# Synthesis and Structure–Activity Relationships of Sweet 2-Benzoylbenzoic Acid Derivatives

Anna Arnoldi, Angela Bassoli,* Gigliola Borgonovo, Lucio Merlini, and Gabriella Morini

Dipartimento di Scienze Molecolari Agroalimentari, Sezione di Chimica, Università di Milano, Via Celoria 2, I 20133 Milano, Italy

Twenty-four analogues of the sweet compound 2-(4-methoxybenzoyl)benzoic acid **1** were synthesized and tasted. The structure–sweet taste relationships were studied by means of principal component analysis and by comparison with the existing sweet receptor models. Three possible glucophores were identified, which could correspond to the sites B, E1, and E2 of the Tinti–Nofre model. Some similarities between this class of compounds and isovanillic sweeteners were found.

**Keywords:** *Sweetness; SAR (structure–activity relationships); taste*

## INTRODUCTION

Some 2-ketocarboxylic aromatic acids, such as 2-acetyl- and 2-benzoylbenzoic acid, have been known to be sweet for about 120 years (Gabriel and Michael, 1877). In the years following several derivatives of these compounds were synthesized and tasted by different authors (Cohn, 1914; Möhler, 1950). The most important term of the series was found to be 2-(4-methoxybenzoyl)benzoic acid, **1**, extensively studied by Möhler (1950). Its sodium salt is 150 times as sweet as sucrose, has a good solubility in water and a good taste profile, although it elicits an unpleasant bitter aftertaste at concentrations greater than 0.2 g/L. Its low chemical reactivity and its thermal stability made this compound a good candidate for use as a sweetener, and in fact it had some commercial applications in Germany in the 1950s with the name of S 23/46 (Möhler, 1950).

The first study of the structure–activity relationship in a series of analogues of **1** was reported by Runti and Galimberti (1957). Some modifications were made on the ketonic and carboxylic groups, but none of the new derivatives resulted in a compound sweeter than **1** and the authors could only draw the conclusion that the concomitant presence of the ketonic and carboxylic functions was necessary to elicit the sweet sensation. The same authors (Runti and Collino, 1964) reported that the *p*-methoxy substituent was more active than any longer alkoxy group on the phenyl moiety. In a paper on the metabolism of compound **1**, Thomas and Savelsberg (1949) noted that it is excreted without modifications both in rats and in humans. Toxicological data for **1** were lacking in an FDA report (Lehaman, 1951), and, to our knowledge, since that time compound **1** has no longer been used for human consumption. The only subsequent literature citation concerning use of **1** as a sweetener is in a review by Hrdlicka (1973).

A general feature of these old references is the poor description of the procedures used during tasting. In the present work we synthesized some of these compounds together with new derivatives, tasted them following currently used procedures, and attempted to rationalize the sweet taste–structure relationship. For this purpose, the general model of sweet taste receptor proposed recently by Tinti and Nofre (1991) was used to rationalize the spatial disposition of the glucophores in the molecule and their interactions with the receptor. The results were also compared with those obtained by a statistical approach using principal component analysis on a homogeneous set of compounds.

Figure 1 shows the compounds that are the subjects of this study together with their relative sweetness and literature references on synthesis and taste (where available).

* Address correspondence to this author (phone +39-2-236 2721; fax +39-2-7063 3062; e-mail DISMA@imiucca.csi.unimi.it).

## MATERIALS AND METHODS

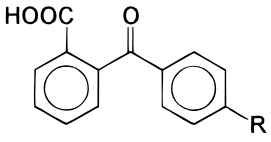
**General Procedures.** Melting points are uncorrected. Phthalic anhydride was sublimed before use. Flash column chromatography was performed on Merck 60 silica gel (230–400 mesh STM). Infrared spectra were recorded on a Perkin-Elmer 1310 IR spectrophotometer. $^1$H NMR spectra were recorded on a Bruker WP80SY (80 MHz) or a Varian XL300 (300 MHz) spectrometer using $Me_4Si$ as internal standard. Chemical shift values are in $\delta$ (ppm) and coupling constant values ($J$) are given in Hz. Mass spectra (electron impact) were recorded on a Finnigan-MAT TSQ70 spectrometer equipped with an ICIS data system.

Compounds **1**–**9**, **11**–**13**, **17**, **20**–**22**, **24**, **25**, and **29** were synthesized following literature methods (Figure 1).

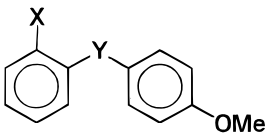
**2-(4-Nitrobenzoyl)benzoic Acid 10.** 4,4-Dimethyl-2-phenyloxazoline (1.9 mL, 11 mmol) in dry THF was added dropwise at −45 °C to a solution of butyllithium (1.6 M in hexane, 10.2 mL, 16 mmol) in THF (40 mL). After being stirred for 1 h at −45 °C, 4-nitrobenzaldehyde (1.70 g, 11 mmol) in THF (20 mL) was added. After 5 h the mixture was treated with a saturated solution of $NH_4Cl$ (40 mL). The organic phase was dried and evaporated, and the crude residue (3.24 g) was purified by column chromatography (hexane:ethyl acetate, 7:3).

2-[2′-(4-Nitrophenyl)hydroxymethylphenyl]-4,4-dimethyloxazoline, **30**, was obtained as a yellow oil (0.60 g, 17%). NMR ($CDCl_3$): $\delta$ 1.0 and 1.5 (2 × 3 H, 2 s, Me), 4.0 (2 H, m, $CH_2$), 6.0 (1 H, broad s, H-2′), 7.0–8.2 (8 H, m, arom.). MS *m*/*z* (%): 326 (43, $M^+$), 254 (100), 160 (78). Compound **30** was added to a solution of pyridinium chlorochromate (0.60 g, 2 mmol) and powdered anhydrous sodium acetate (3 mg, 0.4 mmol) in dry dichlorometane (60 mL) and stirred for 24 h at room temperature. The mixture was then treated with dry diethyl ether and filtered through Florisil. The solvent was removed, and the residue (0.52 g) was purified by column chromatography (hexane:ethyl acetate, 55:45) to give ketone **31** as a white solid (0.20 g, 34%). NMR ($CDCl_3$): $\delta$ 0.97 (2 × 3 H, 2 s, $CH_3$), 3.7 (2 H, s, H-3), 7.2–8.3 (8 H, m, arom.). IR ($cm^{-1}$) 1680 (C=N),

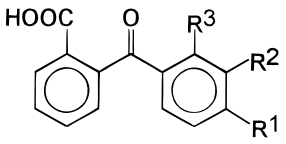
S0021-8561(96)00755-8 CCC: $14.00 



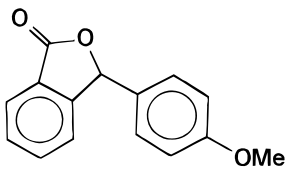

| n | R | taste(1) | potency(2) | ref.(3) |
|---|---|---|---|---|
| 1 | OMe | s | 150 | a |
| 2 | H | t | | a |
| 3 | Me | s | | a |
| 4 | OEt | s | | a |
| 5 | OPr | t | | b |
| 6 | OPh | t | | a |
| 7 | OH | s | | a |
| 8 | SMe | t | | b |
| 9 | $NMe_2$ | s | 200 | a |
| 10 | $NO_2$ | b | | |
| 11 | $CO_2H$ | t | | a |
| 12 | Cl | s | | a |
| 13 | Ph | t | | a |



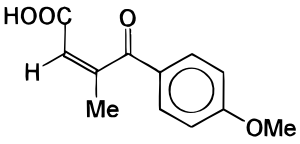

| n | X | Y | taste(1) | potency(2) | ref.(3) |
|---|---|---|---|---|---|
| 20 | $CO_2H$ | $CH_2$ | t | | c |
| 21 | $CO_2H$ | O | t | | d |
| 22 | $CO_2H$ | S | t | | e |
| 23 | $CO_2H$ | SO | s | 50 | |
| 24 | H | CO | t | | |
| 25 | $CO_2Me$ | CO | t | | f |
| 26 | $CH_2OH$ | CHOH | t | | |
| 27 | $CO_2H$ | COO | s | 10 | |




| n | $R^1$ | $R^2$ | $R^3$ | taste | potency | ref. |
|---|---|---|---|---|---|---|
| 14 | OH | OH | H | s | 100 | |
| 15 | OMe | OH | H | s | 250 | |
| 16 | OH | OMe | H | t | | |
| 17 | OMe | OMe | H | t | | a |
| 18 | OMe | H | OMe | b | | a |
| 19 | OMe | H | OH | t | | |




28 t




29 t g

**Figure 1.** Compounds **1**−**29**. (1) s = sweet; t = tasteless; b = bitter. (2) Sweet taste potency relative to 3% sucrose. (3) a = Cohn (1914); b = Runti and Collino (1964); c = Blicke (1932); d = Oi (1988); e = Roberts (1929); f = Runti and Galimberti (1957); g = Bowden and Henry (1971).

1650 (C=O). MS *m*/*z* (%) 324 (4, $M^+$), 295 (100). 2-[2-(4-Nitrophenyl)benzoylphenyl)]-4,4-dimethyloxazoline **31** was hydrolyzed by heating in 3% HCl (30 mL) for 3 h. Ethyl acetate and sodium bicarbonate were added. The aqueous layer was acidified, and acid **10** was extracted as a white solid (0.04 g, 24%). NMR ($CD_3OD$): $\delta$ 7.4 (1 H, d), 7.6−7.8 (3 H, m), 7.9 (2 H, d, $J$ = 16, H-2′ and 6′), 8.1 (1H, s broad, COOH), 8.03 (2 H, d, $J$ = 16, H-3′ and 5′). MS *m*/*z* (%): 271 (35, $M^+$), 210 (37), 181 (40), 149 (100).

**2-(3,4-Dihydroxybenzoyl)benzoic Acid (14) and 2-(3-Hydroxy-4-methoxybenzoyl)benzoic Acid (15).** 2-Methoxyphenol (2.00 g, 16 mmol) was dissolved in dry pyridine at 0 °C. Methanesulfonyl chloride (2.5 mL, 32 mmol) was added, and the mixture was stirred for 2.5 h and then treated with aqueous HCl and filtered. 1-Methoxy-2-methylsulfonyloxybenzene was obtained as a white powder (3.00 g, 94%), mp 30 °C. NMR ($CDCl_3$): $\delta$ 3.25 (3 H, s, $OSO_2Me$), 3.87 (3 H, s, OMe), 6.75−7.30 (4 H, arom.). This compound (1.60 g, 8 mmol) and $AlCl_3$ (1.67 g, 12 mmol) were added in portions to a solution of phthalic anhydride (1.16 g, 8 mmol) in dichloromethane at 0 °C. After being stirred at room temperature for 20 h, ice and HCl were added, the mixture was extracted with ethyl acetate, and the residue (2.30 g) was purified by column chromatography with hexane:ethyl acetate, 6:4. The fraction separated still contained two compounds by TLC [IR ($cm^{-1}$): 1750 (COOH), 1650 (CO), 1350 and 1100 ($SO_2$)], which were hydrolyzed in basic conditions and separated by column chromatography to give compounds **14** (180 mg, 9%) and **15** (240 mg, 11%).

Compound **14** has a melting point of 210 °C (decomp). NMR (DMSO-$d_6$): $\delta$ 6.8−8.1 (7 H, arom.); MS *m*/*z* (%): 258, (35, $M^+$), 257 (20), 211 (100), 151 (90).

Compound **15** has a melting poing of 95 °C. NMR (300 MHz, DMSO-$d_6$): $\delta$ 3.80 (3 H, s, OMe), 6.96 (2 H, s, H-6′ and 5′), 7.15 (1 H, s, H-2′), 7.34 [1 H, d, $J$ = 7, H-6(3)], 7.60 and 7.80 (2 H, 2 t, $J$ = 7, H-5 and H-4), 7.95 (1 H, d, $J$ = 7, H-3(6)], 9.50 (1 H, s, COOH). MS *m*/*z* (%): 272 (31, $M^+$), 178 (58), 151 (100).

**Structure Determination of Compound 15.** Compound **15** was reacted with acetic anhydride and pyridine to give the acetylated derivative **32**. NMR (300 MHz, DMSO-$d_6$): $\delta$ 2.45 (1 H, s, MeCO), 3.85 (3 H, s, OMe), 7.20 (1 H, d, $J$ = 8.5, H-5′), 7.38 [1 H, dd, $J$ = 7.5 and 1.5, H-3(6)], 7.41 (1 H, H-2′), 7.43 (1 H, dd, $J$ = 8.5 and 2, H-6′), 7.64 [1 H, dt, $J$ = 7.5 and 1.5, H-4(5)], 7.71 [1 H, dt, $J$ = 7.5 and 1.5, H-5(4)], 7.98 [1 H, dd, $J$ = 7.5 and 1.5, H-6(3)]. MS *m*/*z* (%): 314 (10, $M^+$), 272 (100), 254 (18), 225 (78), 196 (22), 95 (152). In compound **15** the *ortho* protons H-5′ and H-6′ are isochronous and appear as coincident singlets, and the signal at 7.15 is thus attributed to H-2′. In the acetylated derivative **32** the signal of H-5′ is shifted downfield by 0.24 ppm, H-6′ by 0.47 ppm, and H-2′ by 0.26

ppm. Therefore the acetyl group must be in 3′ position, i.e., compound **15** is 2-(3-hydroxy-4-methoxybenzoyl)benzoic acid.

**2-(4-Hydroxy-3-methoxybenzoyl)benzoic Acid (16) and 2-(3,4-Dimethoxybenzoyl)benzoic Acid (17).** $AlCl_3$ (2.50 g, 17 mmol) and 1,2-dimethoxybenzene (0.7 mL, 7 mmol) were added in portions to a solution of phthalic anhydride (1.00 g, 7 mmol) in dichloromethane (20 mL) cooled to 0 °C, and the resulting solution was stirred at room temperature overnight. The mixture was then acidified with HCl, extracted with ethyl acetate, and chromatographed to give compounds **16** (180 mg) and **17** (140 mg).

Compound **16** has a melting point of 194°C. NMR (DMSO-$d_6$): $\delta$ 3.8 (3 H, s, OMe), 6.78 (1 H, d, $J$ = 8, H-5′), 6.88 (1 H, dd, $J$ = 8 and 1.5, H-6′), 7.35 (1 H, d, $J$ = 1.5, H-2′), 7.55−7.95 (4 H, m, arom.), 10.1 (1 H, s, COOH). MS $m/z$ (%): 272 (55), 197 (25), 151 (100). MS-CI $m/z$ (%): 273 (100), 255 (7). IR ($cm^{-1}$): 1650.

Compound **17** has a melting point of 234 °C. NMR (DMSO-$d_6$): $\delta$ 3.85 (6 H, s, OMe), 6.9−9.1 (8 H, m, arom. and COOH). MS $m/z$ (%): 286 (10), 165 (70), 91 (100). Anal. Calcd for $C_{16}H_{14}O_5$: C, 67.12; H, 4.93. Found: C, 67.29; H, 4.92.

**Structure Determination of Compound 16.** HPLC analysis showed that compounds **16** and **15** were two single isomers. Compound **16** was reacted with acetic anhydride and pyridine to give the acetylated derivative **33**. NMR (DMSO-$d_6$): $\delta$ 2.3 (3 H, s, MeCO), 3.85 (3 H, s, OMe), 6.98 (1 H, dd, $J$ = 8 and 2, H-6′), 7.18 (1 H, d, $J$ = 8, H-5′), 7.52 (1 H, d, $J$ = 2, H-2′), 7.6−8.0 (4 H, arom.). In compound **16** the signal at 6.78 was attributed to H-5′ because it has only one vicinal coupling; in the corresponding acetylated derivative this signal was shifted downfield by 0.4 ppm, while the signals of H-2′ and H-6′ were shifted only slightly (0.16 and 0.1 ppm, respectively). Thus the proton H-2′ is *ortho* to the acetyl group, i.e., compound **16** contains a 4-hydroxy-3-methoxyphenyl group. Also the acetylated isomeric compounds **32** and **33** have been analyzed by HPLC and were shown to be pure.

**2-(2,4-Dimethoxybenzoyl)benzoic Acid (18) and 2-(4-Hydroxy-2-methoxybenzoyl)benzoic Acid (19).** $AlCl_3$ (2.30 g, 17.3 mmol) was added in portions at 0 °C to a solution of phthalic anhydride (1.00 g, 7 mmol) dissolved in 20 mL of dry $CH_2Cl_2$, and then 1,3-dimethoxybenzene was added dropwise. After the addition was complete the mixture was stirred at room temperature for 20 min and then poured into acid water and ice and extracted with ethyl acetate. The organic layer was washed to neutrality and dried, and the solvent was evaporated. Addition of ethanol favored the separation of crystalline phthalic acid which was eliminated by filtration. The crude residue was purified by column chromatography on silica gel (hexane:ethyl acetate, 7:3) to give separated **18** (0.76 g) and **19** (0.26 g).

The dimethyl derivative **18** was purified by crystallization from aqueous ethanol to yield 0.48 g of white crystals (40%); mp 159−160 °C. NMR ($CDCl_3$): $\delta$ 3.5 (3 H, s, OMe), 3.83 (3 H, s, OMe), 6.25−6.60 (2 H, d, $J$ = 9, H-3′ and H-5′), 7.15−8.04 (5 H, m, arom.), 8.75−9.1 (1 H, br, COOH). MS $m/z$ (%) 286 (20, $M^+$), 165 (100), 122 (18).

Compound **19** was crystallized from aqueous ethanol to yield a white solid (0.10 g, 5%), mp 154−156 °C. NMR ($CDCl_3$): $\delta$ 3.88 (3 H, s, OMe), 6.25 (1 H, dd, $J$ = 9 and 2.5, H-5′), 6.5 (1 H, d, $J$ = 2.5, H-3′), 7 (1 H, d, $J$ = 9, H-6′), 7.2−8.2 (4 H, m, arom.), 12.3 (1 H, s, COOH). MS $m/z$ (%): 272 (10, $M^+$), 151 (100). Anal. Calcd for $C_{15}H_{12}O_5$: C, 66.17; H, 4.93. Found: C, 66.16; H, 4.48.

The position of the methoxy group was determined by NMR with a nuclear Overhauser effect (NOE) experiment: the irradiation of the methoxy group showed a 19% NOE on H-3′ easily recognizable for the *meta J* value.

**2-(4-Methoxyphenylthio)benzoic Acid *S*-Oxide (23).** 2-(4-Methoxyphenylthio)benzoic acid **22** (0.30 g, 1 mmol) was stirred in the presence of 3-chloroperoxybenzoic acid (0.30 g, 1.5 mmol) in dichloromethane, at room temperature, for 24 h. The solvent was evaporated, and the residue was crystallized from ethanol to give compound **23** (0.20 g, 73%), mp 228 °C. NMR (DMSO-$d_6$): $\delta$ 3.8 (3 H, s, OMe), 7.0−8.3 (9 H, arom. and COOH). IR (Nujol) $cm^{-1}$: 1700. MS $m/z$ (%) 276 (5), 155 (42), 124 (100). Anal. Calcd for $C_{14}H_{12}O_4S$: C, 60.92; H, 4.35. Found: C, 60.93; H, 4.35.

**2-Hydroxymethylphenyl-4-methoxyphenylmethanol (26).** Methyl 2-(4-methoxybenzoyl)benzoate (0.14 g, 0.53 mmol) in 3 mL of THF was added dropwise to a solution of $LiAlH_4$ (1 M solution, 0.8 mL, 0.8 mmol) in dry THF (5 mL) under nitrogen. The mixture was stirred at room temperature for 3 h and quenched with a saturated solution of $NH_4Cl$. The solvent was removed, and the residue was acidified with concentrated HCl and extracted with ethyl acetate. The crude residue (90 mg) was crystallized from 5 mL of 96% ethanol to give a white solid (57%), mp 55−57 °C. NMR ($CDCl_3$): $\delta$ 3.80 (3 H, s, OMe), 5.25 (2 H, m, $CH_2$), 6.15 (1 H, s, m, CH), 7.1−7.3 (6 H, m, arom.). MS $m/z$ (%): 226 (71, M − $H_2O$), 225 (90), 210 (24), 195 (100), 165 (33), 152 (51), 135 (77).

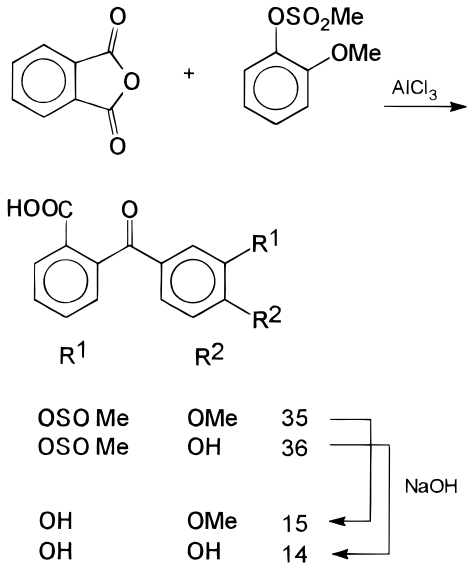
**Phthalic Acid Mono(4-methoxyphenyl) Ester (27).** This compound was prepared by melting together phthalic anhydride (3.60 g, 24 mmol) and 4-methoxyphenol (3.00 g, 24 mmol) for 5 h. The mixture was then treated with 5% NaOH (35 mL) and acidified with 10% HCl to separate unreacted phthalic acid. Addition of hexane to the filtrate precipitated the ester **27** as a white solid (0.37 g, 6%), mp 118−120 °C. NMR (DMSO-$d_6$): $\delta$ 3.8 (3 H, s, OMe), 7.0−7.2 (2 H, 2 d, arom.), 7.5−7.9 (6 H, arom.). MS $m/z$ (%): 149 (7, $M^+$), 124 (100). IR ($cm^{-1}$): 1750 (COOH), 1700 (CO).

**3-(4-Methoxyphenyl)benzo[2]furan-1-one (28).** 4,4-Dimethyl-2-phenyloxazoline (1.9 mL, 11 mmol) in dry THF (5 mL) was added to a solution of *n*-BuLi (1.6 M in hexane, 10.2 mL, 16 mmol) in dry THF (40 mL) at −45 °C. After 1 h, 4-methoxybenzaldehyde (1.38 mL, 11 mmol) in THF (10 mL) was added, and the temperature was kept at −45 °C for 5 h and then at −23 °C for 48 h. The mixture was quenched with an iced saturated solution of $NH_4Cl$ (10 mL). The organic layer was dried and concentrated under reduced pressure, and the crude residue (3.42 g) was chromatographed on silica gel (hexane:ethyl acetate, 7:3). 2-[2-(4-Methoxyphenyl)hydroxymethylphenyl]-4,4-dimethyloxazoline **34** was obtained as a yellow oil (1.46 g, 43%). NMR ($CDCl_3$): $\delta$ 1.0 and 1.5 (2 × 3 H, 2 s, Me), 3.8 (3 H, s, OMe), 4.0 (2 H, dd, $CH_2$), 5.9 (1 H, d broad, CH), 6.8−7.9 (8 H, arom.). MS $m/z$ (%): 311 (80, $M^+$), 280 (58), 240 (100), 181 (38), 152 (42), 137 (95), 135 (88), 105 (65). The lactone **28** was obtained by treating with HCl 3% (70 mL) oxazoline **34** (0.50 g, 1.6 mmol) at reflux for 5 h. The mixture was extracted with ethyl acetate, dried, evaporated, and purified by column chromatography (hexane:ethyl acetate, 7:3). The solid residue (0.30 g) was crystallized from hexane/cyclohexane to give 0.15 g (39%) of pure product, mp 110 °C. NMR ($CDCl_3$): $\delta$ 3.9 (3 H, s, OMe), 6.35 (1 H, s, CH), 6.8−8 (8 H, arom.). MS $m/z$ (%): 240 (100, $M^+$), 135 (48), 104 (40). IR ($cm^{-1}$): 1760 (CO).

**Tasting.** Although some derivatives had already been tasted (see references in Figure 1), the tasting procedures were not exhaustively described. In most cases it was unspecified whether the compounds had been tasted as solids or in water solution, as free acids or as sodium salts and moreover the overall taste was often described as a mixture of different sensations (for instance as "sweet-bitter" or "first sweet, then bitter"), preventing the use of these data even for a qualitative correlation study.

The compounds assayed in this work result from minor structural modifications of compound **1** and therefore were not submitted to toxicological evaluation. However they were tasted only once with the "sip and spit" procedure at a starting concentration of 200 mg/L. The solutions for the tasting trials were obtained by dissolving the compounds (10 mg) in freshly distilled water (50 mL). Derivatives **1**, **7**, **15**, **20**, **22**, **23**, **26**, and **27** were tasted both as free acids and as the corresponding sodium salts; in both cases the taste and the sweetness intensity remained substantially unchanged, but the taste of the sodium salts was a little less persistent. A panel of five to seven untrained people tasted the solutions in comparison with 3% sucrose in water to assess the sweet taste potency. Compounds **3**, **7**, and **12** were described as sweet by Cohn (1914): in our tasting trials they were tasteless at a concentration of 0.2 g/L but were perceived as sweet when tasted as a solid.

**Scheme 1**




RESULTS

**Synthesis of Compounds.** The reaction of phthalic anhydride and $AlCl_3$ with 1,2-dimethoxybenzene gave a 1:1 mixture of two products: the expected product **17** and a monomethylated derivative **16** whose structure was assigned on the basis of the chemical shifts of aromatic protons in the corresponding *O*-acetyl derivative **30**. The demethylation, catalyzed by $AlCl_3$, occurred on compound **17** which was the primary reaction product and not on 1,2-dimethoxybenzene, as shown by monitoring the reaction by TLC and by the fact that the relative amount of this side product increases after aqueous workup and at higher reaction temperature. Similar selective demethylations in the *para* position of aromatic carbonyl derivatives have been already described in the literature (Jaszberenyi, 1988).

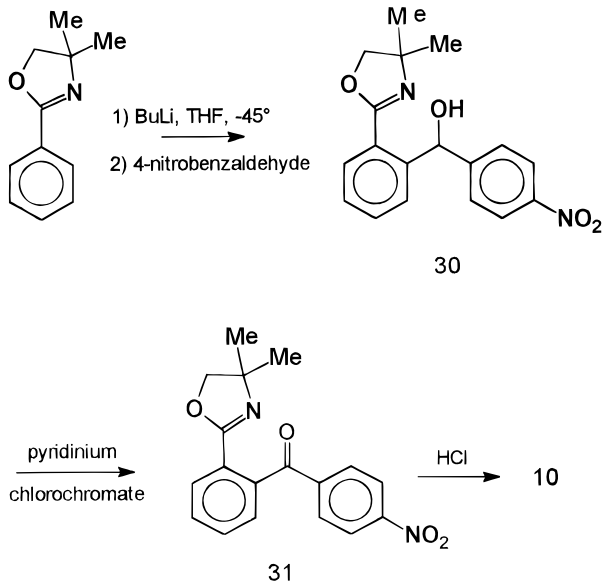
A similar reaction occurs with 1,3-dimethoxybenzene, which gave an 8:1 mixture of compounds **18** and **19**, whose structures were established by NOE experiments.

The synthesis of compound **15** was obtained by protecting guaiacol as the methanesulfonate and reacting the latter with phthalic anhydride (Scheme 1). The resulting mixture of acylated derivatives **35** and **36** was submitted to basic hydrolysis and chromatographed to give compound **15** and the demethylated derivative **14**. Compound **15** is an isomer of compound **16** in which the hydroxy and the methoxy substituents are exchanged; they differ in $^1$H NMR spectra and their acetyl derivatives **32** and **33** in $^1$H NMR and HPLC retention times.

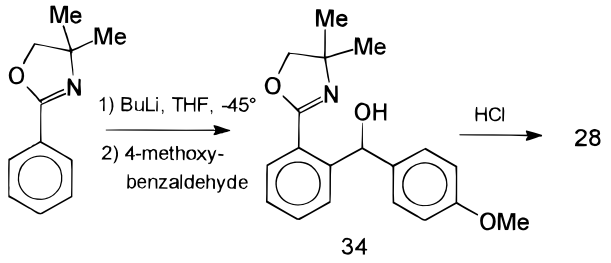
A different synthetic strategy was applied to the preparation of compound **10** which could not be obtained directly by Friedel−Craft acylation (Scheme 2). 2-Phenyloxazoline was lithiated in the *ortho* position (Meyers and Mihelich, 1975) and reacted with 4-nitrobenzaldehyde. The obtained alcohol, **30**, was oxidized to the corresponding ketone **31** with pyridinium chlorochromate and then hydrolyzed to compound **10**.

Lithium 2-phenyloxazoline was reacted also with 4-methoxybenzaldehyde (Scheme 3) to give alcohol **34** whose hydrolysis gave lactone **28** directly. This compound could not be prepared starting directly from **1**: every attempt to reduce the carbonyl function with standard methods ($NaBH_4$, 9-BBN, Li-Selectride) was ineffective. Similarly compound **1** was inert toward reductive amination and hydroxylamine. The low reactivity of the ketonic function in this class of compounds is well-known (Runti and Galimberti, 1957).

**Scheme 2**




**Scheme 3**




Compound **27** was prepared by melting a mixture of phthalic anhydride and 4-methoxyphenol (Bishoff and von Hedenstrom, 1902).

Oxidation of the sulfide **22** with 3-chloroperoxybenzoic acid gave the sulfoxide **23**; further oxidation to the corresponding sulfone could not be obtained.

**Calculation of Steric, Lipophilic, and Electronic Parameters: Principal Component Analysis.** A principal component analysis (PCA) on the physicochemical characteristics of a certain number of sweet and nonsweet compounds of this class was performed.

It is well-known that the taste response depends on the size, shape, and functionalities of the whole molecule, but, as this work concerns a set of derivatives of a parent compound in which the phthalic portion of the structure remains constant, some of the parameters selected refer only to the substituents on the other aromatic ring. Initially a large number of physicochemical and structural properties were identified and considered: the substituent hydrophobic constant ($\pi$), the electronic Hammett constant ($\sigma$), the p$K_a$ value, the van der Waals volume (vdW), the fourth-order molecular connectivity, the molar refractivity (MR), and the five STERIMOL parameters ($L$, $B_1$, $B_2$, $B_3$, and $B_4$) developed by Verloop et al. (1976).

The values of $\pi$, $\sigma$, MR, and the STERIMOL parameters were obtained from files compiled by Hansch and Leo (1979). Dipole moments and van der Waals volumes were obtained by molecular modeling systems. The molecular models were built on a Silicon Graphics IRIS 20, using the program INSIGHT II,1.1.0 (Biosym Technologies, San Diego, CA), and the initial model was

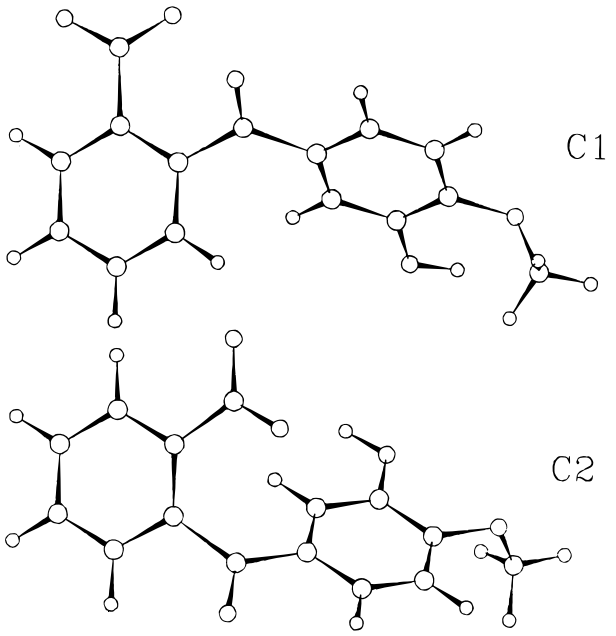



**Figure 2.** Minimum energy conformations C1 and C2 for compound **15**.

refined by molecular mechanics techniques; the DISCOVER program (Biosym Technologies, San Diego, CA) was used to generate low-energy conformations and the electrostatic properties of the molecules were calculated using the program MOPAC 4.0. With these molecular mechanics programs, for all of the compounds two main conformations (C1 and C2) were identified with very similar relative minimum energy values. These conformations are represented in Figure 2 for the sweetest derivative **15**. Dipole moments and van der Waals volumes, which depend on conformation, were calculated for both the conformations.

Molecular connectivity indices of the fourth order, a set of parameters related to molecular shape, size, and electronic properties of the molecules, derived by Kier and Hall (1976), were computed using the program Graph III (Sabljić, 1989).

This study was conducted on the compounds **1**−**17** (Figure 1). This choice was made considering the availability of the physicochemical data (for example, all the analogues that were *ortho*-substituted with respect to the keto group were excluded), and only compounds directly tasted by us or reported unambiguously in the literature as "sweet", "bitter", or "tasteless" were included.

Performing a preliminary statistical analysis, redundant or strongly correlated parameters were found and eliminated. As a result, the properties finally considered were $\pi$, $\sigma$, dipole moment, van der Waals volume, fourth-order molecular connectivity, $L$, and $B_1$.

The data matrix shown in Table 1 was used as starting point for principal component analysis performed with the program Systat 5.0 (Systat Inc.), in the attempt to determine which variables (properties) are important for eliciting the sweet taste and eventually to find a correlation between molecular structure and sweet taste.

Tables 2 and 3 show respectively the component loadings and the percent of total variance explained by the factors obtained for conformation C1. Tables 4 and 5 show the corresponding data obtained for conformation C2. From the data in Tables 2 and 4, factor 1 can be described as a "size" factor, the loadings of molecular connectivity, van der Waals volume, $L$, and $\pi$ being the largest. The loadings of factor 2 are dominated by $\sigma$, therefore by the electronic properties of compounds, whereas in factor 3 dipole moment has the largest loading. Tables 3 and 5 indicate that it may be enough to consider factor 1 and factor 2 for a description of our system: the sum of these two factors accounts in both cases for 75% of the variance in the space defined by these seven variables.

That these results are similar for both conformations C1 and C2 depends from the fact that the only parameter which distinguishes significatively between the two conformations is dipole moment, which is important only in factor 3. For this reason in Figure 3 the score plot of factor 1 versus factor 2 is reported only for conformation C1.

Figure 3 shows that all the sweet compounds, except only for **9**, are clustered in a clearly confined region of the chart. This region also includes the nonsweet compound **16**. This, however, does not happen if factor 3 also is taken into account (Figure 4).

**Superimpositions with Known Receptor Models.** The tridimensional structure of the most relevant compounds of this class was compared with the existing models of the sweet receptor. We referred particularly

**Table 1. Physicochemical Parameters for Principal Component Analysis of Compounds 1−17[a]**

| | | | | | C1 | | C2 | | |
|---|---|---|---|---|---|---|---|---|---|
| compd no.[a] | $\pi$ | $\sigma$ | $L$ | $B1$ | dipole moment | vdW volume | dipole moment | vdW volume | molecular connectivity |
| **S1** | −0.02 | −0.27 | 3.98 | 1.35 | 6.78 | 188.89 | 5.42 | 189.24 | 1.77 |
| **T2** | 0.00 | 0.00 | 2.06 | 1.00 | 6.13 | 166.97 | 4.28 | 167.75 | 1.61 |
| **S3** | 0.56 | −0.17 | 3.00 | 1.52 | 6.08 | 181.85 | 4.08 | 182.17 | 1.73 |
| **S4** | 0.38 | −0.24 | 4.92 | 1.35 | 6.92 | 203.66 | 5.45 | 203.07 | 1.89 |
| **T5** | 1.05 | −0.25 | 6.05 | 1.35 | 6.76 | 218.21 | 5.37 | 218.15 | 1.95 |
| **T6** | 2.08 | −0.03 | 4.51 | 1.35 | 6.93 | 235.95 | 5.57 | 235.22 | 2.32 |
| **S7** | −0.67 | −0.37 | 2.74 | 1.35 | 5.27 | 174.04 | 3.94 | 174.23 | 1.64 |
| **T8** | 0.61 | 0.00 | 4.30 | 1.70 | 6.60 | 201.34 | 5.41 | 200.96 | 2.17 |
| **S9** | 0.18 | −0.83 | 3.53 | 1.50 | 6.71 | 207.90 | 7.16 | 207.98 | 1.93 |
| **B10** | −0.28 | 0.78 | 3.44 | 1.70 | 5.51 | 184.52 | 6.98 | 184.25 | 1.76 |
| **T11** | −0.32 | 0.45 | 3.91 | 1.60 | 5.91 | 187.16 | 6.48 | 187.06 | 1.79 |
| **S12** | 0.71 | 0.23 | 3.52 | 1.80 | 5.44 | 183.14 | 4.83 | 183.11 | 1.75 |
| **T13** | 1.96 | −0.01 | 6.28 | 1.70 | 6.28 | 228.50 | 6.00 | 228.91 | 2.47 |
| **S14** | −1.34 | −0.25 | 2.74 | 1.35 | 4.80 | 180.23 | 4.68 | 180.85 | 1.69 |
| **S15** | −0.69 | −0.15 | 3.98 | 1.35 | 6.49 | 194.89 | 5.75 | 195.40 | 1.85 |
| **T16** | −0.69 | −0.25 | 2.74 | 1.35 | 3.90 | 195.72 | 2.55 | 195.26 | 1.84 |
| **T17** | −0.04 | −0.15 | 3.98 | 1.35 | 5.46 | 211.57 | 4.16 | 211.44 | 2.00 |

[a] C1 and C2 indicate two different minimum energy conformations. s = sweet; t = tasteless; b = bitter.

**Table 2. Component Loadings for Conformation C1**

| | factor | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| vdW volume, C1 | 0.956 | −0.152 | 0.132 | 0.016 |
| molecular connectivity | 0.947 | −0.017 | 0.102 | 0.157 |
| *L* | 0.777 | 0.008 | 0.395 | 0.217 |
| $\pi$ | 0.764 | 0.144 | 0.441 | 0.102 |
| dipole moment, C1 | 0.295 | −0.086 | 0.938 | −0.045 |
| *B*1 | 0.185 | 0.242 | −0.019 | 0.950 |
| $\sigma$ | −0.047 | 0.967 | −0.057 | 0.224 |

**Table 3. Percent of Total Variance Explained for Conformation C1**

| | factor | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| % variance explained | 53.868 | 21.352 | 10.328 | 6.776 |

**Table 4. Component Loadings for Conformation C2**

| | factor | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| vdW volume, C2 | 0.931 | −0.191 | 0.152 | 0.005 |
| molecular connectivity | 0.927 | −0.049 | 0.073 | 0.186 |
| $\pi$ | 0.883 | 0.143 | 0.008 | 0.094 |
| *L* | 0.834 | −0.006 | 0.328 | 0.107 |
| dipole moment, C2 | 0.193 | 0.119 | 0.944 | 0.189 |
| *B*1 | 0.172 | 0.236 | 0.200 | 0.934 |
| $\sigma$ | −0.053 | 0.961 | 0.111 | 0.212 |

**Table 5. Percent of Total Variance Explained for Conformation C2**

| | factor | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| % variance explained | 51.945 | 22.614 | 10.076 | 6.987 |

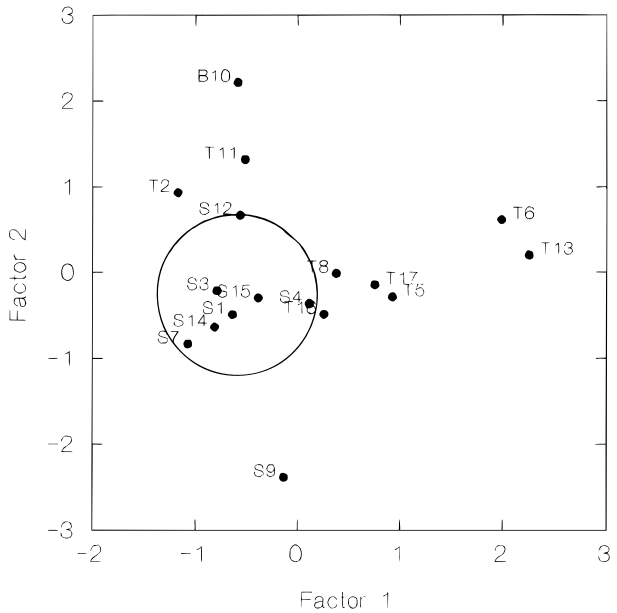



**Figure 3.** Score plot of factor 1 versus factor 2 for compounds **1**−**17**.

to the model by Tinti and Nofre (1991), which has the advantage of a precise description of the topological relationships between the glucophores and a broad effectiveness in explaining the sweet taste of different classes of compounds.

As a starting point two principal "anchor points" were identified, one localized in the *o*-ketocarboxylic acid fragment and the other in the *para* position of the other aromatic ring. All of the structures of the compounds, minimized by molecular mechanics, have a low-energy

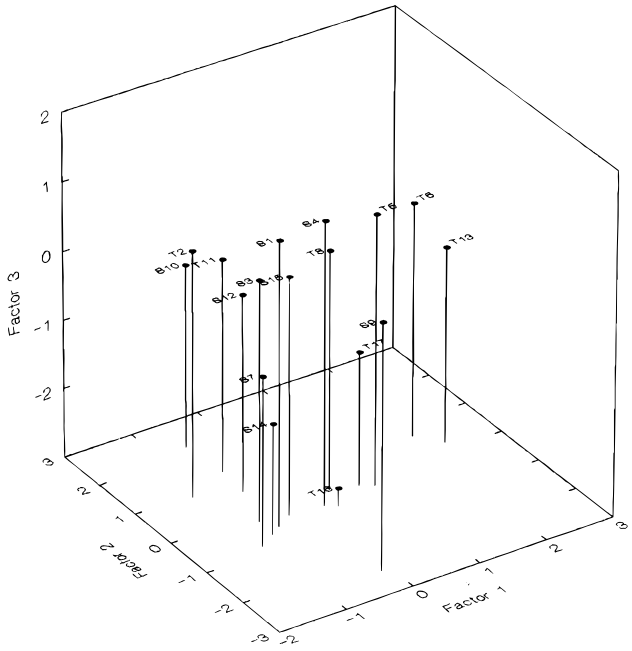



**Figure 4.** Score plot of factors 1, 2, and 3 for compounds **1**−**17**.

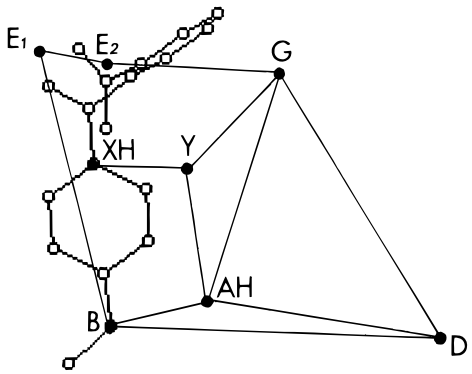



**Figure 5.** Superimposition of compound **1** with the Tinti−Nofre model.

preferred conformation in which the two aromatic rings are staggered by an angle of ca. 48° for C1 and −141° for C2; the loss of stability due to the poor superimposition of $\pi$ orbitals is compensated by minor steric interactions.

These structures were compared with the Tinti−Nofre model, matching the possible binding sites either as chemical functions or as relative distances from each other. The best result obtained for the lead compound **1** in its conformation C1 (which gave the best superimposition with the model) is shown in Figure 5. In this superimposition, the carboxylic and the ketonic groups should correspond to sites E2 and E1 of the model, while the methoxy substituent should correspond to site B.

In the proposed orientation the distances between the three glucophores are consistent with the geometrical requirements of the Tinti−Nofre model.

## DISCUSSION

In substituted 2-benzoylbenzoic acids the existence of a ring−chain tautomeric equilibrium between the open form **1** and the lactone **1a** (Figure 6) has been suggested (Runti and Galimberti, 1957). However, according to Bowden and Taylor (1971), this equilibrium was shown to be largely shifted toward the open form. To eliminate

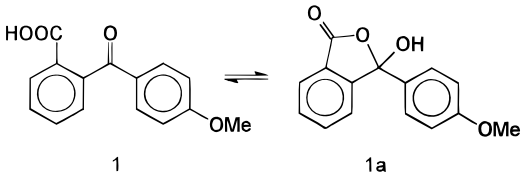



**Figure 6.** Compound **1**: equilibrium between the open (**1**) and closed forms (**1a**).

the possibility that the small percentage of the lactonic form could be responsible for the sweet taste, the lactonic model derivative **28** was synthesized, which was indeed tasteless.

As already observed in older literature, the strong cooperative effect of the carboxylic and ketonic function is confirmed. All of the derivatives **20**−**22**, which lack the C═O group, and the ester **25** and alcohol **26**, which lack the COOH group, are in fact tasteless. The only exception is compound **23**, in which the carbonylic function has been submitted to only a slight modification from C═O to S═O. The characteristics of these two glucophores seem to be consistent with those described in the Tinti−Nofre model for sites E1 and E2. In fact, these sites are suggested to be two H-bond acceptors as oxygen atoms in a CO or SO group, which usually act simultaneously and possibly in cooperation with the hydrophobic region G, this last being only partially represented in our compounds by the benzoic acid aromatic ring.

A hydrophobic part is required to maintain sweetness: in fact, compound **29**, in which the main part of this region has been eliminated, is tasteless. The preferred conformation of this compound is similar to that of **1** (Bowden and Henry, 1972).

The second effect to be explained is the role of the *para* substituent on the phenyl ring, which appears to be the glucophore B by comparison with the Tinti−Nofre model. Site B is another, generally stronger, H-bond acceptor. The methoxy substituent on an aromatic ring has been suggested to act as an efficient site B also in other classes of sweet compounds such as the isovanillyl derivatives (Van der Weel et al., 1987). In the course of those studies, it has been already noticed that there are some steric requirements around this group, namely, that the positive interaction is decreasing from methoxy to ethoxy and propoxy group and disappears with longer or larger substitutions, an effect that was also observed in the corresponding alkoxy derivatives of **1**. Another analogy with the class of isovanillyl sweet compounds (Arnoldi et al., 1993) is that the sweet taste disappears when $SCH_3$ is substituted for $OCH_3$ (compare **1** with **8**).

Consistently, an important contribution of steric factors is apparent from PCA analysis. This steric effect is presumably related to the dimensions of the cavity which accommodates glucophore B in the protein receptor and in particular with the length of the group.

It is possible to explain the failure of PCA analysis in predicting the sweetness of compound **9**. This compound has a high van der Waals volume value, and this probably explains why it is outside the region of sweet compounds. However, the compound is sweet, thus it must interact with the receptor. It seems reasonable that site B could accommodate a −NMe portion of the chain (which has an appropriate length) with the second methyl residue remaining outside. The electronic nature of this glucophore is also important:

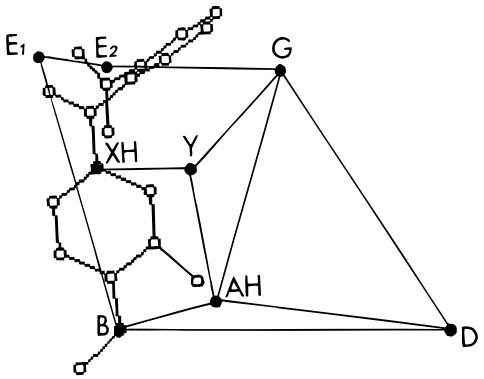



**Figure 7.** Superimposition of compound **15** with the Tinti−Nofre model.

compounds with substituents with negative $\sigma$ values such as $-OCH_3$, $-CH_3$, or $-NMe_2$ are sweeter than compounds with substituents with a zero or positive $\sigma$ value such as −H, SMe, $-NO_2$, or −COOH.

To confirm the hypothesis that the *para* substituent on the phenyl group corresponds to glucophore B, a fourth binding site, namely, site AH, was introduced in compound **1**. This was accomplished by preparing compound **15**, which has an isovanillic substitution pattern on phenyl ring, i.e., a hydroxy substituent in **3**, that can act as site AH of the Tinti−Nofre model (Figure 7). This obviously cannot hold for its isomer **16** in which the hydroxy and methoxy substituents are exchanged. As expected, compound **15** is sweet and its relative sweetness is greater than that of **1**, whereas **16** is tasteless. Thus the increase in sweetness of derivative **15** can be explained by the introduction of the AH glucophore in the appropriate position.

Moreover as seen before the PCA analysis explains, in some respect, the differences existing between the two isomers. In fact, they have essentially the same parameter values for the hydrophobic constant, vdW volume, and molecular connectivity; **16** has a $\sigma$ value lower than **15**, but this value is comparable to that of other sweet compounds such as **4** and **12**. The only noticeable difference is in the dipole moment of **16**, which is much lower than any other in the series.

With a different substituent such as methoxyl (compound **17**) in the 3′ position, corresponding to the site AH, the derivative is tasteless; the same substituent in the 2′ position (compound **18**) confers bitterness (**7**). Substituting the 3′-OH, 4′-OMe system with two adjacent hydroxyl groups, we still have a sweet compound (**14**) but with a decreased intensity.

In conclusion, the results of this work can be summarized as follows:

A very specific interaction with the sweet taste receptor is localized on the carboxyl and ketonic functions, which could likely be sites E1 and E2 according to the Tinti−Nofre model;

A third binding site is present, which presumably corresponds to site B and has some specific requirements concerning shape, volume, and electronic properties;

This class of sweet compounds shows some analogies with that of the isovanillyl sweeteners.

## LITERATURE CITED

Arnoldi, A.; Bassoli, A.; Merlini, L.; Ragg, E. Isovanillyl sweeteners. Synthesis and sweet taste of sulfur heterocycles. *J. Chem. Soc., Perkin Trans. 1*, **1993**, 1359−1366.

Bishoff, C. A.; von Hedenstrom, A. Ueber Bernsteinsäure Phenyl und Benzyl Ester. *Chem. Ber.* **1902**, *35*, 4073−4099.

Blicke, F. F.; Weinkauff, O. J. The preparation of several substituted anthrones. *J. Org. Chem.* **1932**, *54*, 1460−1464.

Bowden, K.; Taylor, G. R. Reactions of carbonyl compounds in basic solution. Part II. The mechanism of the alkaline hydrolysis of methyl 2-benzoylbenzoates. *J. Chem. Soc. B* **1971**, 145−148.

Bowden, K.; Henry, M. P. Ring−chain tautomerism. Part III. Substituted *cis*-3-benzoylacrylic acids. *J. Chem. Soc., Perkin Trans 2*, **1972**, 201−206.

Cohn, G. Zur Kenntnis organischer Geschmacksstoffe. *Pharm. Zentralhalle* **1914**, *55*, 735−762.

Gabriel, S.; Michael, A. Über die Einwirkung von wasserentziehenden Mitteln auf Säureanhydride. *Chem. Ber.* **1877**, *10*, 1551−1562.

Hansch, C.; Leo, A. *Substituent constants for correlation analysis in chemistry and biology*; John Wiley & Sons: New York, 1979.

Hrdlicka, J. Umelá sladidla a jejich vlastnosti. *Vyz. Lidu* **1973**, *28*, 9−11.

Lehaman, A. J. *Assoc. Food Drug Officials U.S. Q. Bull.* **1951**, *15*, 82−89; *Chem. Abstr.* **1952**, *46*, 2702i.

Jaszberenyi, C. A novel synthesis of precocenes. Efficient synthesis and regioselective O-alkylation of dihydroxy 2,2-dimethyl-4-chromanones. *J. Heterocycl. Chem.* **1988**, *25*, 871−877.

Kier, L. B.; Hall, L. H. *Molecular connectivity in chemistry and drug design*; Academic Press: New York, 1976.

Meyers, A. I.; Mihelich, E. D. Oxazolines. XVII. Regioselective metalation of 2-aryl oxazolines. A route to polydeuteriobenzoic acids. *J. Org. Chem.* **1975**, *40*, 3158−3159.

Möhler, K. Die Verwendung der *p*-Methoxy-*o*-benzoylbenzoesäure als Sü*β*stoff: Nachweis und Bestimmung. *Z. Lebensm. Unters.* **1950**, *90*, 431−434.

Oi, R.; Shimakawa, C.; Takenaka, S. Ullmann ether synthesis in DMI. Preparation of *m*-phenoxybenzyl alcohol. *Chem. Lett.* **1988**, *5*, 899−900.

Roberts, K. C.; Smiles, S. Methoxy derivatives of thioxanthone. *J. Chem. Soc.* **1929**, 863−872.

Runti, C.; Galimberti, S. Rapporti fra costituzione chimica e sapore dolce: Derivati dell' acido *p*-metossibenzoilbenzoico. *Ann. Chim. (Roma)* **1957**, *44*, 250−259.

Runti, C.; Collino, F. Rapporti fra costituzione chimica e sapore dolce: Analoghi dell' acido *p*-metossibenzoilbenzoico e derivati malondiamidici. *Ann. Chim. (Roma)* **1964**, *54*, 431−441.

Sabljić, A. GRAPH III, version 1.03; Institute Ruder Bosković, Zagreb, 1989.

Thomas, K.; Savelsberg, W. *Istanbul Univ. Tip Fak. Mecm.* **1949**, *12*, 353−358; *Chem. Abstr.* **1950**, *44*, 728f.

Tinti, J. M.; Nofre, C. Why does a sweetener taste sweet? In *Sweeteners: Discovery, molecular design, and chemioreception*; Walter, D. E., Orthoefer F. T., Dubois, G., Eds.; American Chemical Society: Washington DC, 1991, pp 206−213.

Van der Weel, H.; Van der Heijden, A.; Peer, H. G. Sweeteners. *Food Rev. Int.* **1987**, *3* (3), 193−268.

Verloop, A.; Hoogenstraaten, W.; Tipker, J. Development and application of new steric substituent parameters in drug design. In *Drug design*; Ariens, E. J., Ed.; Academic Press: New York, 1976; Vol. 7, pp 165−207.

Received for review October 4, 1996. Accepted March 11, 1997.[⊗] Financial support by EC Contract AIR3-CT94-2107 is gratefully acknowledged.

JF960755I

[⊗] Abstract published in *Advance ACS Abstracts,* May 15, 1997.