

Molecular structure analysis of xanthine alkaloids using terahertz spectroscopy

Ningyi Wang, Xinghao Huang, Jiamin Zhang, Xu Wu^{*}, Yan Peng, Yiming Zhu

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

ARTICLE INFO

Keywords:

Terahertz spectroscopy
Xanthine alkaloid
Density functional theory
Quantitative detection

ABSTRACT

Caffeine, theophylline and theobromine are representative xanthine alkaloids, commonly used as stimulants due to their effects on the central nervous system. Despite their similar molecular structures, they have different pharmacological effects, necessitating a rapid and accurate identification method. In this study, terahertz time-domain spectroscopy (THz-TDS) was used to measure the absorption spectra of these three xanthine alkaloids within the range of 2.0–17.0 THz. The characteristic absorption peaks were visualized and analyzed based on the quantum chemical calculations using Hartree-Fock (HF), Møller–Plesset perturbation theory (MP2) and density functional theory (DFT). Caffeine exhibited unique absorption peaks at 4.24, 5.00, and 11.13 THz. Theophylline showed characteristic peaks at 9.25, 12.20, and 15.09 THz. While theobromine exhibited characteristic peaks at 4.45, 7.68, and 11.21 THz. The results demonstrate that combining THz-TDS with DFT calculation can non-destructively, efficiently, and accurately identify these xanthine alkaloids, and providing valuable information for further understanding their pharmacological functions.

1. Introduction

Xanthine alkaloids are a class of naturally occurring nitrogen-containing organic compounds known for their central nervous system stimulatory activity [1,2]. They are widely used in clinical and health-care settings. The most commonly used xanthine alkaloids include caffeine, theophylline, and theobromine. These compounds share similar molecular structures, as they all contain purine rings. However, they differ in the position and number of their specific substituents. Caffeine ($C_8H_{10}N_4O_2$) is a trimethyl xanthine with methyl groups located at the 1, 3, and 7 positions. Theophylline ($C_8H_{10}N_4O_2$) is a dimethylxanthine with two methyl groups at positions 1 and 3. Theobromine ($C_7H_8N_4O_2$) is also a dimethylxanthine, but its methyl groups are at positions 3 and 7. These structural differences result in varying pharmacological effects. Caffeine has strong central nervous system stimulating activity. It is commonly used to treat headaches and increase heart rate, with side effects including insomnia, anxiety, and palpitations [3–5]. Theophylline acts as a bronchodilator by inhibiting of phosphodiesterase, leading to an increase in intracellular cyclic adenosine monophosphate levels. It is often used to treat asthma and chronic obstructive pulmonary disease, with side effects such as gastrointestinal discomfort and insomnia [6–8]. Theobromine has mild central nervous

system stimulatory activity and cardiac excitatory effects. It has fewer side effects, including mild gastrointestinal distress and tachycardia [9–11]. Despite their structural similarities, these compounds exhibit significantly different pharmacological effects as shown in Table 1. Therefore, the accurate identification of these xanthine alkaloids is of great scientific and clinical significance. This ensures the rational use of drugs, helps avoid misdiagnosis and adverse reactions, and improves therapeutic efficacy.

Common methods employed to identify different xanthine alkaloids include high-performance liquid chromatography (HPLC), mass spectrometry (MS), and Raman spectroscopy. HPLC method utilizes the distribution equilibrium between the stationary and mobile phases to sequentially separate and detect components at different retention times. In 2023, Gonzales-Yépez et al. [12] achieved the identification of caffeine, theophylline, and theobromine in beverages using HPLC coupled diode-array detection. Mass spectrometry identifies different xanthine alkaloids by measuring the mass-to-charge ratio of fragment ions. The major characteristic fragment ions for caffeine are 195, 138, and 110 m/z , for theophylline are 181, 163, and 138 m/z , and for theobromine are 181, 124, and 96 m/z [13,14]. Raman spectroscopy identifies the molecular structures by analyzing their characteristic fingerprint spectra. The major Raman characteristic peaks for caffeine

^{*} Corresponding author.

E-mail address: wuxu@usst.edu.cn (X. Wu).

<https://doi.org/10.1016/j.jms.2024.111936>

Received 6 June 2024; Received in revised form 11 July 2024; Accepted 6 August 2024

Available online 8 August 2024

0022-2852/© 2024 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

are 553, 1316, and 1605 cm^{-1} , for theophylline are 520, 440, and 1587 cm^{-1} , and for theobromine are 639, 1316, and 1355 cm^{-1} [15,16]. While HPLC, MS, and Raman spectroscopy have their advantages in identifying structurally similar biomolecules, they possess limitations including complexity, cost, and accuracy issues. Therefore, it is important to develop new, efficient and economical methods for identifying these structurally similar alkaloids.

Terahertz (THz) waves are located in the 0.1–20.0 THz band, positioned between the millimeter and infrared bands. They possess unique properties that distinguish them from other electromagnetic bands. THz waves are highly sensitive to the low-frequency motion patterns of biomolecules, such as crystal vibrations and weak interactions like hydrogen bonding and van der Waals forces. This sensitivity allows for the effective identification of biomolecules with similar structures, including chiral enantiomers and isomers [17]. Additionally, the photon energy of THz radiation is on the order of millielectron volts, which is low enough to avoid photoionizing or damaging biomolecules [18–20]. This enables nondestructive detection. Due to these attributes, THz spectroscopy has been employed in the identification of structurally similar drug molecules across a wide range of pharmaceuticals [21–24]. Building upon this foundation, the combination of THz spectroscopy and quantum chemical calculations emerges as a potent analytical methodology, further enhancing the understanding of the intricacies of drug molecules [24–28].

Although THz spectroscopy has demonstrated its potential in identifying structurally similar drug molecules, research on its application in detecting xanthine alkaloids is still in its infancy. In 2017, Weiling Fu et al. [29] used THz time-domain spectroscopy to obtain the absorption curves of caffeine, theophylline, and theobromine at 0.2–2.0 THz. Based on their characteristic absorption peaks in this band, they achieved preliminary identification of these three compounds. In 2021, Yang et al. [30] successfully identified the caffeine component in a compound acetaminophen drug using THz spectroscopy combined with a machine learning algorithm. In 2022, Phatham et al. [31] successfully identified the caffeine component of quinic acid/nicotinic acid/caffeine mixture by analyzing the characteristic absorption fingerprint spectrum of caffeine in the range of 0.3–3.0 THz. However, most current studies focus on narrower low-frequency bands, and the depth of THz response characterization for different types of xanthine alkaloids remains shallow.

This study investigates the THz characteristic peaks of three xanthine alkaloids across a broad frequency range through a combination of experimental testing and quantum chemical calculations. The analysis includes identifying molecular vibrational modes corresponding to characteristic absorption peaks and exploring the correlation between their structural properties and characteristic absorption frequencies.

The research aims to provide a new detection method utilizing THz spectroscopy for the rapid identification of structurally similar xanthine alkaloids. The anticipated outcomes are expected to advance the application of THz spectroscopy in pharmaceutical analysis and identification.

2. Materials and methods

2.1. Materials

Caffeine powder was provided by the Chongqing Institute for Food and Drug Control. Anhydrous theophylline powder (CAS: 487-21-8) and theobromine powder (CAS: 83-67-0) were purchased from Shanghai Aladdin Biochemical Technology Co. Ltd., both with a purity exceeding 99 %. Copolymers of cycloolefin (COC) powder, purchased from the Shanghai Institute of Nuclear Research, had an average particle size below 60 μm . These raw materials were not further purified before use.

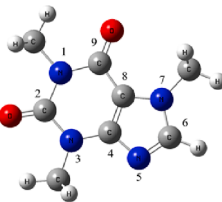
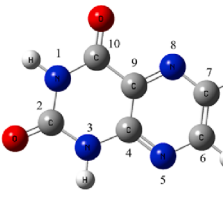
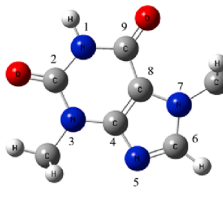
2.2. Sample preparation

The COC powder exhibits low absorption in the 2–20 THz, making it suitable as a diluent for preparing alkaloid samples. Initially, 60 mg of COC powder was finely ground and compressed under a pressure of 3 tons to form a tablet with smooth and parallel surfaces. This COC tablet served as the reference samples for subsequent THz measurements, providing reference signals. Subsequently, 5 mg of alkaloid powder (caffeine, anhydrous theophylline, or theobromine) was thoroughly mixed with 60 mg of COC powder. The mixtures were then ground and compressed under a 3 tons pressure for 3 min, to form tablets with an average diameter of 13 mm and a thickness in range of 500–580 μm . The sample loss during preparation was controlled to below 1 %.

2.3. Experimental apparatus

THz spectroscopy was conducted using a commercial Fourier Transform Infrared Spectrometer (FTIR, Bruker VERTEX 80v). The light source was a water-cooled mercury lamp, and the detector was a DTGS/polyethylene detector. The effective spectral region covered from 2.0 to 20.0 THz, with a Signal to Noise Ratio (SNR) exceeding 10000:1. All spectra were measured under vacuum at room temperature ($\sim 22^\circ\text{C}$) in order to minimize the effect of water vapor on the experiment results. The resolution was 4 cm^{-1} , the average number of scans was 32, and the scanning speed was 5 kHz.

Table 1
Molecular structure and medicinal properties of the three Xanthine alkaloids [3–11].

| Name | Caffeine | Anhydrous theophylline | Theobromine |
|----------------------|---|--|--|
| Molecular structure |  |  |  |
| Medicinal properties | Common medicinal properties | Central nervous system stimulation Caffeine > Theophylline > Theobromine [6] | |
| | Unique medicinal properties | Vasoconstrictor, used to promote smooth muscle relaxation and reduce headaches [3,6] | Bronchodilator, used in the treatment of asthma and chronic obstructive pulmonary diseases [6–8] |
| Side effects | | Insomnia, anxiety, and palpitations [3–5] | Gastrointestinal discomfort and insomnia [7,8] |
| | | | Vasodilator, used to lower blood pressure and improve blood circulation [10,11] |
| | | | Mild gastrointestinal discomfort and tachycardia [9–11] |

2.4. Theoretical calculation

THz wave exhibit high sensitivity to low-frequency molecular motions, such as backbone vibrations and weak interaction, enabling them to reflect the unique molecular structures. To understanding the characteristic absorption properties of caffeine, theobromine, and anhydrous theophylline in the THz region, quantum chemical calculations were employed in this study.

The initial structure models of the three xanthine alkaloids were obtained from the PubChem Structure database. The caffeine model has PubChem CID: 2519, anhydrous theophylline model has PubChem CID: 10250, and theobromine has PubChem CID: 5429. These molecular models were input into Gaussian 09 W software for molecular structure optimization, theoretical spectrum calculations, and vibrational mode analysis. Structure optimization and the frequency calculation are carried out under the same calculation method and basis set. Four computational methods were used: HF/6-31g(d), MP2/6-31g(d), DFT/B3LYP/6-31g(d), and DFT/B3LYP/6-311g(d,p). These methods used three ab initio frameworks, the Hartree-Fock (HF) method, Møller–Plesset perturbation theory (MP2) and density functional theory (DFT), with different basis set compositions. By using these methods, the theoretical vibrational spectra of the three xanthine alkaloids in the range of 2.0–17.0 THz were obtained. Subsequently, all the vibrational frequencies were Gaussian broadened according to a half-peak width of 4 cm^{-1} to obtain their theoretical spectra. The molecular vibrational modes corresponding to each vibrational frequency were then resolved. In this paper, our theoretical simulation results are also focused on this specific band to facilitate a meaningful comparison with the experimental test outcomes.

3. Analysis and discussion

3.1. Experimental THz absorbance spectra for xanthine alkaloids

Fig. 1 shows the experimental absorption spectra of caffeine, anhydrous theophylline, and theobromine in the range of 2.0–17.0 THz. In the THz spectrum of caffeine, twelve absorption peaks were observed, including strong peaks at 5.00, 6.59, 11.13, 11.71, 12.72, 13.32, 14.42 THz, weak peaks at 3.18, 7.22, 9.31, 16.50 THz, and a shoulder peak at 4.24 THz. Peaks exceeding 10 % of the maximum peak's height are classified as strong; others are weak. Three absorption peaks at 4.24, 5.00, and 11.13 THz were found only in caffeine but not in the other two xanthine alkaloids. Anhydrous theophylline showed fewer and weaker

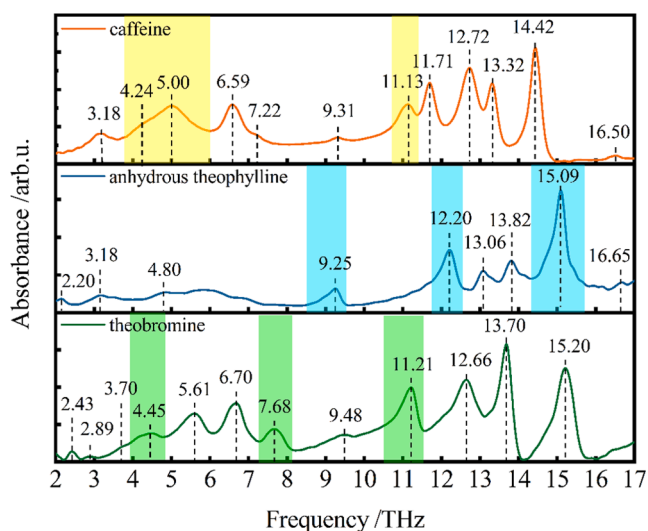


Fig. 1. THz absorption spectra of caffeine, anhydrous theophylline and theobromine.

absorption peaks in the 2.0–8.0 THz region, but exhibited three strong absorption peaks at 9.25, 12.20, and 15.09 THz. The experimental spectrum of theobromine exhibited twelve absorption peaks, including unique absorption peaks at 4.45, 7.68, and 11.21 THz. The experimental results demonstrated that the caffeine, anhydrous theophylline, and theobromine have both common and unique THz characteristic peaks. The unique characteristic peaks of each xanthine alkaloid are indicated in the colored section of Fig. 1.

3.2. Theoretical THz spectra for xanthine alkaloids

Fig. 2 shows the comparison between the experimental and theoretical spectra of caffeine at 2.0–17.0 THz. Four different computational methods were used to obtain theoretical spectra: HF/6-31g(d), MP2/6-31g(d), DFT/B3LYP/6-31g(d), and DFT/B3LYP/6-311g(d,p). The HF/6-31g(d) method calculated 17 vibrational frequencies for caffeine, resulting in 10 theoretical absorption peaks in 2.0–17.0 THz after Gaussian broadening. Similarly, the MP2/6-31g(d) method calculated 17 vibrational frequencies, showing 12 theoretical absorption peaks. The DFT/B3LYP/6-31g(d) method calculated 16 vibrational frequencies, showing 12 theoretical absorption peaks. The DFT/B3LYP/6-311g(d,p) method calculated 16 vibrational frequencies, resulting in 11 theoretical absorption peaks.

Table 2 presents the statistical deviations between the experimental and theoretical peaks of caffeine, obtained using different computational methods. The deviations of the HF/6-31g(d) method range from -0.32 to 0.55 THz, with a total absolute deviation of 2.73 THz. For the MP2/6-31g(d) method, the deviations range from -0.28 to 0.62 THz, and the sum of absolute deviation is 2.61 THz. For the DFT/B3LYP/6-31g(d) method, the deviations range from -0.10 to 0.50 THz, and the sum of absolute deviation is 2.35 THz. While the DFT/B3LYP/6-311g(d,p) method shows deviations ranging from -0.21 to 0.49 THz, with a total absolute deviation of 2.81 THz in comparison. Among these computational methods, the theoretical spectrum obtained using the

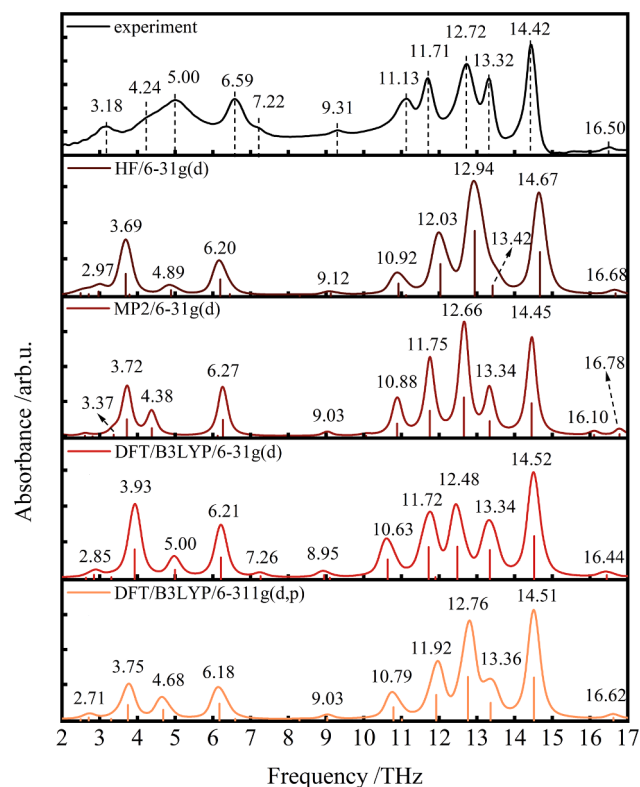


Fig. 2. Comparison of experimental and theoretical spectra of caffeine in 2.0–17.0 THz.

Table 2
Comparison of caffeine experimental and theoretical peaks under different computational methods.

| Experimental peak (THz) | HF/6-31g(d) | | MP2/6-31g(d) | | DFT/B3LYP/6-31g(d) | | DFT/B3LYP/6-311g(d,p) | |
|-------------------------|------------------------|-----------------|------------------------|-----------------|------------------------|-----------------|------------------------|-----------------|
| | Theoretical peak (THz) | Deviation (THz) | Theoretical peak (THz) | Deviation (THz) | Theoretical peak (THz) | Deviation (THz) | Theoretical peak (THz) | Deviation (THz) |
| 3.18 | 2.97 | 0.21 | 3.37 | -0.19 | 2.85 | 0.33 | 2.71 | 0.47 |
| 4.24 | 3.69 | 0.55 | 3.72 | 0.52 | 3.93 | 0.31 | 3.75 | 0.49 |
| 5.00 | 4.89 | 0.11 | 4.38 | 0.62 | 5.00 | 0.00 | 4.68 | 0.32 |
| 6.59 | 6.20 | 0.39 | 6.27 | 0.32 | 6.21 | 0.38 | 6.18 | 0.41 |
| 7.22 | — | — | — | — | 7.26 | -0.04 | — | — |
| 9.31 | 9.12 | 0.19 | 9.03 | 0.28 | 8.95 | 0.36 | 9.03 | 0.28 |
| 11.13 | 10.92 | 0.21 | 10.88 | 0.24 | 10.63 | 0.50 | 10.79 | 0.34 |
| 11.71 | 12.03 | -0.32 | 11.75 | -0.05 | 11.72 | -0.01 | 11.92 | -0.21 |
| 12.72 | 12.94 | -0.22 | 12.66 | 0.06 | 12.48 | 0.24 | 12.76 | -0.04 |
| 13.32 | 13.42 | -0.10 | 13.34 | -0.01 | 13.34 | -0.02 | 13.36 | -0.04 |
| 14.42 | 14.67 | -0.25 | 14.45 | -0.03 | 14.52 | -0.10 | 14.51 | -0.09 |
| 16.50 | 16.68 | -0.18 | 16.78 | -0.28 | 16.44 | 0.06 | 16.62 | -0.12 |

DFT/B3LYP/6-31g(d) method closely matches the experimental spectrum for caffeine. Therefore, the vibrational mode analysis of caffeine was conducted based on this method. The experimental peaks at 3.18, 4.24, 5.00, 6.59, 7.22, 9.31, 11.13, 11.71, 12.72, 13.32, 14.42, and 16.50 THz correspond to the theoretical peaks at 2.85, 3.93, 5.00, 6.21, 7.26, 8.95, 10.63, 11.72, 12.48, 13.34, 14.52, and 16.44 THz respectively (DFT/B3LYP/6-31 g(d)).

The comparison between the experimental and theoretical spectra of anhydrous theophylline, is presented in Fig. 3. The HF/6-31 g(d) method calculated 9 vibrational frequencies for anhydrous theophylline, resulting in 6 theoretical absorption peaks in 2.0–17.0 THz after Gaussian broadening. The MP2/6-31 g(d) method calculated 10 vibrational frequencies for anhydrous theophylline, resulting in 8 theoretical absorption peaks. The DFT/B3LYP/6-31 g(d) method calculated 10 vibrational frequencies, showing 7 theoretical absorption peaks. Similarly, the DFT/B3LYP/6-311 g(d,p) method calculated 10 vibrational frequencies, resulting in 7 theoretical absorption peaks.

The statistical deviations between the experimental and theoretical peaks of anhydrous theophylline, obtained using different computational methods, were presented in Table 3. The deviations of the HF/6-31g(d) method range from -1.16 to 0.58 THz, with a total absolute deviation of 3.67 THz. For the MP2/6-31g(d) method, the deviations range from -0.33 to 0.71 THz, and the sum of total absolute deviation is 2.13 THz. For the DFT/B3LYP/6-31g(d) method, the deviations range from -0.92 to 0.79 THz, and the sum of total absolute deviation is 3.64 THz. While the DFT/B3LYP/6-311g(d,p) method shows deviations ranging from -0.98 to 0.68 THz, with a total absolute deviation of 3.72 THz in comparison. Among these computational methods, the theoretical spectrum obtained using the MP2/6-31g(d) method closely matches the experimental spectrum for anhydrous theophylline. Therefore, the experimental peaks at 2.20, 3.18, 4.80, 9.25, 12.20, 13.06, 13.82, and 15.09 THz correspond to the theoretical peaks at 2.53, 3.49, 4.36, 9.24, 11.49, 13.08, 14.07 and 15.02 THz respectively based on MP2/6-31g(d) method.

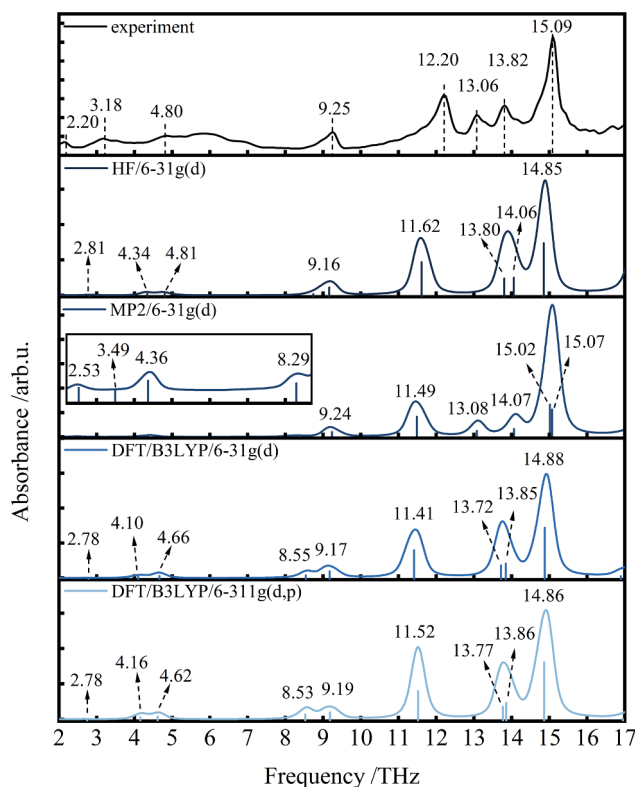


Fig. 3. Comparison of experimental and theoretical spectra of anhydrous theophylline in 2.0–17.0 THz.

Fig. 4 presents the experimental and theoretical spectra of theobromine. The HF/6-31g(d) method calculated 13 vibrational frequencies for anhydrous theophylline, resulting in 10 theoretical absorption peaks in 2.0–17.0 THz after Gaussian broadening. The MP2/6-31g(d) method calculated 13 vibrational frequencies for anhydrous theophylline, resulting in 9 theoretical absorption peaks. The DFT/B3LYP/6-31g(d) method calculated 12 vibrational frequencies, showing 9 theoretical absorption peaks. Similarly, the DFT/B3LYP/6-311g(d,p) method calculated 12 vibrational frequencies, resulting in 9 theoretical absorption peaks.

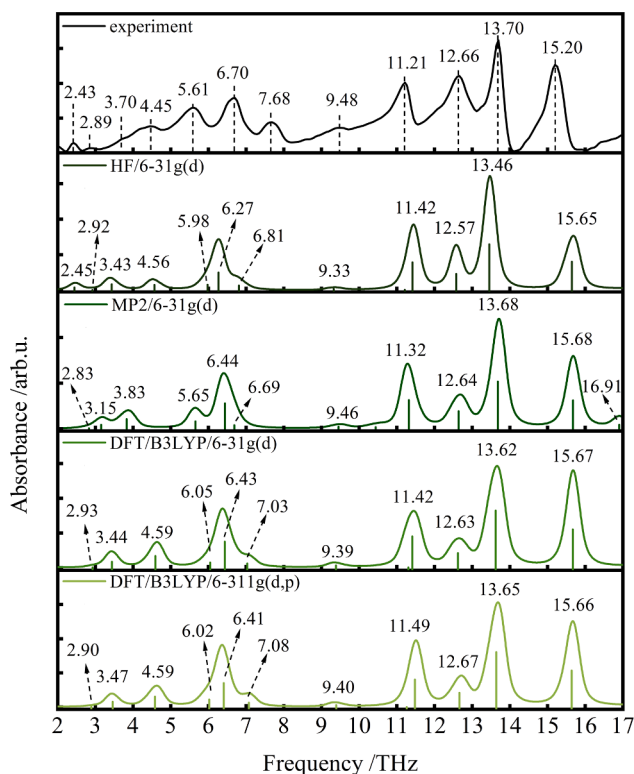
The statistical deviations between the experimental and theoretical peaks of theobromine, obtained using different computational methods, were presented in Table 4. The deviations of the HF/6-31g(d) method range from -0.45 to 0.87 THz, with a total absolute deviation of 3.22 THz. For the MP2/6-31g(d) method, the deviations range from -0.48 to 0.99 THz, and the sum of total absolute deviation is 3.17 THz. For the DFT/B3LYP/6-31g(d) method, the deviations range from -0.47 to 0.65 THz, and the sum of total absolute deviation is 2.77 THz. While the DFT/B3LYP/6-311g(d,p) method shows deviations ranging from -0.46 to 0.60 THz, with a total absolute deviation of 2.56 THz in comparison. Among these computational methods, the theoretical spectrum obtained using the DFT/B3LYP/6-311g(d,p) method closely matches the experimental spectrum for theobromine. The experimental peaks at 2.89, 3.70, 4.45, 5.61, 6.70, 7.68, 9.48, 11.21, 12.66, 13.70, and 15.20 THz correspond to the theoretical peaks at 2.90, 3.47, 4.59, 6.02, 6.41, 7.08, 9.40, 11.49, 12.67, 13.65, and 15.66 THz respectively based on DFT/B3LYP/6-311g(d,p) method.

Different computational methods exhibit different accuracies when applied to xanthine alkaloids. The HF/6-31g(d), MP2/6-31g(d) and DFT/B3LYP/6-31g(d) methods used the same basis set but differ in their theoretical frameworks. For the three xanthine alkaloids analyzed, the

Table 3

Comparison of anhydrous theophylline experimental and theoretical peaks under different computational methods.

| Experimental peak (THz) | HF/6-31g(d) | | MP2/6-31g(d) | | DFT/B3LYP/6-31g(d) | | DFT/B3LYP/6-311g(d,p) | |
|-------------------------|------------------------|-----------------|------------------------|-----------------|------------------------|-----------------|------------------------|-----------------|
| | Theoretical peak (THz) | Deviation (THz) | Theoretical peak (THz) | Deviation (THz) | Theoretical peak (THz) | Deviation (THz) | Theoretical peak (THz) | Deviation (THz) |
| 2.20 | 2.81 | -0.61 | 2.53 | -0.33 | 2.78 | -0.58 | 2.74 | -0.54 |
| 3.18 | 4.34 | -1.16 | 3.49 | -0.31 | 4.10 | -0.92 | 4.16 | -0.98 |
| 4.80 | 4.81 | -0.01 | 4.36 | 0.44 | 4.66 | 0.14 | 4.62 | 0.18 |
| 9.25 | 9.16 | 0.09 | 9.24 | 0.01 | 9.17 | 0.08 | 9.19 | 0.06 |
| 12.20 | 11.62 | 0.58 | 11.49 | 0.71 | 11.41 | 0.79 | 11.52 | 0.68 |
| 13.06 | 13.80 | -0.74 | 13.08 | -0.02 | 13.72 | -0.66 | 13.77 | -0.71 |
| 13.82 | 14.06 | -0.24 | 14.07 | -0.25 | 13.85 | -0.03 | 13.86 | -0.04 |
| 15.09 | 14.85 | 0.24 | 15.02 | 0.07 | 14.88 | 0.21 | 14.86 | 0.23 |

**Fig. 4.** Comparison of experimental and theoretical spectra of theophylline in 2.0–17.0 THz.

HF/6-31g(d) method took 5–11 min with frequency deviations ranging from -1.16 to 0.87 THz, the MP2/6-31g(d) method took 53 min to 9.5 h with frequency deviations ranging from -0.48 to 0.99 THz,

Table 4

Comparison of theophylline experimental and theoretical peaks under different computational methods.

| Experimental peak (THz) | HF/6-31g(d) | | MP2/6-31g(d) | | DFT/B3LYP/6-31g(d) | | DFT/B3LYP/6-311g(d,p) | |
|-------------------------|------------------------|-----------------|------------------------|-----------------|------------------------|-----------------|------------------------|-----------------|
| | Theoretical peak (THz) | Deviation (THz) | Theoretical peak (THz) | Deviation (THz) | Theoretical peak (THz) | Deviation (THz) | Theoretical peak (THz) | Deviation (THz) |
| 2.43 | 2.45 | -0.02 | — | — | — | — | — | — |
| 2.89 | 2.92 | -0.03 | 2.83 | 0.06 | 2.93 | -0.04 | 2.90 | -0.01 |
| 3.70 | 3.43 | 0.27 | 3.15 | 0.55 | 3.44 | 0.26 | 3.47 | 0.23 |
| 4.45 | 4.56 | -0.11 | 3.83 | 0.62 | 4.59 | -0.14 | 4.59 | -0.14 |
| 5.61 | 5.98 | -0.37 | 5.65 | -0.04 | 6.05 | -0.44 | 6.02 | -0.41 |
| 6.70 | 6.27 | 0.43 | 6.44 | 0.26 | 6.43 | 0.27 | 6.41 | 0.29 |
| 7.68 | 6.81 | 0.87 | 6.69 | 0.99 | 7.03 | 0.65 | 7.08 | 0.60 |
| 9.48 | 9.33 | 0.15 | 9.46 | 0.02 | 9.39 | 0.09 | 9.40 | 0.08 |
| 11.21 | 11.42 | -0.21 | 11.32 | -0.11 | 11.42 | -0.21 | 11.49 | -0.28 |
| 12.66 | 12.57 | 0.09 | 12.64 | 0.02 | 12.63 | 0.03 | 12.67 | -0.01 |
| 13.70 | 13.46 | 0.24 | 13.68 | 0.02 | 13.62 | 0.08 | 13.65 | 0.05 |
| 15.20 | 15.65 | -0.45 | 15.68 | -0.48 | 15.67 | -0.47 | 15.66 | -0.46 |

whereas the DFT/B3LYP/6-31g(d) method took 12–23 min with frequency deviations ranging from -0.92 to 0.79 THz. The HF method offers a foundational mean-field approximation for electronic correlations [32]. This is further refined by the MP2 method, which simultaneously incorporates electron correlation and dispersion energy [33]. In contrast, the DFT method excels in incorporating electronic information for xanthine alkaloids, offering a balance of calculation accuracy and efficiency.

The difference between DFT/B3LYP/6-31g(d) and DFT/B3LYP/6-311g(d,p) methods lies in their basis set composition, specifically the splitting group of valence orbitals and the inclusion of polarization functions. For the xanthine alkaloids analyzed, the DFT/B3LYP/6-31g(d) method took 12–23 min, while the DFT/B3LYP/6-311g(d,p) method took 24–45 min. The 6-31g(d) basis set provided higher accuracy for caffeine and anhydrous theophylline, whereas the 6-311g(d,p) basis set exhibited higher accuracy for theobromine. Although the 6-311g(d,p) basis set offers slightly better accuracy but at the cost of increased computation time. Balancing efficiency with accuracy, the 6-31g(d) basis set is recommended. Therefore, the DFT/B3LYP/6-31g(d) method is deemed more appropriate for theoretical studies of xanthine alkaloids.

Additionally, frequency deviations were observed between some experimental and theoretical peaks. For example, the theoretical peak at 4.10 THz for anhydrous theophylline was experimentally observed at 3.18 THz, and the theoretical peak at 7.03 THz for theobromine appeared at 7.68 THz experimentally. These frequency deviations likely result from the theoretical calculations considering single molecules at 0 K, while experimental measurements are conducted on polymolecule systems at room temperature. Such difference in conditions affects the accuracy of molecular motion calculations in THz wave region [34].

3.3. Common vibrational modes of xanthine alkaloids

Caffeine, anhydrous theophylline, and theobromine exhibit structural similarities and therefore possess similar vibrational modes. The

origins of these common characteristic THz peak were investigated in detail. Some THz peaks arise from the combination of multiple vibrational modes. The contribution of each vibrational mode was quantified by dividing its vibrational intensity by the sum of the intensities of all contributing vibrational modes, allowing for the identification of the most dominant vibrational mode. For the terahertz peak from a single vibration mode, its contribution is inherently 100 %.

These xanthine alkaloids exhibit a common characteristic THz peak in the range of 13.3–13.8 THz. The experimental peaks closely correspond to the theoretical peaks: 13.32 THz for caffeine (theoretical: 13.34 THz), 13.70 THz for theobromine (theoretical: 13.65 THz), and 13.82 THz for anhydrous theophylline (theoretical: 13.85 THz), as illustrated in Fig. 5. These peaks originate from the in-plane stretching vibration of the nitrogen-containing six-membered ring. Caffeine and theobromine exhibit greater similarity in vibrational modes compared to anhydrous theophylline, owing to the structural transition in anhydrous theophylline from a five-membered ring to a six-membered ring.

Both caffeine and theobromine exhibit common characteristic THz peaks around 6.6–6.7 THz and 12.6–12.7 THz. The experimental peaks closely correspond to the theoretical peaks: 6.59 THz for caffeine (theoretical: 6.21 THz), 6.70 THz for theobromine (theoretical: 6.41 THz), as illustrated in Fig. 6 (a–b). Additionally, 12.72 THz for caffeine (theoretical: 12.48 THz), 12.66 THz for theobromine (theoretical: 12.67 THz), as illustrated in Fig. 6 (c–d). The peak around 6.6–6.7 THz mainly arises from the in-plane rocking vibration of the 7-position methyl group, although the vibrational orientations differ between caffeine and theobromine. Similarly, the peak around 12.6–12.7 THz mainly originates from the in-plane rocking vibration involving both methyl groups (3- and 7-position) and hydroxyl group (2-position), although their vibrational orientations differ between caffeine and theobromine.

Both anhydrous theophylline and theobromine possess nitrogen-containing six-membered rings and neither has a methyl group at the 1-position. The theoretical peak at 4.66 THz for anhydrous theophylline corresponds to the experimental peak at 4.80 THz, as shown in Fig. 6(a), and the theoretical peak at 6.02 THz for theobromine corresponds to the experimental peak at 5.61 THz, as shown in Fig. 6(b). Both vibrational modes are related to the out-of-plane bending vibration of nitrogen-containing six-membered rings, though their vibrational orientations differ.

Both anhydrous theophylline and theobromine exhibit a common characteristic THz peak in the range of 4.8–5.6 THz. The experimental peak at 4.80 THz for anhydrous theophylline corresponds to the theoretical peak at 4.66 THz, and the experimental peak at 5.61 THz for theobromine corresponds to the theoretical peak at 6.02 THz, as shown in Fig. 7. These peaks are related to the out-of-plane bending vibration of nitrogen-containing six-membered rings, despite their opposite vibrational orientations.

3.4. Unique vibrational modes of xanthine alkaloids

Caffeine has a distinctive methyl group at 1-position, which triggers some unique vibrational modes in the 2.0–17.0 THz region. For example: the experimental peak at 4.24 THz (theoretical: 3.93 THz) mainly originates from the out-of-plane rocking vibration of the methyl groups, with a contribution degree of 98.17 %, as shown in Fig. 8(a). The experimental peak at 5.00 THz (theoretical: 5.00 THz) originates solely from in-plane rocking vibration of the methyl groups, as shown in Fig. 8 (b). The experimental peak at 11.13 THz (theoretical: 10.63 THz) originates solely from in-plane rocking vibration of the 1-position methyl group, as shown in Fig. 8(c).

Anhydrous theophylline has no methyl group, and the anisotropy of its five-membered ring into a six-membered ring, triggers some unique vibrational modes in the 2.0–17.0 THz region. For example: the experimental peak at 9.25 THz (theoretical: 9.17 THz) originates solely from the in-plane rocking vibration of the 10-position carbonyl group, as shown in Fig. 8(d). The experimental peak at 12.20 THz (theoretical: 11.41 THz) arises solely from the in-plane rocking vibration of the 2- and 10-position carbonyl groups, as shown in Fig. 8(e). The experimental peak at 15.09 THz (theoretical: 14.88 THz) originates solely from the in-plane stretching vibration of the nitrogen-containing six-membered ring, as shown in Fig. 8(f).

Theobromine possess two methyl groups at 3- and 7- positions, which trigger some unique vibrational modes in the 2.0–17.0 THz region. For example: the experimental peak at 4.45 THz (theoretical: 4.59 THz) originates solely from the out-of-plane rocking vibration of the 7-position methyl group, as shown in Fig. 8(g). The experimental peak at 7.68 THz (theoretical: 7.08 THz) arises solely from the out-of-plane rocking vibration of the methyl groups, as shown in Fig. 8(h). The experimental peak at 11.21 THz (theoretical: 11.49 THz) mainly originates from in-plane rocking vibration of the 7-position methyl group along with the 9-position carbonyl group, with a contribution of 99.92 %, as shown in Fig. 8(i).

The characteristic peaks of caffeine at 4.24, 5.00, and 11.13 THz, anhydrous theophylline at 9.25, 12.20, and 15.09 THz, and theobromine at 4.45, 7.68, and 11.21 THz correspond to low-frequency vibrational modes associated with distinct functional groups. By combining these unique THz peaks, a more precise identification of various xanthine alkaloids can be achieved.

4. Conclusion

This paper investigated the characteristic spectra and vibrational modes of three xanthine alkaloids (2.0–17.0 THz) using THz spectroscopy and theoretical calculations. Theoretical spectra closely match measured THz spectra, validating the accuracy of quantum chemical calculations. Considering both calculation accuracy and efficiency, the DFT/B3LYP/6-31g(d) method is more suitable for theoretical calculations of xanthine alkaloids. Characteristic absorption frequencies were

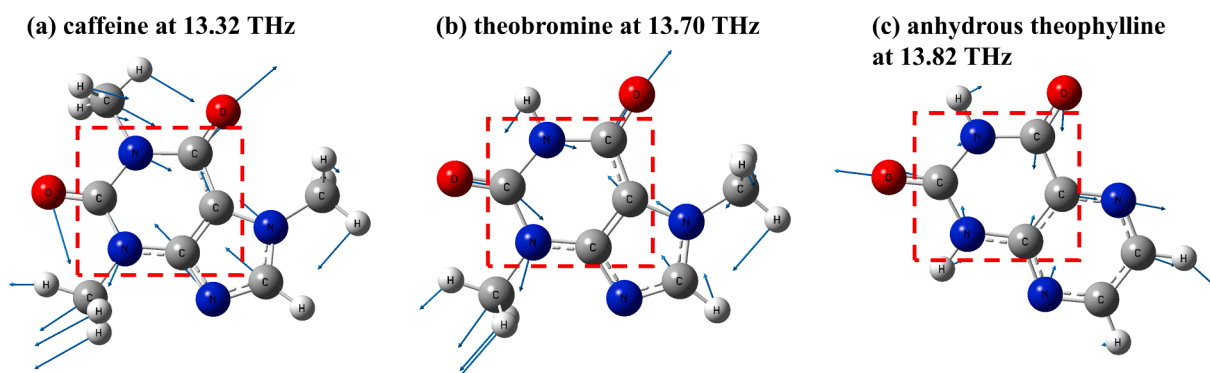


Fig. 5. The common vibration modes of three xanthine alkaloids.

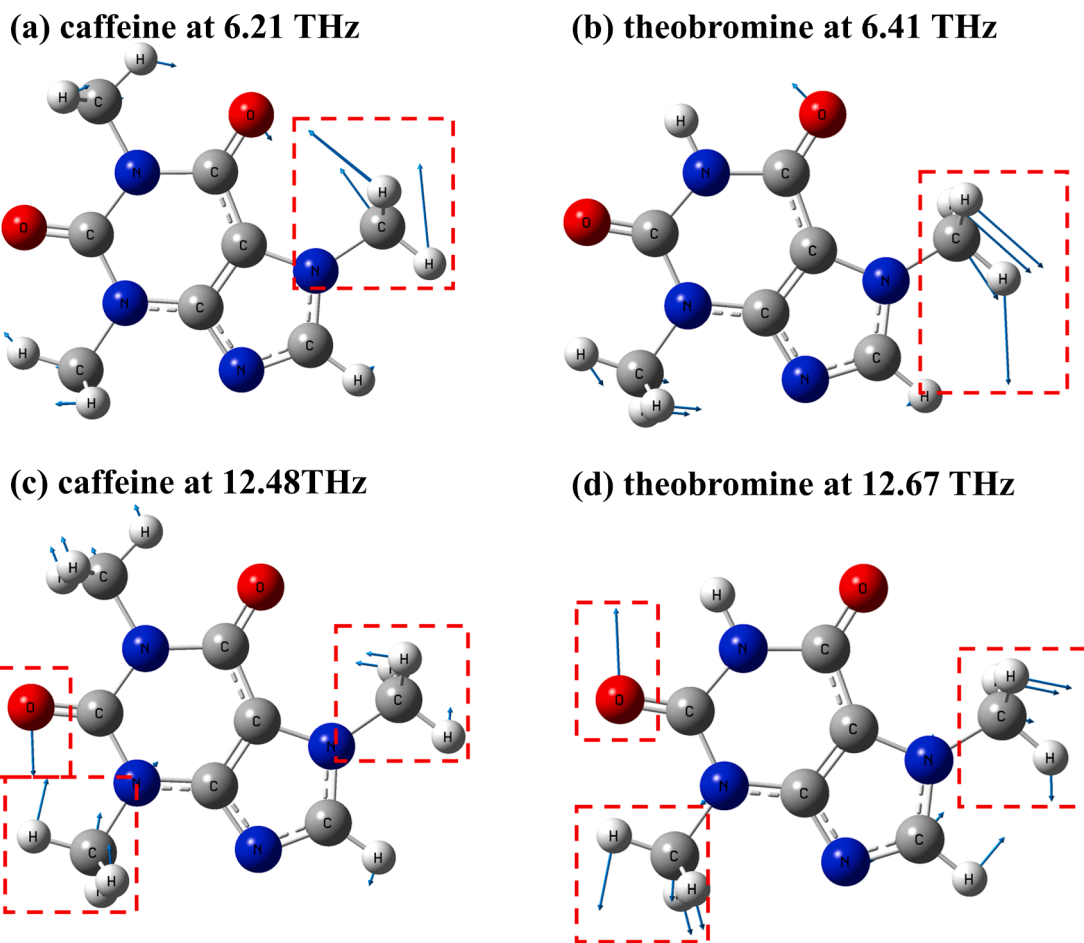


Fig. 6. The common vibration mode of caffeine and theobromine.

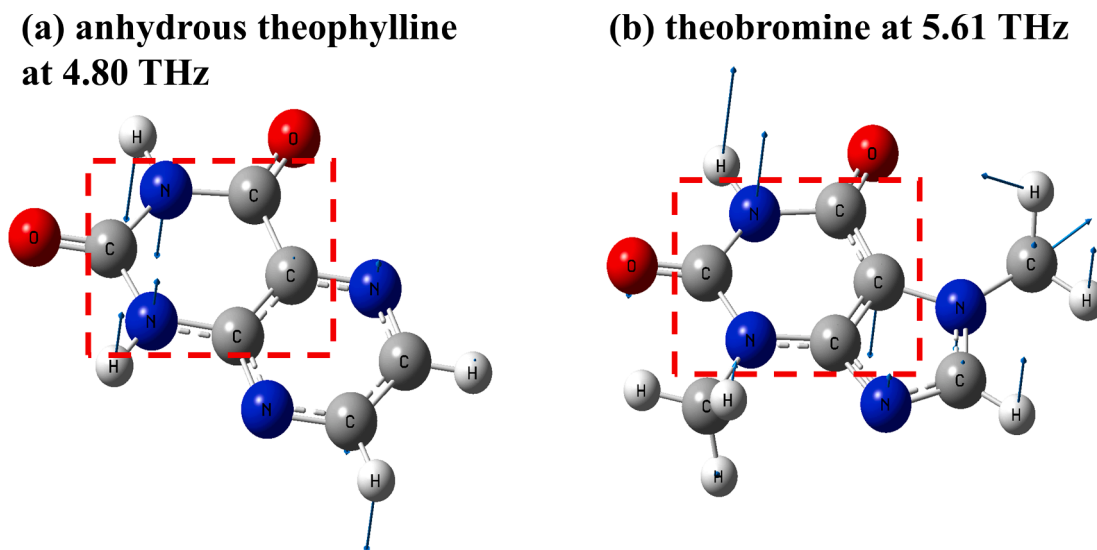


Fig. 7. The common vibration mode of anhydrous theophylline and theobromine.

identified, with common vibrational modes from the same functional groups producing similar absorption peaks. A common THz peak for the three xanthine alkaloids was observed at 13.3–13.8 THz. Caffeine and theobromine exhibited common THz peaks around 6.6–6.7 THz and 12.6–12.7 THz. Anhydrous theophylline and theobromine shared a common THz peak in the range of 4.8–5.6 THz. Unique vibrational

modes from distinct functional groups produced unique absorption peaks, such as those observed at 4.24, 5.00, and 11.13 THz for caffeine; 9.25, 12.20, and 15.09 THz for anhydrous theophylline; and 4.45, 7.68, and 11.21 THz for theobromine. These peaks correspond to low-frequency vibrational modes of their respective functional groups. This method addresses issues with existing detection techniques,

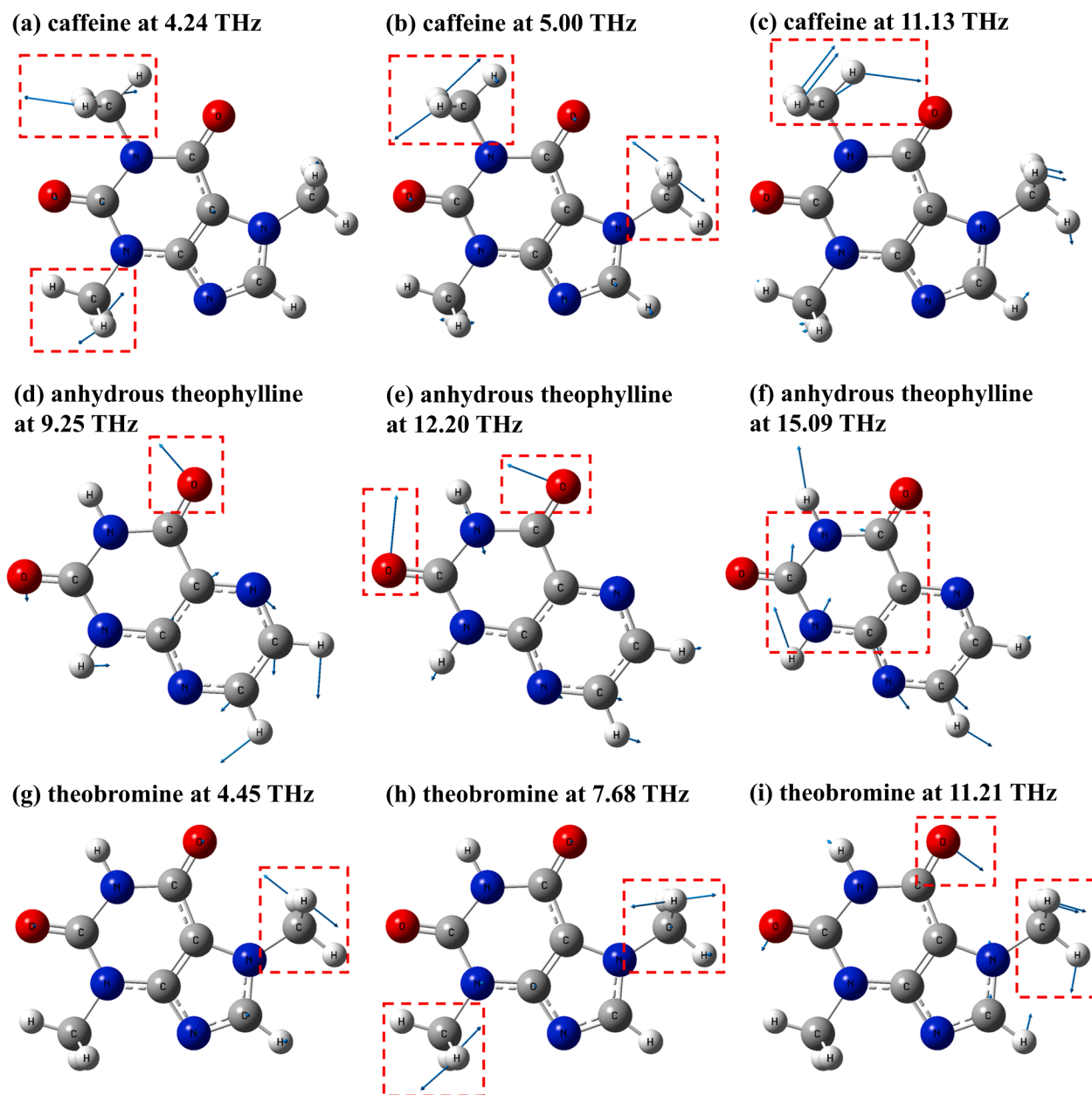


Fig. 8. The unique vibration modes of three xanthine alkaloids.

offering a new approach for THz spectroscopy in pharmaceuticals by solving problems related to sample pretreatment, high loss rate, long time consumption, and high cost.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (NSFC) [grant numbers 61988102, 61805140, 62335012];

National Key Research and Development Program of China [2022YFA1404004].

References

- [1] K. Yadav, D. Yadav, R. Yadav, Xanthine: synthetic strategy and biological activity, *Biointerface Res. Appl. Chem.* 12 (6) (2021) 7438–7468, <https://doi.org/10.33263/briac126.74387468>.
- [2] H. Ashihara, K. Mizuno, T. Yokota, A. Crozier, Xanthine Alkaloids: Occurrence, Biosynthesis, and Function in Plants, in: A.D. Kinghorn, H. Falk, S. Gibbons, J.i. Kobayashi (Eds.), *Progress in the Chemistry of Organic Natural Products* 105, Springer International Publishing, Cham, 2017, pp. 1–88.
- [3] J. dePaula, A. Farah, Caffeine consumption through coffee: content in the beverage, metabolism, health benefits and risks, *Beverages* 5 (2) (2019) 37, <https://doi.org/10.3390/beverages5020037>.
- [4] R.P. Barcelos, F.D. Lima, N.R. Carvalho, G. Bresciani, L.F.J.N.r. Royes, Caffeine effects on systemic metabolism, oxidative-inflammatory pathways, and exercise performance, *Nutrition Research* 80 (2020) 1–17. Doi: 10.1016/j.nutres.2020.05.005.
- [5] A. Wozniowiczka, M. Krychowiak-Maśnicka, G. Gołuński, A. Felberg, A. Borowik, D. Wyrzykowski, J. Piosik, Modulatory effects of caffeine and pentoxifylline on

- aromatic antibiotics: a role for hetero-complex formation, *Molecules* 26 (12) (2021) 3628, <https://doi.org/10.3390/molecules26123628>.
- [6] J. Monteiro, M.G. Alves, P.F. Oliveira, B.M. Silva, Pharmacological potential of methylxanthines: retrospective analysis and future expectations, *Crit. Rev. Food Sci. Nutr.* 59 (16) (2019) 2597–2625, <https://doi.org/10.1080/10408398.2018.1461607>.
- [7] D. Janitschke, A.A. Lauer, C.M. Bachmann, H.S. Grimm, T. Hartmann, M.O. W. Grimm, Methylxanthines and neurodegenerative diseases: an update, *Nutrients* 13 (3) (2021) 803, <https://doi.org/10.3390/nu13030803>.
- [8] J.P. Monteiro, M.G. Alves, P.F. Oliveira, B.M. Silva, Structure-bioactivity relationships of methylxanthines: trying to make sense of all the promises and the drawbacks, *Molecules* 21 (8) (2016) 974, <https://doi.org/10.3390/molecules21080974>.
- [9] M. Zhang, H. Zhang, L. Jia, Y. Zhang, R. Qin, S. Xu, Y. Mei, Health benefits and mechanisms of theobromine, *J. Funct. Foods* 115 (2024) 106126, <https://doi.org/10.1016/j.jff.2024.106126>.
- [10] F.C. Cadoná, R.F. Dantas, G.H.D. Mello, F.P. Silva-Jr, Natural products targeting into cancer hallmarks: An update on caffeine, theobromine, and (+)-catechin, *Food Sci. Nutr.* 62 (26) (2021) 7222–7241, <https://doi.org/10.1080/10408398.2021.1913091>.
- [11] N.S. Dutra, C.M. da Silva D'Avila, T.C. da Silva, T. de Oliveira Mendes, I. C. Livinalli, A.C.Z. Bertocelli, F.K. Saccol, F.C. Cadona, Biological properties of caffeine, (+)-catechin, and theobromine: an in silico study, *3 Biotech* 14 (4) (2024) 94, <https://doi.org/10.1007/s13205-024-03934-7>.
- [12] K.A. Gonzales-Yepe, J.L. Vilela, O. Reategui, Determination of caffeine, theobromine, and theophylline by HPLC-DAD in beverages commonly consumed in Lima, Peru, *Int. J. Food Sci.* 2023 (2023) 4323645, <https://doi.org/10.1155/2023/4323645>.
- [13] J. Zhai, Z. Gao, Y. Sai, C. Yang, L. Liu, J. Ssun, Simultaneous determination of caffeine, theophylline and theobromine in drinks by liquid chromatography – mass spectrometry, *Chinese J. Health Laborat. Technol.* 30 (02) (2020) 144–147.
- [14] A. Aqel, A. Almulla, A. Al-Rifai, S. Wabaidur, Z. Alothman, A. Badjah-Hadj-Ahmed, Rapid and sensitive determination of methylxanthines in commercial brands of tea using ultra-high-performance liquid chromatography-mass spectrometry, *Int. J. Anal. Chem.* 2019 (2019) 2926580, <https://doi.org/10.1155/2019/2926580>.
- [15] X. Cui, F. Lu, Identification of structural analogues xanthine, theophylline and theobromine by surface-enhanced Raman spectroscopy, *J. Pharm. Pract.* 38 (3) (2020) 6, <https://doi.org/10.12206/j.issn.1006-0111.202001005>.
- [16] M. Zareef, M. Mehedi Hassan, M. Arslan, W. Ahmad, S. Ali, Q. Ouyang, H. Li, X. Wu, Q. Chen, Rapid prediction of caffeine in tea based on surface-enhanced Raman spectroscopy coupled multivariate calibration, *Microchemical J.* 159 (2020) 105431, <https://doi.org/10.1016/j.microc.2020.105431>.
- [17] O.A. Smolyanskaya, N.V. Chernomyrdin, A.A. Konovko, K.I. Zaytsev, V.V.J.P.i.Q.E. Tuchin, Terahertz biophotonics as a tool for studies of dielectric and spectral properties of biological tissues and liquids, *Prog. Quantum Electron.* 62 (2018) 1–77, <https://doi.org/10.1016/j.pquantelec.2018.10.001>.
- [18] X. Wu, Y. Dai, L. Wang, Y. Peng, L. Lu, Y. Zhu, Y. Shi, S.J.B.o.e. Zhuang, Diagnosis of methylglyoxal in blood by using far-infrared spectroscopy and o-phenylenediamine derivation, *Biomedical Optics Express* 11(2) (2020) 963-970. Doi: 10.1364/boe.381542.
- [19] H. Chen, J. Han, J. Liu, L. Gao, S. Ma, Identification of chiral lansoprazole drugs using THz fingerprint spectroscopy, *Chem. Pap.* 77 (2022) 887–893, <https://doi.org/10.1007/s11696-022-02529-x>.
- [20] S. Zhang, X. Chen, K. Liu, H. Li, Y. Xu, X. Jiang, Y. Xu, Q. Wang, T. Cao, Z. Tian, Nonvolatile reconfigurable terahertz wave modulator, *Photonix* 3 (1) (2022) 7, <https://doi.org/10.1186/s43074-022-00053-5>.
- [21] S. Hu, C. Sun, X. Wu, Y. Peng, Polarization-independent terahertz surface plasmon resonance biosensor for species identification of Panax and Paeonia, *Photonics* 10 (3) (2023) 250, <https://doi.org/10.3390/photonics10030250>.
- [22] Z. Wang, Y. Peng, C. Shi, L. Wang, X. Chen, W. Wu, X. Wu, Y. Zhu, J. Zhang, G. Cheng, S. Zhuang, Qualitative and quantitative recognition of chiral drugs based on terahertz spectroscopy, *Analyst* 146 (12) (2021) 3888–3898, <https://doi.org/10.1039/d1an00500f>.
- [23] Z. Zheng, S. Zhao, Y. Liu, F. Zeng, Terahertz spectroscopy and weak interaction analysis of cinnamic acid derivatives, *J. Infrared Millim. Waves* 43 (02) (2024) 207–214.
- [24] X. Wu, L. Wang, Y. Peng, F. Wu, J. Cao, X. Chen, W. Wu, H. Yang, M. Xing, Y. Zhu, Y. Shi, S. Zhuang, Quantitative analysis of direct oral anticoagulant rivaroxaban by terahertz spectroscopy, *Analyst* 145 (11) (2020) 3909–3915, <https://doi.org/10.1039/d0an00268b>.
- [25] T. Chen, X. Zhong, Z. Li, F. Hu, Analysis of intermolecular weak interactions and vibrational characteristics for vanillin and ortho-vanillin by terahertz spectroscopy and density functional theory, *IEEE Trans. Terahertz Sci. Technol.* 11 (3) (2021) 318–329, <https://doi.org/10.1109/TTHZ.2020.3039462>.
- [26] J.W. Bennett, M.E. Raglione, S.M. Oburn, L.R. MacGillivray, M.A. Arnold, S. E. Mason, DFT computed dielectric response and THz spectra of organic co-crystals and their constituent components, *Molecules* 24 (5) (2019) 959, <https://doi.org/10.3390/molecules24050959>.
- [27] H. Takeda, H. Yoshioka, H. Minamide, Y. Oki, C. Aadachi, 0.5–4.5THz band terahertz spectroscopy of thermally activated delayed fluorescence molecules, *Optics Communications* 476 (2020) 126339, <https://doi.org/10.1016/j.optcom.2020.126339>.
- [28] A. Squires, E. Constable, J. Horvat, D. Appadoo, R. Plathe, R.A. Lewis, K.C. Rule, Terahertz vibrational dynamics and DFT Calculations for the quantum spin chain linearite, PbCuSO₄(OH)₂, *Chem. A Eur. J.* 128 (10) (2024) 1767–1775, <https://doi.org/10.1021/acs.jpca.3c06926>.
- [29] F. Weiling, L. Yang, Terahertz technology and its biological applications, *People's Medical Publishing House, Beijing*, 2017.
- [30] Q. Yang, L. Wu, C. Shi, X. Wu, X. Chen, W. Wu, H. Yang, Z. Wang, L. Zeng, Y. Peng, Qualitative and quantitative analysis of caffeine in medicines by terahertz spectroscopy using machine learning method, *IEEE Access* 9 (2021) 140008–140021, <https://doi.org/10.1109/access.2021.3116980>.
- [31] P. Loahavilai, S. Datta, K. Prasertsuk, R. Jintamethasawat, P. Rattanawan, J.Y. Chia, C. Kingkan, C. Thanapirom, T.J.A.o. Limpanuparb, Chemometric Analysis of a Ternary Mixture of Caffeine, Quinic Acid, and Nicotinic Acid by Terahertz Spectroscopy, *ACS Omega* 7(40) (2022) 35783-35791. Doi: 10.1021/acso.2c03808.
- [32] T. Sun, Y. Wang, M. Li, D. Hu, Raman Spectroscopic Study of Five Typical Plasticizers Based on DFT and HF Theoretical Calculation, 12(15) (2023) 2888. Doi: 10.3390/foods12152888.
- [33] M.K. Bilonda, L. Mammino, Intramolecular Hydrogen Bonds in Conformers of Quinine and Quinidine: An HF, MP2 and DFT Study, 22(2) (2017) 245. Doi: <https://doi.org/10.3390/molecules22020245>.
- [34] M. Wan, J. Fang, J. Xue, J. Liu, J. Qin, Z. Hong, J. Li, Y. Du, Pharmaceutical cocrystals of ethenzamide: molecular structure analysis based on vibrational spectra and DFT calculations, *Int. J. Mol. Sci.* 23 (15) (2022) 8850, <https://doi.org/10.3390/ijms23158550>.