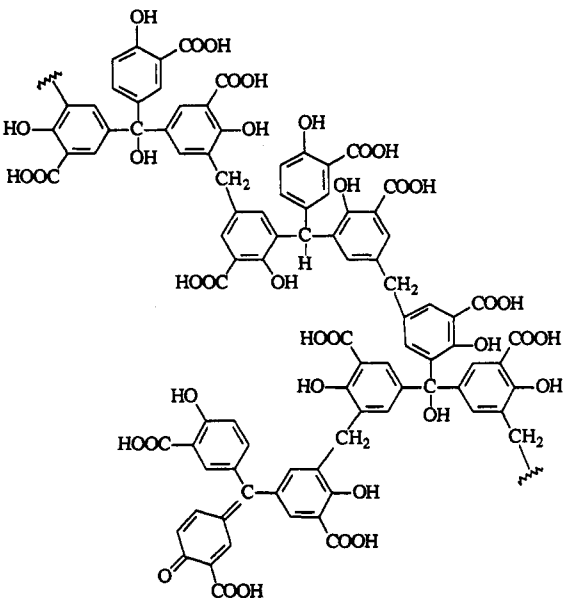
Aurintricarboxylic acid (ATA) is a solid substance that forms when salicylic acid (1) is treated with sulfuric acid, formaldehyde, and sodium nitrite.[1] ATA is a potent inhibitor of many biochemical processes that are dependent upon the binding of nucleic acids to proteins. For example, in cell-free systems it inhibits a variety of enzymes that process nucleic acids, including DNA polymerases,[2,3] RNA polymerases,[4-8] reverse transcriptases,[5,9,10] nucleases,[11-19] ribonucleotide reductases,[20] and aminoacyl-*t*RNA synthetase.[21] In addition, ATA inhibits protein synthesis by interfering with the binding of mRNA to ribosomes,[22,23] and it also inhibits the binding of glucocorticoid-receptor complexes and dihydroxyvitamin D_3-receptor complexes to DNA.[24,25] The evidence indicates that ATA inhibits these protein nucleic acid interactions by binding to the protein, not the nucleic acid.[4,12] Recent interest in ATA has resulted from the discovery that it prevents the cytopathic effect of the AIDS viruses, HIV-1 and HIV-2, in cell cultures.[26-28] This antiviral effect appears to be due to the binding of ATA to the CD4 receptor on the cell surface as well as to gp120 on the surface of the virus, which prevents the initial binding of the virus to the cell surface.[27,28]



1



2



3

(1) Caro, N. *Chem. Ber.* **1892**, *25*, 939-949.
(2) Seki, S.; Tsutsui, K.; Oda, T. *Biochem. Biophys. Res. Commun.* **1977**, *79*, 179.
(3) Nakane, H.; Balzarini, J.; De Clercq, E.; Ono, K. *Eur. J. Biochem.* **1988**, *177*, 91.
(4) Blumenthal, T.; Landers, T. A. *Biochem. Biophys. Res. Commun.* **1973**, *55*, 680.
(5) Liao, L.-L.; Horwitz, S. B.; Huang, M.-T.; Grollman, A. P.; Steward, D.; Martin, J. *J. Med. Chem.* **1975**, *18*, 117.
(6) Hughes, S.; Burns, J. H.; Birnie, G. D. *Methodol. Surv.* **1982**, *11*, 178.
(7) Iapalucci-Espinoza, S.; Haim, L.; Franze-Fernández, M. *Mol. Cell. Biochem.* **1983**, *55*, 41.
(8) Ananiev, E.; Karagyozov, L. *Physiol. Vég.* **1984**, *22*, 555.
(9) Givens, J. F.; Manly, K. F. *Nucleic Acids Res.* **1976**, *3*, 405.
(10) Baba, M.; Schols, D.; Pauwels, R.; Balzarini, J.; De Clercq, E. *Biochem. Biophys. Res. Commun.* **1988**, *155*, 1404.
(11) Hallick, R. B.; Chelm, B. K.; Gray, P. W.; Orozco, E. M. J. *Nucl. Acids Res.* **1977**, *4*, 3055.
(12) González, G.; Haxo, R. S.; Schleich, T. *Biochemistry* **1980**, *19*, 4299.
(13) Schulz-Harder, B.; Tata, J. R. *Biochem. Biophys. Res. Commun.* **1982**, *104*, 903.
(14) Kumar, A.; Ali, R. *J. Appl. Biochem.* **1984**, *6*, 76.
(15) Stern, D. B.; Newton, K. J. *Plant Mol. Biol. Rep.* **1984**, *2*, 8.
(16) Ramón, D.; Ferrer, S.; Vicente, E.; Uruburu, F. *FEBS Microbiol. Lett.* **1986**, *36*, 9.
(17) Kumar, A.; Ali, R. *Ind. J. Exptl. Biology* **1986**, *24*, 79.
(18) Eslami, B. S.; Konno, S.; Wu, J. M. *Biochem. Internatl.* **1986**, *13*, 163.
(19) Batistatou, A.; Greene, L. A. *J. Cell Biol.* **1991**, *115*, 461.
(20) Baumann, H.; Hofmann, R.; Lammers, M.; Schimpff-Weiland, G.; Follmann, H. Z. *Naturforsch.* **1983**, *39c*, 276.
(21) Igarashi, S. J.; Zmean, J. A. *Can. J. Biochem.* **1975**, *53*, 120.
(22) Grollman, A. P.; Stewart, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **1968**, *61*, 719.
(23) Igarashi, S. J.; Zmean, J. A. *Can. J. Biochem.* **1975**, *53*, 124.
(24) Moudgil, V. K.; Caradonna, V. M. *J. Steroid. Biochem.* **1985**, *23*, 125.
(25) Mellon, W. S. *Biochem. Pharmacol.* **1984**, *33*, 1047.
(26) Balzarini, J.; Mitsuya, H.; De Clercq, E.; Broder, S. *Biochem. Biophys. Res. Commun.* **1986**, *136*, 64.
(27) Schols, D.; Baba, M.; Pauwels, R.; Desmyter, J.; De Clercq, E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 3322.
(28) Cushman, M.; Wang, P.; Chang, S.; Wild, C.; De Clercq, E.; Schols, D.; Goldman, M. E.; Bowen, J. A. *J. Med. Chem.* **1991**, *34*, 329.

0022-3263/92/1957-3861$03.00/0

Table I. R_f Values of ATA Components with MW Smaller than 1000[a]

ATA component	R_f	mp (°C)
A-1	0.810	
A-2	0.606	239–241
A-3	0.538	
A-4	0.478	268–269
A-5	0.349	
A-6	0.302	
A-7	0.266	
A-8	0.240	270–272
A-9	0.186	184–186
A-10	0.180	280–282
A-11	0.152	
A-12	0.131	266–268
A-13	0.103	202–204
A-14	0.076	
A-15	0.065	197–199
A-16	0.038	
A-17	0.023	
A-18	0.016	
A-19	0.006	
A-20	0.000	

[a] This MW is based on the ultrafiltration membrane MW cutoff value. The TLC was performed on silica gel 60, 0.2 mm, eluting with $CHCl_3$, THF, and formic acid (100:5:2), developed one time at room temperature.

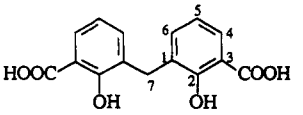
Aurintricarboxylic acid was originally thought to be the triphenylmethane dye 2, and this incorrect structure has appeared in much of the ATA literature. ATA is actually a heterogeneous mixture of polymers.[29] ATA has been fractionated, and the potencies of the fractions have been correlated with molecular weight in a number of assays, including prevention of the cytopathic effect of HIV-1 and HIV-2 in cell culture, syncytium formation, prevention of the binding of the OKT4A monoclonal antibody to the CD4 receptor, inhibition of binding of anti-gp120 monoclonal antibody binding to gp120, inhibition of attachment of HIV-1 virions to MT-4 cells, and inhibition of HIV-1 reverse transcriptase.[28]

Despite the biological significance of ATA, very little is actually known regarding the structures of its components. This information would be useful in attempting to understand the biological activity of ATA in chemical terms. A general structure 3 was previously proposed on the basis of a ^{13}C NMR study of three ATA fractions.[29] Prior fractionation schemes have not resulted in the isolation and characterization of any ATA components.[28-30] This is not surprising in view of the extremely complex mixture comprising ATA, as well as the similar structures and polarities of its components. The goal of the present study was to isolate and elucidate the structures of low molecular weight ATA components with the expectation that the results would also provide insight into the structures of the higher molecular weight components.

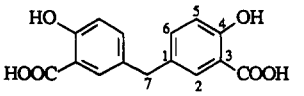
Results and Discussion

The ATA preparation used in this study was prepared according to a previously published procedure, which employs a limited reaction time at 0 °C.[28] The product was subjected to ultrafiltration through a membrane having a molecular weight cutoff of 1000. After considerable experimentation involving a variety of separation methods, it was determined that a relatively good resolution of the components of the ultrafiltrate could be achieved by thin-layer chromatography on silica gel, employing a mixture of chloroform, formic acid, and tetrahydrofuran (100:5:2) as the mobile phase. At least 20 components were detected and designated A-1–A-20 in order of decreasing R_f value (Table I). As indicated by TLC, the major components were A-1, A-4, A-8, A-13, and A-15. After repeated chromatography, nine samples were isolated: A-1, A-2, A-4, A-8, A-9, A-10, A-12, A-13, and A-15.

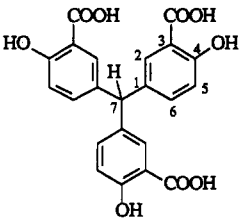
Compound A-1 proved to be salicylic acid (1). Compound A-2 is a methylenedisalicylic acid isomer 4. In the



4 (A-2)



5 (A-4)



6 (A-8)

^1H NMR spectrum, the H-4 protons appear as a doublet of doublets at δ 7.68 (J = 7.8 and 1.7 Hz), the H-5 protons correspond to an apparent triplet at δ 6.84 (J = 7.7 Hz), the H-6 protons produce a doublet of doublets at δ 7.27 (J = 7.5 and 1.7 Hz), and the H-7 methylene protons appear as a singlet at δ 3.91. On the basis of the 2D NMR heteronuclear (proton–carbon) correlation (HETCOR) spectrum, the four ^{13}C NMR signals corresponding to the four different types of hydrogen-bearing carbons can be assigned: δ 128.2 (C-4), 118.6 (C-5), 135.9 (C-9), and 28.5 (C-7). The ^{13}C NMR signals of C-1, C-2, and C-3 are assigned from the long-range HETCOR spectrum. There were three carbons with long-range correlations (multiple bond couplings) to the methylene protons attached to C-7 (δ 127.7, 135.9, and 159.5). The correlation to δ 135.9 was previously assigned to C-6. The two remaining signals, δ 127.7 and 159.5, arise from coupling to C-1 and C-2, respectively. Both H-4 and H-6 correlated with the signal at δ 159.4, which can be assigned to C-2. Two long-range correlations with H-5 were observed (δ 112.2 and 127.7). Previous assignment of δ 127.7 to C-1 left δ 112.2 as the C-3 resonance. An authentic sample of A-2 previously prepared by us had identical spectral properties.[31]

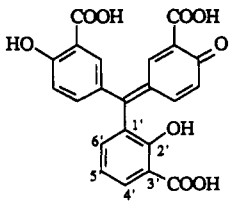
Compound A-4 is the methylenedisalicylic acid isomer 5. The protons can be assigned to signals in the ^1H NMR spectrum on the basis of their chemical shifts and coupling constants. Protons H-2 appeared as a doublet at δ 7.61 (J = 2.3 Hz), the H-5 protons were a doublet at δ 6.87 (J = 8.5 Hz), the H-6 protons corresponded to a doublet of doublets at δ 7.34 (J = 8.5 and 2.3 Hz), and the H-7 protons appeared as a singlet at δ 3.83. The four methine carbons were assigned on the basis of HETCOR as follows: δ 136.2 (C-6), 129.8 (C-2), 117.3 (C-5), and 38.7 (C-7). The assignment of the remaining signals of A-4 was completed by analysis of the long-range HETCOR spectrum. It was observed that H-7 correlated with C-2 at δ 129.8 ($^3J_{CH}$) and C-6 at 136.2 ($^3J_{CH}$); H-7 correlated with C-1 at δ 132.1

(29) González, R.; Blackburn, B. J.; Schleich, T. *Biochim. Biophys. Acta* 1979, *562*, 534.
(30) Tsutsui, K.; Seki, S.; Tsutsui, K.; Oda, T. *Biochim. Biophys. Acta* 1978, *517*, 14.
(31) Cushman, M.; Kanamathareddy, S.; De Clercq, E.; Schols, D.; Goldman, M.; Bowen, J. A. *J. Med. Chem.* 1991, *34*, 337.

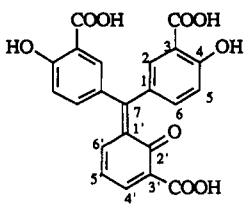
($^2J_{CH}$); H-2 correlated with C-4 at δ 159.6, the carboxyl carbon at δ 171.9 ($^3J_{CH}$), and C-6 at δ 136.2 ($^3J_{CH}$); H-5 correlated with to C-1 at δ 132.1 ($^3J_{CH}$) and C-3 at 112.7 ($^3J_{CH}$); and H-6 correlated with C-4 (δ 159.6) ($^3J_{CH}$) and C-2 (δ 129.82) ($^3J_{CH}$). The properties of A-4 also agreed with those of an authentic sample previously synthesized in our laboratory.[32]

Substance A-8 was assigned the triphenylmethane structure **6**. The four different types of protons attached to carbons were assigned in the ^1H NMR spectrum as follows: δ 7.49 (d, J = 2.4 Hz, H-2), 6.92 (d, J = 8.6 Hz, H-5), 7.23 (dd, J = 8.6 and 2.4 Hz, H-6), and 5.59 (s, H-7). The protonated aryl carbons were assigned on the basis of the HETCOR data: δ 136.1 (C-6), 130.0 (C-2), 117.4 (C-5), and 52.32 (C-7). The ^{13}C NMR assignments were completed through analysis of the long-range HETCOR data set. Proton H-7 displayed correlations to C-2 (δ 130.0) and C-6 (δ 136.1), two-bond correlation with C-1 (δ 134.3); H-2 had three-bond correlations with C-7 (δ 52.3), C-4 (δ 159.6), and C-6 (δ 136.1) and the carboxyl carbon (δ 171.6); H-5 correlated through three bonds with C-1 (δ 134.3) and C-3 (δ 112.7); and H-6 had three-bond correlations with C-2 (δ 130.0) and C-4 (δ 159.6). The properties of A-8 were in accord with those of authentic **6** prepared by us before.[31]

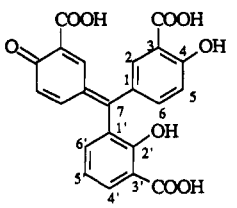
Component A-9 is evidently a mixture of three tautomers **7a–c**. Based on the ^1H NMR spectrum, the six different types of protons attached to carbon can be assigned as follows: δ 7.77 (dd, J = 7.8 and 1.7 Hz, H-4′), 7.67 (d,



7a (A-9)



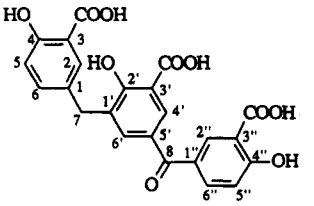
7b (A-9)



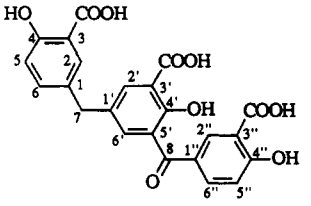
7c (A-9)

J = 2.6 Hz, H-2), 7.37 (dd, J = 7.8 and 1.5 Hz, H-6′), 7.33 (dd, J = 8.8 and 2.6 Hz, H-6), 6.89 (apparent t, J = 7.7 Hz, H-5′), and 6.88 (d, J = 8.7 Hz, H-5). From HETCOR analysis, the ^{13}C NMR signal at δ 135.1 was assigned to C-6, the signal at 133.9 was assigned to C-6′, the signal at δ 129.8 was assigned to C-4′, the signal at δ 129.0 was assigned to C-2, the signal at δ 118.2 was assigned to C-5′, and the signal at δ 116.2 was assigned to C-5. Since separate signals for the two nonequivalent tautomers **7ac** and **7b** were not observed, the three tautomers **7a–c** are evidently in rapid equilibrium on the NMR time scale.

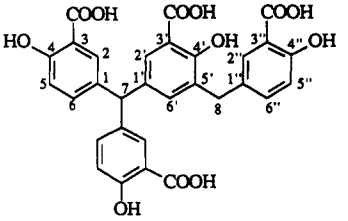
Component A-10 has been assigned structure **8**. From the ^1H–^1H COSY NMR spectrum the proton signals were coupled in three groups. Couplings between δ 8.16–7.88–7.10, 8.05–7.85, and δ 7.72–7.44–6.89 were observed. The magnitudes of the coupling constants were measured from the ^1H survey spectrum. All of the protons in the NMR spectrum of **8** could be assigned: δ 8.16 (d, J = 2.2 Hz, H-2″), 8.05 (d, J = 2.2 Hz, H-4′), 7.88 (dd, J = 8.7 and 2.2 Hz, H-6″), 7.85 (d, J = 1.9 Hz, H-6′), 7.72 (d, J = 2.1 Hz, H-2), 7.44 (dd, J = 8.6 and 2.21 Hz, H-6), 7.10 (d, J = 8.7 Hz, H-5″), 6.89 (d, J = 8.5 Hz, H-5), and 3.96 (s, 7 H). The nine carbons directly bonded to these protons were assigned from the HETCOR NMR spectrum: δ 136.5 (C-6″), 136.4 (C-6′), 136.2 (C-6), 132.7 (C-2″), 130.8 (C-2), 130.0 (C-4′), 117.5 (C-5″), 117.2 (C-5), and 33.6 (C-7). From the long-range HETCOR analysis, H-2″ couples through three bonds with C-4″ and the carbon of the carboxy group at δ 164.2 and 171.1, respectively. Proton H-6′ correlated with C-2′ at 162.6 ($^3J_{CH}$). Proton H-4′ correlated with C-2′ and carboxyl carbon at 162.6 and 171.9 ($^3J_{CH}$). Proton H-2 correlated with C-4 and carboxyl carbon at 159.5 and 171.8 ($^3J_{CH}$). It was observed that H-7 couples through three bonds to the phenolic hydroxyl-bearing carbon at C-2′. This H-7 to C-2′ correlation excludes the isomeric structure **9** for A-10.



8 (A-10)



9



10 (A-12)

The spectra of A-12 are consistent with structure **10**. The ^1H NMR proton assignments of 10 different types of hydrogens in **10** were made with the assistance of H^1–H^1 COSY NMR data: δ 7.60 (d, J = 2.1 Hz, H-2″), 7.48 (m, H-2), 7.36 (d, J = 2.1 Hz, H-2′), 7.27 (dd, J = 8.5 and 2.1 Hz, H-6″), 7.22 (m, H-6′), 7.20 (m, H-6), 6.90 (d, J = 8.5 Hz, H-5), 6.83 (d, J = 8.5 Hz, H-5″), 5.56 (s, H-7), 3.82 (s, H-8). The HETCOR data revealed the ^{13}C NMR chemical shift values of the hydrogen-bearing carbon atoms: δ 136.8 (C-6′), 136.1 (C-6), 135.9 (C-6″), 130.1 (C-2), 129.9 (C-2″), 128.17 (C-2′), 117.42 (C-5), 117.05 (C-5″), 52.42 (C-7), and 33.86 (C-8). From the long-range HETCOR spectrum, proton H-2 (δ 7.48) has three-bond correlations with a carboxylic acid carbon (δ 171.7), C-4 (δ 159.7), and C-6 (δ 136.1); H-5 (δ 6.90) has three-bond correlations to C-1 (δ 134.4) and C-3 (δ 112.7) and a one-bond correlation with C-5 (δ 117.42); H-6 (δ 7.20) displays three-bond correlations

(32) Cushman, M.; Kanamathareddy, S. *Tetrahedron* **1990**, *46*, 1491.

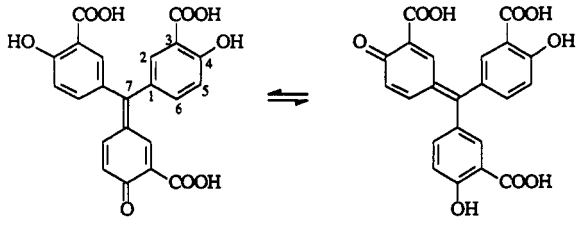
Table II. The Molar Absorptivity of A-13 at 556 nm as a Function of Time[a]

time (min)	molar absorptivity in acetone: water added	molar absorptivity in acetone: no water
1.5	7837	8182
3.5	4157	8182
5.5	2530	8182
8.0	1506	8182
11.0	964	8182
14.0	663	8182
17.0	422	8182
21.0	331	8182
26.0	331	8182
32.0	331	8182

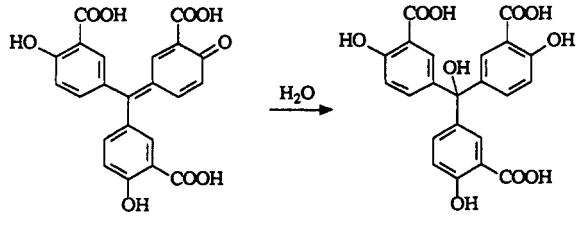
[a] A sample of A-13 (0.9 mg) was dissolved in dry ($CaSO_4$) acetone (9 mL), and water (1 mL) was added. Molar absorptivity at 556 nm was measured at the indicated times after addition of water.

with C-2 (δ 130.1) and C-4 (δ 159.7) and a one-bond correlation with C-6 (δ 136.1); H-2′ (δ 7.36) is coupled through three bonds with a carboxylic acid carbon at δ 172.3, C-4′ (δ 158.0), and C-6′ (δ 136.8); H-7 (δ 5.56) displays two-bond correlation with C-1′ (δ 133.9) and with C-1 (δ 134.4) and three-bond correlations with C-6 (δ 136.1), C-2 (δ 130.1), and C-6′ (δ 136.8); H-2″ (δ 7.60) has three-bond correlations with C-4″ (δ 159.5), C-6″ (135.9), and a carboxylic acid carbon (δ 171.9); H-5″ has three-bond correlation with C-1″ (δ 130.9) and C-3″ (δ 112.5); H-6″ (δ 7.27) has a three-bond correlation with C-4″ (δ 159.5); and H-8 (δ 3.82) shows a three-bond correlation with C-2″ (δ 129.9), a two-bond correlation with C-1″ (δ 130.9), and three-bond couplings with C-6″ (δ 135.91), C-6′ (δ 136.8), and C-4′ (δ 158.0). One additional correlation with H-8 (δ 3.82) must come from C-5′ (δ 129.4). The C-3 signal can be assigned to δ 112.2 by the process of elimination.

Component A-13 is a mixture of three rapidly equilibrating tautomers **2a**, **2b**, and **2c**. In DMSO, the rapid equilibration of these tautomers is indicated by the fact that only three types of aromatic hydrogens and eight types of carbons are evident in the ^1H and ^{13}C NMR spectra. Compound A-13 is of special interest because



2a (A-13) **2b** (A-13)



2c (A-13) **11**

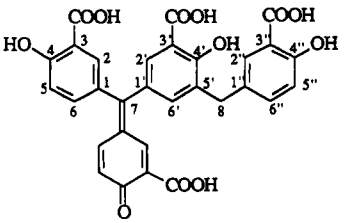
ATA was long believed to be a single compound having structure **2**. The three types of hydrogens on aromatic carbons in **2a–c** can be easily assigned on the basis of the chemical shifts and coupling constants: δ 7.63 (d, J = 2.6 Hz, H-2), 7.29 (dd, J = 8.7 and 2.6 Hz, H-6), and 6.92 (d, J = 8.8 Hz, H-5). From HETCOR analysis, the three carbons attached to the three aromatic hydrogens can be assigned: δ 135.1 (C-6), 128.81 (C-2), and 116.6 (C-5). On the basis of the long-range HETCOR analysis, the remaining carbon atoms can be assigned. H-5 correlated through three bonds with C-1 and C-3 at δ 138.2 and 111.9. H-2 correlated with C-4 at 159.9 ($^3J_{CH}$). We had observed that A-8 gradually oxidized to A-13 after standing in air at room temperature for several weeks, which was another confirmation for the validity of the structure of A-13. We also found that when A-13 is dissolved in acetone containing water, it is converted to the covalent hydrate **11**. The conversion of **2** to the covalent hydrate **11** is accompanied by a change in the color of the solution from red to almost colorless. This change in color could be quantitated by observing the decrease in molar absorptivity of the solution at 553–560 nm (Table II). No change in molar absorptivity occurred in the absence of water. The conversion of the quinone methide **2** to the covalent hydrate **11** could also be followed by ^1H NMR as detailed in Table III. The properties of the stable covalent hydrate **11** are in agreement with an authentic sample previously prepared by us during an attempted synthesis of **2**.[32]

Table III. ^1H NMR Data of A-13 Quinone (2) and Hydrated Form (11) in Acetone-d_6 at Different Times

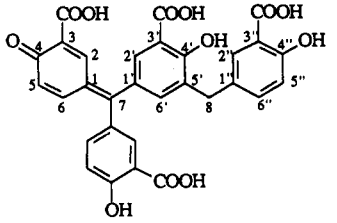
structure	chemical shift (ppm)	integration values[a] at different times[b] (h): 0.2	0.4	0.6	0.8	3.0	9.0	24.0
quinone (2)	7.02	23.0	21.0	20.5	19.0	12.0	7.0	3.8
	7.61	23.0	21.0	20.5	19.0	12.0	7.0	3.8
	8.08	23.0	21.0	20.5	19.0	12.0	7.0	3.8
hydrated (11)	6.94	12.0	14.0	14.5	16.0	23.0	28.0	31.2
	7.47	12.0	14.0	14.5	16.0	23.0	28.0	31.2
	7.89	12.0	14.0	14.5	16.0	23.0	28.0	31.2

[a] The integration values (change in integrator trace height in mm) were measured directly from the spectra. [b] Time after dissolution of the sample.

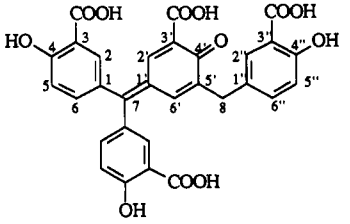
Component A-15 is compound **12a–c**. As with **2**, the molar absorptivity of A-15 at 527 nm decreases over time in aqueous acetone solutions, supporting the quinone methide structure of A-15. We also observed that A-12 was gradually oxidized to A-15 after standing in air at room temperature for several weeks. In addition, the observation of only eight types of aromatic protons in the ^1H NMR spectrum of A-15 indicates that it is a mixture of three rapidly equilibrating tautomers **12a–c**. On the basis of the ^1H NMR spectrum, six of the eight aromatic protons can be assigned as follows: δ 7.64 (d, J = 2.4 Hz, H-2), 7.27 (d, overlap, H-6′), 7.26 (dd, overlap, H-6), 7.24 (dd, overlap, H-6″), 6.89 (d, J = 8.7 Hz, H-5), and 6.82 (d, J = 8.4 Hz, H-5″). Two protons left at δ 7.60 (d, J = 2.3 Hz), 7.50 (d, J = 2.4 Hz) were from H-2′ and H-2″ and cannot at present be distinguished. The HETCOR spectrum led to the assignment of six of the eight aromatic carbons in the ^{13}C NMR spectrum: δ 135.80 (C-6″), 135.75 (C-6′), 135.03 (C-6), 128.83 (C-2), 116.92 (C-5″), 116.50 (C-5). The HETCOR spectrum also showed that the ^1H signals at δ 7.60 and 7.50 correlated with carbons at δ 129.82 and 129.96, respectively. From the long-range HETCOR spectrum it was found that the ^1H signal δ 7.60 correlated with C-6″ (δ 135.80, $^3J_{CH}$). Consequently, the proton at δ 7.60 must be H-2″ and the other ^1H signal δ 7.50 must be H-2′, since H-2′ was found to be correlated with C-6′ (δ 135.75, $^3J_{CH}$). Proton H-5″ correlated with C-3″ and C-1″ at δ 112.43 and 130.74 ($^3J_{CH}$). Proton H-5 correlated with C-1 (δ 138.18, $^3J_{CH}$) and C-3 (δ 111.91, $^3J_{CH}$). Proton H-2 correlated with C-4 (δ 159.93, $^3J_{CH}$) and a carboxyl carbon (δ 171.74, $^3J_{CH}$). Proton H-2″ correlated with C-4″



12a (A-15)



12b (A-15)



12c (A-15)

(δ 159.38, $^3J_{CH}$) and a carboxyl carbon (δ 171.78, $^3J_{CH}$).

The anti-HIV-1 activities of the various components were investigated and are listed in Table IV. The values for compounds **4** (A-2), **5** (A-4), **6** (A-8), and **11** were previously reported on samples synthesized by us.[31] The remaining values were determined on ATA components isolated in the present study. Compounds **4** (A-2), **5** (A-4), **7** (A-9), and **12** (A-15) proved to be inactive. The remaining substances inhibited the cytopathic effect of HIV-1 in CEM lymphocyte cell culture at relatively high concentrations that were near the cytotoxic concentrations. All of these low molecular weight components were considerably less potent in preventing the cytopathic effect of HIV-1 than unfractionated ATA.[28] This is not surprising because prior investigations of the anti-HIV activities of ATA fractions have shown that the antiviral potency increases with the molecular weight of the ATA fraction.[28] However, the low molecular weight components having anti-HIV-1 activity described here offer the advantage that, in contrast to ATA, they are structurally defined, pure compounds. These materials offer structural leads for the development of new potential anti-AIDS agents.

A previously published fractionation of ATA also afforded a low molecular weight fraction with a number average molecular weight of 396.[29] The ^{13}C NMR analysis of that fraction led to the conclusion that "monomeric ATA" (**2**) does not exist, at least in that preparation.[29] The present study demonstrates that monomeric ATA and related low molecular weight compounds can in fact be isolated from ATA. It has been recognized before that the composition of ATA varies depending of the exact reaction conditions.[29] The structures of these compounds will prove to be of value in defining the structures of higher molecular weight components of ATA and in understanding the molecular basis for the biological activities of ATA. It is possible that the polyanionic nature of ATA at physiological pH may resemble the phosphodiester backbone of nucleic acids, and therefore ATA may bind to enzymes that process nucleic acids. In this sense, ATA may be considered to be a relatively metabolically stable oligonucleotide mimetic.

Table IV. Anti-HIV-1 Activities of Low Molecular Weight ATA Components

compd	EC_{50}[a] (M)	IC_{50}[b] (M)
2	NA[c]	NA[c]
4 (A-2)	NA[c]	NA[c]
5 (A-4)	NA[c]	NA[c]
6 (A-8)	7.7×10^{-5}	6.0×10^{-4}
7 (A-9)	NA[c]	NA[c]
8 (A-10)	3.8×10^{-4}	1.0×10^{-4}
10 (A-12)	9.2×10^{-5}	2.0×10^{-4}
11	1.1×10^{-4}	5.6×10^{-4}
12 (A-15)	NA[c]	2.0×10^{-4}

[a] The EC_{50} is the 50% inhibitory concentration for cytopathicity of HIV-1 in CEM cells. The values are averages of at least four determinations. All compounds were tested as their ammonium salts. [b] The IC_{50} is the 50% cytotoxic concentration for mock-infected CEM cells. The values are averages of at least four determinations. All compounds were tested as their ammonium salts. [c] No activity was observed at a concentration of at least 3.5×10^{-4} M.

Experimental Section

Melting points are uncorrected. ^1H NMR spectra were recorded at 500 MHz, and ^{13}C NMR spectra were obtained at 125.7 MHz. The ^1H COSY, HETCOR, and long-range HETCOR spectra were obtained using standard Varian pulse sequences. The NMR samples were prepared in DMSO-d_6 and were referenced to TMS as an internal standard. Low-resolution chemical ionization mass spectra (CIMS) were determined using isobutane as the reagent gas. Microanalyses were performed by the Purdue Microanalytical Laboratory.

Ultrafiltration of ATA through a MW 1000 Cutoff Membrane. The ammonium salt of ATA was prepared as described previously.[28] An Amicon 8200 ultrafiltration apparatus with a 200-mL capacity cell and a YM2 ultrafiltration membrane (Amicon, Danver, MA) with molecular weight cutoff 1000 Da were used for fractionation of ATA to obtain material of MW less than 1000. The ammonium salt of ATA (2.0 g) was dissolved in distilled water (150 mL), the solution was placed in the cell, and 40 psi of pressure was applied. A total of 500 mL of ultrafiltrate was collected from five operations. The ultrafiltrate was evaporated to a volume of 200 mL, and the solution was acidified with 1.2 N HCl to precipitate the ATA components. The precipitate was collected by filtration and washed four times with distilled water. The wet sample was frozen and then dried under reduced pressure (1.5 mm) for 24 h at room temperature. A red powder (1.3 g) was obtained and subjected to column chromatography and TLC for pure oligomer isolation.

Isolation of Pure ATA Oligomers. The fractionated ATA sample with MW < 1000 was dissolved in acetone (50 mL) and added to silica gel (8 g, Aldrich 60, 230–400 mesh). The solvent was then evaporated on a rotary evaporator. The mixture was placed on top of a wet column of the same type of silica gel (300 g) as above, and the column was eluted with a solvent mixture containing $CHCl_3$, HCOOH, and THF (100:4:0.3). Column chromatography was monitored by TLC observing fluorescent spots under UV light while the TLC plates were still wet with solvent. The TLC plates were developed with the same solvent mixture employed for column chromatography. When A-4 had eluted, the solvent mixture was changed to $CHCl_3$, HCOOH, and THF (100:6:2), and chromatography was continued until A-15 was obtained. Fractions containing a single component as indicated by TLC were then combined, washed with water, and evaporated

to dryness, leaving residues that were purified further by preparative TLC. During this process some pure compounds were obtained as indicated by analytical TLC. However, most of the isolates still contained mixtures of two or three compounds. These mixtures were purified by repeated preparative TLC on silica gel until pure compounds were obtained.

3,3'-Dicarboxy-2,2'-dihydroxydiphenylmethane (4, component A-2): mp 239–241 °C (dec, lit.[31] mp 284–286 °C); UV (ethanol) 273 (ε 19 287); IR (KBr) 3460, 3225, 3178, 3049, 2919, 2578, 2320, 1660, 1449, 1243, 1185, 1091, 873, 761 cm^{-1}; ^1H NMR (500 MHz, DMSO-d_6) δ 13–14 (broad peak, exchangeable with D_2O, 2 H), 11.6–11.9 (broad peak, exchangeable with D_2O, 2 H), 7.68 (dd, J = 7.8 and 1.7 Hz, 2 H), 7.27 (dd, J = 7.5 and 1.7 Hz, 2 H), 6.84 (t, J = 7.7 Hz, 2 H), 3.92 (s, 2 H); ^{13}C NMR (125.7 MHz, DMSO-d_6) δ 172.54, 159.45, 135.94, 128.27, 127.68, 118.61, 112.26, 28.42 ppm; CIMS m/z (relative intensity) 289.1 (MH^+, 100), 271.1 (92). Anal. Calcd for $C_{15}H_{12}O_6$: C, 62.50; H, 4.16. Found: C, 62.34; H, 4.05.

3,3'-Dicarboxy-4,4'-dihydroxydiphenylmethane (5, component A-4): mp 268–269 °C (dec, lit.[32] mp 266–269 °C); UV (ethanol) 312 (ε 6400); IR (KBr) 3217, 3028, 2911, 2863, 2615, 1671, 1618, 1583, 1489, 1441, 1282, 1212, 1188, 1129, 1088, 899, 834, 799, 740, 681 cm^{-1}; ^1H NMR (500 Hz, DMSO-d_6) δ 10.8–11.6 (broad peak, exchangeable with D_2O, 2 H), 7.61 (d, J = 2.3 Hz, 2 H), 7.34 (dd, J = 8.5 and 2.3 Hz, 2 H), 6.87 (d, J = 8.5 Hz, 2 H), 3.83 (s, 2 H); ^{13}C NMR (125.7 MHz, DMSO-d_6) δ 171.90, 159.58, 136.16, 132.08, 129.82, 117.31, 112.72, 38.72; CIMS m/z (relative intensity) 289.0 (MH^+, 100), 271.0 (78), 151.0 (16). Anal. Calcd for $C_{15}H_{12}O_6$: C, 62.50; H, 4.16. Found: C, 62.29; H, 4.17.

3,3',3''-Tricarboxy-4,4',4''-trihydroxytriphenylmethane (6, component A-8): mp 270–272 °C (dec, lit.[31] mp 280–282 °C); UV (ethanol) 314 (ε 13 327); IR (KBr) 3457, 3261–2859, 2524, 2328, 1668, 1619, 1584, 1486, 1295, 1231, 1206, 1083, 897, 843, 793 cm^{-1}; ^1H NMR (500 MHz, DMSO-d_6) δ 10.8–11.4 (broad peak, exchangeable with D_2O, 3 H), 7.49 (d, J = 2.4 Hz, 3 H), 7.23 (dd, J = 8.6 and 2.4 Hz, 3 H), 6.92 (d, J = 8.6 Hz, 3 H), 5.59 (s, 1 H); ^{13}C NMR (125.7 MHz, DMSO-d_6) δ 171.62, 159.63, 136.11, 134.33, 130.00, 117.40, 112.67, 52.32; FABMS m/z (relative intensity) 425 (MH^+, 7), 407 (21), 389 (7), 371 (6). Anal. Calcd for $C_{22}H_{16}O_9$: C, 62.26; H, 3.77. Found: C, 62.29; H, 3.62.

Component A-9 (7): mp 184–186 °C dec; UV (ethanol) 307 (ε 11 816), 535 (ε 608); IR (KBr) 3600–2373, 1685, 1675, 1658, 1610, 1490, 1438, 1363, 1285, 1205, 1084, 840, 800, 765, 671 cm^{-1}; ^1H NMR (500 MHz, DMSO-d_6) δ 12.2–11.9 (broad peak, exchangeable with D_2O, 1 H), 11.1–11.3 (broad peak, exchangeable with D_2O, 2 H), 7.77 (dd, J = 7.7 and 1.7 Hz, 1 H), 7.67 (d, J = 2.5 Hz, 2 H), 7.37 (dd, J = 7.7 and 1.5 Hz, 1 H), 7.33 (dd, J = 8.8 and 2.5 Hz, 2 H), 6.89 (apparent t, J = 7.7 Hz, 1 H), 6.88 (d, J = 8.7 Hz, 2 H); ^{13}C NMR (125.7 MHz, DMSO-d_6) δ 172.52, 171.83, 159.88, 159.03, 136.81, 135.06, 133.97, 133.91, 129.76, 129.02, 118.16, 116.22, 113.23, 111.75, 78.63; FABMS (negative ion mode) m/z (relative intensity) 421.5 [$(M - H)^-$, 95.5]. Anal. Calcd for $C_{22}H_{14}O_9 \cdot H_2O$: C, 60.00; H, 3.66. Found: C, 60.12; H, 3.82.

Component A-10 (8): mp 280–282 °C dec; UV (ethanol) 306 nm (ε 17 764); IR (KBr) 3700–2370, 1666, 1617, 1592, 1488, 1446, 1355, 1285, 1212, 1083, 908, 838, 797, 761, 702, 627, 598, 518 cm^{-1}; ^1H NMR (500 MHz, DMSO-d_6) δ 8.16 (d, J = 2.2 Hz, 1 H), 8.05 (d, J = 2.2 Hz, 1 H), 7.88 (dd, J = 8.7 and 2.2, 1 H), 7.85 (d, J = 1.9 Hz, 1 H), 7.72 (d, J = 2.1 Hz, 1 H), 7.44 (dd, J = 8.6 and 2.1 Hz, 1 H), 7.10 (d, J = 8.7 Hz, 1 H), 6.89 (d, J = 8.5 Hz, 1 H), 3.96 (s, 2 H); ^{13}C NMR (125.7 MHz, DMSO-d_6) δ 191.70, 171.87, 171.76, 171.12, 164.19, 162.61, 159.53, 136.48, 136.39, 136.16, 132.74, 130.80, 130.31, 130.04, 129.88, 128.17, 127.60, 117.47, 117.17, 113.01, 112.54, 112.29, 33.62; FABMS m/z (relative intensity) 453.0, (MH^+, 5.8), 435.0 ($MH^+ - H_2O$, 2.1), 246.0 (13.0), 219.3 (15.6), 185.3 (100).

Component A-12 (10): mp 266–268 °C, UV (ethanol) 315 (ε 20 377); IR (KBr) 3010–2370, 1676, 1618, 1491, 1438, 1265, 1197, 901, 840, 797, 670 cm^{-1}; ^1H NMR (500 MHz, DMSO-d_6) δ 11.7–11.5 (broad peak, exchangeable with D_2O, 1 H), 11.3–11.0 (broad peak, exchangeable with D_2O, 2 H), 7.60 (d, J = 2.1 Hz, 1 H), 7.48 (d, J = 2.2 Hz, 2 H), 7.36 (d, J = 2.1 Hz, 1 H), 7.27 (dd, J = 8.5 and 2.2 Hz, 1 H), 7.22 (d, J = 2.1 Hz, 1 H), 7.20 (dd, J = 8.5 and 2.2 Hz, 2 H), 6.90 (d, J = 8.5 Hz, 2 H), 6.83 (d, J = 8.5 Hz, 1 H), 5.56 (s, 1 H), 3.82 (s, 2 H); ^{13}C NMR (125.7 MHz, DMSO-d_6) δ 172.31, 171.90, 171.74, 159.72, 159.51, 158.01, 136.82, 136.14, 135.91, 134.42, 133.94, 130.87, 130.09, 129.89, 129.35, 128.17, 117.42, 117.05, 112.69, 112.48, 112.22, 52.42, 33.86; FABMS m/z (relative intensity) 575 (MH^+, 4.9), 557 ($MH^+ - H_2O$, 7.7). Anal. Calcd for $C_{30}H_{22}O_{12} \cdot H_2O$: C, 60.81, H, 4.08. Found: C, 61.00; H, 4.00.

Component A-13 (2): mp 202–204 °C dec; UV (ethanol) 309 (ε 13 090), 526 (ε 2802); the molar absorptivity of A-13 at 526 nm gradually decreased over time; IR (KBr) 3651–2373, 1676, 1580, 1489, 1436, 1354, 1297, 1211, 1086, 964, 846, 800, 675, 613 cm^{-1}; ^1H NMR (500 MHz, DMSO-d_6) δ 14.1–13.4 (broad peak, exchangeable with D_2O, 2 H), 11.1–11.3 (broad peak, exchangeable with D_2O, 3 H), 7.63 (d, 2 H J = 2.6, 3 H), 7.29 (dd, J = 8.7 and 2.6 Hz, 3 H), 6.92 (d, J = 8.8 Hz 3H); FABMS m/z (relative intensity) 423.2 (MH^+, 3), 405.2 (1), 387.3 (1); ^{13}C NMR (125.7 MHz, DMSO-d_6) δ 171.67, 159.90, 138.20, 135.08, 128.81, 116.58, 111.87, 79.00. Anal. Calcd for $C_{22}H_{14}O_9 \cdot H_2O$: C, 60.00; H, 3.66. Found: C, 60.29; H, 3.66.

Component A-15 (12): mp 197–199 °C dec; UV (ethanol) 309 (ε 17 313), 527 (ε 1726); the molar absorptivity of A-15 at 527 nm gradually decreased over time; IR (KBr) 3600–2300, 1678, 1586, 1491, 1439, 1354, 1293, 1206, 1084, 844, 801, 672 cm^{-1}; ^1H NMR (500 MHz, DMSO-d_6) δ 11.65 (s, exchangeable with D_2O, 1 H), 11.23 (s, exchangeable with D_2O, 2 H), 11.09 (s, exchangeable with D_2O, 1 H), 7.64 (d, J = 2.5 Hz, 2 H), 7.60 (d, J = 2.2 Hz, 1 H), 7.50 (d, J = 2.5 Hz, 1 H), 7.28 (d, 1 H), 7.27 (dd, 2 H), 7.25 (dd, 1 H), 6.90 (d, J = 8.8 Hz, 2 H), 6.80 (d, J = 8.5 Hz, 1 H), 3.83 (s, 2 H); ^{13}C NMR (125.7 MHz, DMSO-d_6) δ 172.30, 171.78, 171.74, 159.93, 159.38, 158.21, 138.18, 137.68, 135.80, 135.75, 135.03, 130.74, 129.82, 128.830, 128.34, 129.96, 116.92, 116.50, 112.43, 111.91, 111.44, 79.07, 33.95; FABMS m/z (relative intensity) 572 (MH^+, 39), 307 (18), 257 (16), 232 (13), 215 (31), 202 (50), 185 (100). Anal. Calcd for $C_{30}H_{20}O_{12} \cdot 1.5H_2O$: C, 60.11; H, 3.86. Found: C, 60.46; H, 3.78.

Acknowledgment. This investigation was supported by Contract NO1-CM-17513, awarded by the National Cancer Institute, DHHS. Anti-HIV screening was performed under the auspices of the Developmental Therapeutics Program, Drug Synthesis and Chemistry Branch, Dr. Ven L. Narayanan, Chief, and the results were provided by Dr. John P. Bader, Chief, Antiviral Evaluations Branch.

Registry No. 1, 69-72-7; **2a**, 4431-00-9; 4, 5578-67-6; 5, 122-25-8; 6, 129749-38-8; **7a**, 141511-21-9; **7b**, 141511-26-4; 8, 141511-22-0; 9, 141511-23-1; 10, 141511-24-2; 11, 639-56-5; **12a**, 141511-25-3; **12c**, 141511-27-5.