Pyruvate Aldol Condensation Product: A Metabolite That Escaped Synthetic Preparation for Over a Century

Andro C. Rios,* Partha P. Bera, Jennifer A. Moreno, and George Cooper

ABSTRACT: The homoaldol condensation product of pyruvate, 2-methyl-4-oxopent-2-enedioic acid (OMPD), has been recently implicated as a catabolic intermediate in the bacterial degradation of lignin and previously identified from other biological sources in reports ranging over 60 years. Yet, while a preparation of the pyruvate homoaldol product precursor, 4-hydroxy-4-methyl-2-oxoglutaric acid (HMOG/Parapyruvate), was first reported in 1901, there has not been a complete published synthesis of OMPD. Analyses of reaction mixtures have helped identify zymonic acid, the lactone of HMOG, as the direct precursor to OMPD. The reaction appears to proceed through an acid- or base-mediated ring opening that does not involve formal lactone hydrolysis. In addition to a preparative protocol, we provide a proposed mechanism for the formation of methylsuccinic acid that arises from the nonoxidative decarboxylation of OMPD. Finally, we calculated the relative stability of the isomers of OMPD and found Z-OMPD to be the lowest in energy. These computations also support our observations that Z-OMPD is the most abundant isomer across a range of pH values.

INTRODUCTION

Keto acids are a vital class of molecules in central carbon metabolism, and their biological chemistry as well as their synthetic building block utility have been extensively studied.1–3 Pyruvate (1) or, in its anionic form, pyruvate is the most prominent of these keto acids given its role as an anabolic/catabolic junction for many metabolic pathways. The degradation of lignin by soil bacteria catalyzes large insoluble aromatic materials into primary central carbon metabolites such as pyruvate. Understanding the detailed molecular transformations that the lignocellulosic material undergoes during this process has been of interest to identifying renewable sources of biofuel feedstock chemicals.4 Enzymes from soil bacteria responsible for these transformations funnel this aromatic material into precursor compounds that are easily converted into pyruvate.5 One of these precursors, 4-carboxy-2-hydroxy-pyenta-2,4-dienoate (CHPD, 7/8), is the enol tautomer of the homoaldol condensation product of pyruvate, 2-methyl-4-oxopent-2-enedioic acid (OMPD, 4/5) (Scheme 1), and had been proposed to occur in these transformations, over twenty years ago.6 Only recent evidence of its occurrence in the 5,5′-dehydrovanilllate (DDVA) pathway had been reported.6,7 The first study7 reported an enzymatic production of 7/8 from its aldol product precursor, 4-hydroxy-4-methyl-2-oxoglutaric acid (HMOG/parapyruvic acid, 2), with analysis by high-performance liquid chromatography (HPLC) but without verification with an authentic analytical standard. The more recent work6 also reported evidence of CHPD based on the enzymatic fragmentation of an upstream lignin-derived aromatic compound that leads to the suspected CHPD and 5-carboxyvanilllate. While a commercially available standard was used for verification of 5-carboxyvanilllate, the identity of CHPD was gained by analyzing a possible matching mass spectrum, the use of the known aromatic precursor of CHPD in the enzymatic reaction, along with an HPLC analysis using 2,4-DNP as a diagnostic derivatization reagent. In a review of the literature to determine if an analytical standard for CHPD/OMPD might have been accessible, we found that a synthetic procedure has never been reported. This was surprising given that the chemistry of pyruvate has been studied since the nineteenth century.8

There is one recent publication related to the production of esters of 4-MEOG (6) and E-OMPD (5) (Scheme 1) while exploring the scope of Lewis acid-based zeolite catalysts.9 Ethyl pyruvate was subjected to various synthesized zeolite catalysts, and reported yields of the esters ranged from 50 to 68% as determined by gas chromatography–mass spectrometry (GC–MS) analysis of crude reaction mixtures. Purification by chromatography was mentioned, but isolated yields were not reported for any compound, and the presented mass and NMR spectra were incomplete. The primary objective appeared to have been on the demonstration of the chemistry of zeolite catalysis rather than a complete synthesis of any one compound. There have been reports documenting the isolation of OMPD or 4-MEOG from biological sources. In 1955, 4-MEOG was isolated from the seedling plants that produce peanuts, and its
structure was derived from chemical tests and transamination studies.\textsuperscript{10} In the late 1980s, Z-OMPD (4) was isolated from tulips and the structure was confirmed by NMR, but a synthetic protocol was not reported.\textsuperscript{11} Thus, we concluded that the community could benefit from a complete and simple preparation of this metabolite.

Scheme 1. Structures of All Compounds Used in This Study\textsuperscript{a}

| Compound | Structure |
|----------|-----------|
| 1 | Pyruvic acid |
| 2 | HMOG or Parapyruvic acid |
| 3 | Zymonic acid |
| 4 | (Z)-OMPD |
| 5 | (E)-OMPD |
| 6 | 4-MEOG |
| 7 | (E)-CHPD |
| 8 | (Z)-CHPD |
| 9 | Methylsuccinic acid |
| 10 | Methyl pyruvate |
| 11 | Methyl zymonate |
| 12 | Monoester acid of OMPD |
| 13 | Citraconic acid |
| 14 | Mesaconic acid |
| 15 | Itaconic Acid |
| 16 | Dimethyl-2-methylsuccinate (DMMS) |

\textsuperscript{a}Isomers of the homoaldol condensation product of pyruvate are 2-methyl-4-oxopent-2-enedioic acid (OMPD, 4 and 5), 4-methylene-2-oxoglutamic acid (4-MEOG, 6), and 4-carboxy-2-hydroxypenta-2,4-dienoate (CHPD, 7 and 8).

Scheme 2. Summary of the Reaction Pathways of Pyruvic Acid under Aqueous Conditions Previously Known\textsuperscript{2,12–15,21,22} and Newly Identified Here\textsuperscript{a}

\textsuperscript{a}While all keto acid structures are shown as neutral and unhydrated at the keto position to maintain simplicity, their actual states (hydration, tautomer, and ionic) vary with pH in aqueous solutions.

\textbf{RESULTS AND DISCUSSION}

The synthesis of HMOG (2) is known and has been updated\textsuperscript{12} and appears to be the obvious starting point to OMPD, yet attempting to access the condensation product is not straightforward. Acidification of HMOG is known to facilitate a lactonization reaction forming an isotetronic acid derivative called zymonic acid (3)\textsuperscript{13} (Scheme 2). In fact, even in a neat liquid of pyruvic acid, zymonic acid itself is a prominent species that spontaneously forms and, over extended periods of time in...
cold storage, can crystallize directly in commercial source bottles (Figure S18). With synthetic access to HMOG and zymonic acid, we were able to explore the conditions that might lead to the production of the condensation products. Particularly important to this endeavor was the utility of liquid chromatography–mass spectrometry (LC–MS) to facilitate compound identification and reaction monitoring (See the Supporting Information and Table S1). Subjecting freshly prepared HMOG to 0.1 M HCl at room temperature resulted in the expected formation of zymonic acid (Figure S1A). Directly heating (60 °C) this acidic HMOG solution facilitated the progress of the lactonization, but we also observed the generation of new peaks containing parent masses consistent with condensation products (Figure S1B). Oxidation of this mixture using hydrogen peroxide resulted in the production of citraconic (13) and a smaller abundance of itaconic (15) acid, which are the expected oxidative decarboxylation products from two of the three possible isomers of the keto diacid condensation precursors (Figure S2). This led us to postulate that zymonic acid itself may be the more important precursor to the desired products. Subjecting zymonic acid to the same reaction conditions (0.1 M HCl) revealed that an acid-catalyzed ring opening (Figure S3) may in fact be accessing an alternative pathway (Scheme 2) to the condensation products, 4–8.

Exploring the reactivity under alkaline conditions (aqueous carbonate buffer 0.1 M, pH 10) demonstrated that zymonic acid produces both the expected HMOG from lactone hydrolysis and OMPD peaks (Figure S4A). However, subjecting HMOG alone under these conditions, we observed mostly the formation of pyruvate, resulting from a base-catalyzed retro-aldol reaction rather than the elimination step (Figure S4B). The prominence of zymonic acid formation in aqueous solutions from HMOG across a range of pH values was studied in a recent report using NMR spectroscopy. It was demonstrated that at low pH and ambient temperature, dominant species consisted mainly of the lactone and HMOG tautomers with very low signals of a possible aldol condensation product, 4. The report is consistent with our results as we observed that zymonic acid had to be subjected to higher temperatures (>60 °C) for extended periods (>2 h) before the appearance of condensation products.

With this knowledge, we decided to focus on the utility of a zymonic acid derivative as the precursor to the preparation of the desired compound, OMPD. Since homoaldol reactions of pyruvate esters in organic solvents are reported to be prone to spontaneous lactonization, we sought the use of methyl pyruvate ester, 10, to experiment with base-mediated, heated, aldol reactions. After some experimentation with solvents (acetoneitrile, dichloromethane, and N,N-dimethyl formamide) and organic bases (DBU, triethylamine) discussed in the previous reports, we found that simply heating methyl pyruvate, 10, in a methanolic suspension of sodium carbonate leads to the cleanest production of the monoester acid condensation product, 12 (Scheme 3). Monitoring of the reaction mixture by thin-layer chromatography revealed that the first detectable species to form was the methyl zymonate lactone ester, 11, and only upon extended reflux did the monoester acid product, 12, appear with the simultaneous disappearance of the lactone ester, 11. We also subjected a freshly prepared racemic 11 (Scheme 3) to the reaction conditions and found the same result leading to the monoester acid, 12. In both cases, the growing abundance of 12 coincided with a yellow-orange color in the reaction mixture. The mid to moderate yields of 12 in both routes (a and b) were attributed to the significant formation of dimethyl-2-methylsuccinate (DMMS, 16), resulting from the subsequent decarboxylation of 12, detailed in Figure S19. The abundance of DMMS expectantly increased as reflux times or temperatures were increased. We minimized the formation of DMMS in route b using lower temperatures but longer reaction times to improve the yield of 12 in comparison to route a. Taking advantage of the ionizable group in 12 that is not accessible for 16, a simple purification using pH-controlled liquid–liquid extraction resulted in its isolation. Subjecting 12 to acidic hydrolysis leads to Z-OMPD, and the stereochemistry was confirmed by an oxidative decarboxylation producing only citraconic acid (Figure S5). The Z-OMPD synthetic standard was matched to the dominant chromatographic peak identified in reaction mixtures starting from zymonic acid, as shown in Figure 1.

Attempts to selectively isolate E-OMPD (5) or 4-MEOG (6) have so far been unsuccessful. Chromatographic identification of these isomers was made however by the observed oxidation products from mixtures that contained the known Z-OMPD and one of the other isomers. Z-OMPD dominates the product mixtures likely due to isomerization from 5 and 6 (see computational discussion below), but some general observations can be made. A qualitatively higher abundance of 6 can be observed by LC–MS in reaction mixtures from heated solutions of zymonic acid at pH 4–5 (Figure 1) or from a solvent specific hydrolysis of compound 2 (Figure S6); however, the persistence of 6 in the reaction mixture is not long and appears to isomerize in solution after 3–4 h. At pH values > 7, this interconversion is even faster. Heating a solution of zymonic acid at pH 7 leads mostly to the production of Z-OMPD, but a qualitatively small amount of the suspected 5 isomer was formed and identified by the production of mesaconic acid, 14, upon oxidation (Figure S7). Only tentative assignments can be made for the E/Z enol tautomers of CHPD, 7/8 (Table S1), being the suspected unresolved shoulder that emerges from the Z-OMPD peak in Figure 1.

Extended heating of solutions of HMOG, zymonic and Z-OMPD will ultimately result in the production of methysuccinic acid (9) via a decarboxylation reaction. Methysuccinic acid had
historically been associated with the pyrolysis of tartaric acid, of which pyruvic acid was also a known decomposition product. The observation that 9 was originating from HMOG was first reported from studies that subjected pyruvic acid to high-temperature and high-pressure conditions. This may very well be the same pathway for how methylsuccinic acid is produced from tartaric acid. Reaction monitoring of heated (75 °C) acidic HMOG solutions eventually leads to 9 via the prior emergence of zymonic and Z-OMPD (Figure 2A). This was also supported by subjecting zymonic acid under the same conditions and observing the formation of Z-OMPD followed by methylsuccinic acid (Figure S8B). Expectantly, subjecting isolated Z-OMPD directly to the heated acidic solution (0.1 M HCl, 75 °C) produced methylsuccinic acid (Figure S8C). We propose that the decarboxylation reaction may be occurring predominantly from a hydrated keto Z-OMPD precursor (Figure 2B). Keto acids typically exist as hydrates at low pH, and the hydrated Z-OMPD resembles a β-γ unsaturated dicarboxylic acid structure. As a class, β-γ unsaturated acids are known to undergo thermal decarboxylation reactions, therefore, the hydrated Z-OMPD appears to provide a reasonable pathway to the saturated methylsuccinic acid.

To understand the energetic relationship among the isomers of OMPD and decarboxylation leading to methylsuccinic acid, we used density functional theory (DFT) calculations using the B3LYP hybrid density functional along with the 6-31+G* basis set to explore their differences. A comparison of the five isomers is shown in Figure 3, and reveals Z-OMPD, 4, as the lowest in energy. The trans keto isomer, E-OMPD, 5, is 8.6 kcal/mol higher and 4-MEOG, 6, is another 2.9 kcal/mol above the trans. E-CHPD, 7, sits at the highest energy state among all isomers being 13.7 kcal/mol above Z-OMPD, while the cis-enol, Z-CHPD, 8, is slightly lower than 5 by about 0.9 kcal/mol. Of the keto isomers, 4 and 5 can each access hydrated states that promote a decarboxylation reaction as described previously in Figure 2B. Yet, for 5, an additional increase in 1.1 kcal/mol to its hydrated state may be less favorable in comparison to an alternative, lower-energy route. Keto-enol tautomerization facilitates isomer interconversion (detailed in Scheme S1) and might be responsible for interconverting 5 through an energetic downhill pathway via the Z-CHPD enol, 8, to Z-OMPD and ultimately funneling a decarboxylation through the Z-OMPD hydrate isomer. This downhill isomeric interconversion might also help explain the prominence of Z-OMPD in our reaction mixtures and the observed decrease in abundance of the other keto isomers over time.

**CONCLUSIONS**

In conclusion, while the aldol product of pyruvate has been synthetically accessible for over a hundred years, the preparation of OMPD had been more elusive. The unexpected pathway to access 4 via the lactone, 3, and not formally the aldol product, 2, could be one of the reasons for this delay. The ease of
methylsuccinic acid, 9, production arising from decarboxylation of 4 in heated acidic solutions may be another. OMPD was initially discovered from niche areas of biology,10,11 but recently has been found to be generated by a class of bacterial enzymes participating in catabolic pathways that break down widely abundant lignin.4–8 Generating OMPD/CHPD may benefit organisms by providing yet another entry point to central carbon metabolism by ultimately producing two molecules of pyruvate. Our preparation of Z-OMPD as a free acid reveals this isomer to be the dominant of five possible isomeric forms in solution, a result supported by gas-phase computational evidence. At pH 7 and higher, however, there might be an increase in the enol (7/8) content, which was observed by a developing chromatographic shoulder peak of Z-OMPD (in Figure 1) and visually with an increased orange color to these solutions. In terms of identities at physiological pH, it is reasonable to propose that both Z-OMPD, 4, and CHPD, 7/8, isomers are the most relevant metabolic species.

■ EXPERIMENTAL SECTION

Materials and Methods. Reagents were purchased from Sigma-Aldrich, VWR, Acros and were used without further purification. Solvents and NMR solvents were purchased from VWR. Synthesis reactions were monitored with analytical thin-layer chromatography (TLC) performed on precoated silica gel aluminum-backed plates (Merck Kieselgel 60 F254). Column chromatography was performed with silica gel particle size 40–63 μm. Chemicals obtained from synthases were purified by flash column chromatography or described otherwise. Origin 2018 graphing program analysis software (OriginLab Corporation, North Hampton, MA) was used to replot data from chromatograms, mass spectra, and reaction monitoring graphs. NMR spectra were obtained on a Bruker 500 MHz NMR with an Avance III HD console located at the NMR Facility at the University of California Santa Cruz. LC–MS/electrospray ionization-MS analyses for commercial and synthesized compounds were performed on a Thermo Finnigan LCQ Deca Max XP (Thermo-Fisher San Jose) equipped with an electrospray ionization source probe assembly with detection in the negative mode polarity. Detailed methods are described in the Supporting Information.

Experimental Details for Figures 1–3. Figure 1 Panel A: In a 5 mL volumetric flask was added a solution containing 51 mg of zymonic acid, 3, in 0.1 M sodium acetate buffer (pH 4) and readjusted to pH 4 with the dropwise addition of 1 M sodium hydroxide before filling to the mark. The solution was transferred to a reaction vial and placed in a prewarmed 60 °C heating block and allowed to react with periodic sampling, about once every 24 h. Sampling consisted of removing 150 μL of the reaction mixture, and remaining procedures were followed as described under LC–MS methods in the Supporting Information. The chromatogram shown in A was taken after 69 h of reaction time.

Figure 1 Panel B: In a 5 mL volumetric flask was added 15 mg of synthetically prepared Z-OMPD, 4, oil dissolved in water, and pH was not adjusted. Sampling consisted of removing 150 μL of the reaction mixture, and the remaining procedures were followed as previously described under LC–MS methods in the SI.

Figure 2 Panel A: In a 10 mL volumetric flask was added 55 mg of HMOG disodium salt, 2, and dissolved in 0.1 M hydrochloric acid. After immediate, thorough mixing, the solution was transferred to a reaction vial with a cap and placed in a preheated 75 °C heating block and allowed to react. At designated time points, 50 μL of the reaction mixture was removed quickly before recapping the reaction vial and the aliquot was added to 200 μL of deionized water and allowed to cool further for 1 min before continuing with sample procedures previously described under LC–MS methods in the SI.

Figure 3 Density functional theory calculations were performed using the widely used B3LYP density functional along with the 6-31+G* basis set. The zero-point vibrational energy-corrected energies of the reactants, intermediates, and products were plotted in the reaction diagram.

■ ASSOCIATED CONTENT

* Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.0c00877.

General procedures; LC–MS methods and summary table; additional supporting figures and schemes; synthetic experimental and NMR Spectra (PDF)

■ AUTHOR INFORMATION

Corresponding Author
Andro C. Rios – Exobiology Branch, Space Science and Astrobiology Division and Center for the Emergence of Life, NASA Ames Research Center, Moffett Field, California 94035, United States; Blue Marble Space Institute of Science, Seattle, Washington 98154, United States; orcid.org/0000-0001-5779-8306; Email: andro.c.rios@nasa.gov

https://dx.doi.org/10.1021/acsomega.0c00877
ACS Omega 2020, 5, 15063–15068
15067
The authors would like to thank Joseph Solvason, Mastewal Abate, Samantha Yim, and Melissa Thang for early explorations in this work. This investigation was supported by a grant (14-EXO14_2-0064) from the Exobiology Program solicitation #NNH14ZDA001N-EXO and an Internal Scientist Funding Model award for “Evolutionary processes that drove the emergence and early distribution of life” from the National Aeronautics and Space Administration.

Notes
The authors declare no competing financial interest.