Quantification analysis of progesterone based on terahertz spectroscopy

Yongni Shao, Yutian Wang, Zhi Zhu, Jialing Liu, Xin Xiong, Yiming Zhu, Ruoqing Zhai, Ming Gu, Zhengan Tian*, Yan Peng*

Abstract—**Progesterone is a natural progestational hormone secreted by the corpus luteum and placenta which protects the growth of the fetus in the early stage. It is also applied in the study on the treatment of human central nervous system diseases, proved to have high clinical medicinal value. Present detection methods of progesterone concentration, such as High Performance Liquid Chromatography (HPLC), Enzyme-Linked Immunosorbent Assay (ELISA), have disadvantages of high consumption, cumbersome procedures and expensive equipment. Here, the terahertz spectroscopy is applied to detect the medicinal concentration of progesterone. Firstly, the characteristic peaks of progesterone are 1.24 THz, 1.64 THz and 2.12 THz, determined by density functional theory calculation and experimental collection of terahertz spectra. Then the peak height and peak area of the above characteristic peaks are used to establish a linear correlation model with the concentration of progesterone injection, and the coefficient of determination (R^2), in this model is greater than 0.95. Meanwhile, we adopt Raman spectroscopy technology, and got the R^2 of 0.71, which is not so good as terahertz spectroscopy. Furthermore, Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) is used to analyze the composition of progesterone in progesterone capsules. The result shows that the error rate between the predicted concentration of progesterone and the real one is less than 5%. Therefore, terahertz spectroscopy provides a new way for the detection of progesterone drug content.**

Index Terms: **Progesterone, Terahertz spectroscopy, Density functional theory (DFT), Multivariate curve resolution-alternating least squares (MCR-ALS), Raman technology**

I. INTRODUCTION

Progesterone is an endogenous steroid hormone secreted by the corpus luteum, which has significant morphological effects on the endometrium stimulated by estrogen, and thus can protect the endometrium [1]. It is generally believed that progesterone is mainly related to the establishment and maintenance during mammalian pregnancy, releasing uterine tension and inhibit the excitability, turning the endometrial proliferation caused by progesterone into secretory phase, providing favorable conditions for the implantation of pregnant eggs and maintaining pregnancy. Meanwhile, progesterone promotes the development of breast and inhibits the ovulation, which can support and ensure the early growth of fetus. Within the nervous system, the neuroprotective and promyelinating effects of progesterone are promising, not only for preventing but also for reversing age-dependent changes and dysfunctions. There is indeed strong evidence that the aging nervous system remains at least to some extent sensitive to these beneficial effects of progesterone. [2-4]. Therefore, the research of progesterone is still valuable.

Most of the existing progesterone detection methods are chemical methods, such as Electrochemiluminescence (ECL), High Performance Liquid Chromatography (HPLC) and Enzyme-Linked Immunosorbent Assay (ELISA). ECL method is a form of chemiluminescence in which the light emitting chemiluminescent reaction is preceded by an electrochemical reaction. [5]. HPLC is based on the separation of material (rely on the differences in the adsorption characteristics, surface charge, ligand specificity, and molecular size of protein molecules) [6-7]. ELISA is the specific combination of antigen and antibody [8-9]. However, these methods are complicated to prepare, require the configuration of chemical reagents, and are time-consuming and costly to assay.

In addition to the above-mentioned chemical detection methods of progesterone, spectroscopy technology is increasingly used in biomedical research due to its rapid and non-destructive detection, including infrared spectroscopy and Raman spectroscopy and so forth. The infrared spectrum triggers transition between different energy levels mainly through the molecular vibration, stretching, bending and other state changes. Since different substances have different absorptive intensity and wavelength, their absorption peaks and intensity reflected in infrared spectroscopy vary from each other [10-12]. Hai et al. used qualitative research method to study a variety of estrogen and progesterone by applying the near-infrared spectroscopy [13]. Wang et al. used infrared spectroscopy to verify that there is a good correlation between

Yongni Shao, Yiming Zhu, Yan Peng are with the Terahertz Technology Innovation Research Institute, Terahertz Spectrum and Imaging Technology Cooperative Innovation Center, Shanghai Key Lab of Modern Optical System, University of Shanghai for Science and Technology, Shanghai 20009, China, and also with the Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai 200092, China. (email: ynshao@zju.edu.cn; ymzhu@usst.edu.cn; py@usst.edu.cn)

Yutian Wang, Zhi Zhu, Jialing Liu, Xin Xiong are with the Terahertz Technology Innovation Research Institute, Terahertz Spectrum and Imaging Technology Cooperative Innovation Center, Shanghai Key Lab of Modern

Optical System, University of Shanghai for Science and Technology, Shanghai 20009, China. (email: 193720569@st.usst.edu.cn; zhuzhi@usst.edu.cn; 183740584@st.usst.edu.cn; 183740611@st.usst.edu.cn)

Zhengan Tian is with the Shanghai International Travel Healthcare Center, Shanghai Customs District P.R. 200335, China. (email: tianzhenganciq@163.com)

Ruoqing Zhai, Ming Gu are with the School of Foreign Languages, Tongji University, Shanghai 200092, China. (email: 1933078@tongji.edu.cn; 2033159@tongji.edu.cn)

(Corresponding author: Yan Peng; Zhengan Tian)

solvent-induced infrared band shifts (SIFS) of the two C=O in the progesterone molecule and the solvent AN value [14]. In addition, Raman spectroscopy is an effective method for studying molecular vibrations, since it can reflect the differences in the chemical composition and molecular structure of samples at the molecular level. The detection principle of Raman spectroscopy is different from that of infrared spectroscopy, because it has strong recognition for non-polar groups (such as C = C, C-C, etc.). Wang et al. used Raman spectroscopy to distinguish between type I and type II progesterone crystals [15]. These studies show that the spectral analysis of progesterone has certain feasibility and practical value.

With the development of modern ultrafast optics, terahertz spectroscopy becomes more and more mature. Terahertz (THz) waves are electromagnetic waves between far-infrared waves and microwaves, with a frequency range from 0.1 to 10 THz, and have characteristics of both infrared waves and microwaves. Compared with other spectrum techniques, terahertz spectroscopy has more advantages in the vibrational and rotational levels of macromolecule, because many biological samples show molecular vibrations and obvious terahertz characteristic peaks within the terahertz range, such as hormones and amino acids [16-17]. Terahertz spectrum is very sensitive to polar groups. Contrary to Raman spectrum, which mainly reflects intramolecular vibration, the spectral information in terahertz region is abundant in weak intermolecular interaction and skeleton vibration of macromolecular, which is directly related to molecular structure. Therefore, terahertz spectroscopy is increasingly applied to identification and analysis of substance [18-21].

In recent years, many researchers have been using terahertz technology to study progesterone. For example, Son et al. adopted the terahertz time-domain spectroscopy system to distinguish the two polymorphic forms of progesterone, using the metastable existence of progesterone as a sign, and the apparent absorption pattern was used to recognize the mesenchymal transition of progesterone [22]. It shows that terahertz technology can be used to study the preparation of progesterone crystals. Smirnova et al. adopted THz-TDS and Raman spectroscopy to study the influence of changes in molecular bonds in molecular crystals on the vibrational spectrum, and modelled in DMol3 code, thus revealing the correspondence between molecular vibration and frequency [23]. It shows that theoretical calculations can be considered as a prediction of actual test results, and it can also analyze the mode of low-frequency vibration as well as the reasons that actually affect the vibration frequency from the differences between theory and practice. The results of the research also provide us a theoretical analyzing basis. Lee et al. combined Nanoslot-based THz sensor and THz-TDS technology to provide a new detection method for the qualitatively distinguish between trace progesterone and 17α-OH progesterone [24]. However, how to apply terahertz spectroscopy to the quantitative detection of progesterone concentration still remains unknown.

Experiment in this study proposes a rapid and non-destructive detection method of progesterone drug concentration based on terahertz spectroscopy technology. First of all, it used the B3LYP hybrid function and 6-31G basis set in Gaussion09 software to simulate the vibrational absorption spectra of progesterone molecules by density functional theory (DFT), and obtained the identification of the characteristic peak positions of progesterone in terahertz. Then the experiment set the progesterone in seven different concentrations within the range of the injection concentration, and collected the terahertz spectra of different concentrations of progesterone using THz-TDs. Hence, the correlation model between the peak height or area of the characteristic peak of progesterone and the concentration of progesterone solution is established. In order to compare the predicted accuracy of the model, we also used Raman spectroscopy to detect the same batch of samples, and established a correlation model based on the intensity of the Raman characteristic peak and the concentration of the progesterone solution. Finally, the predicted accuracy of the two methods were compared. In addition, this study used the same proportion of progesterone capsules on the market (the proportion of progesterone and excipients was consistent), and the spectral curve of progesterone and the corresponding concentration percentage of progesterone were analyzed by multivariate curve resolution-alternating least squares (MCR-ALS), and the concentration of progesterone was calculated and compared with the real concentration.

II. MATERIALS AND METHODS

A. Chemical Reagent

Progesterone is insoluble in water. Its molecular formula is $C_{21}H_{30}O_2$, and the molecular weight is 314. The progesterone in this study is purchased from Aladdin in solid powder form, the purity of which is more than 98%, and its CAS number is 57-83-0. The polyethylene is purchased from the Sigma in solid powder form with a particle diameter of 40-48um, and the CAS number is 9002-88-4.

B. Density functional theory

The THz absorption peak is derived from the resonance absorption between the irradiated THz wave and the vibration/rotation of atoms/functional groups in the molecule [25]. For the analysis of molecular vibration/rotation mode, density functional theory is currently used more frequently. It is a quantum mechanical method for studying the electronic structure of multi-electron systems. It is widely used in the fields of physics and chemistry, and is particularly suitable for molecular and condensed analysis. The study of the properties of matter is one of the most commonly used methods in the fields of condensed matter physical computational materials science and computational chemistry. The Gaussion-09 package is currently the most popular and widely used comprehensive quantum chemistry calculation program in the field of computational chemistry. It is developed on the basis of quantum mechanics and it is committed to applying the theory of quantum mechanics to practical problems. It can verify and predict almost all the properties of the target system by some basic commands. The vibrational mode of

progesterone molecule can be used to calculate the theoretical terahertz spectrum.

C. Terahertz time-domain spectroscopy system

TAS7400SP (standard system) terahertz time-domain spectroscopy system produced by Advantest of Japan is used in the experiment. the single scanning time is 200 ms; and the dynamic range of peak frequency is more than 60 dB. We applied the transmission spectrum mode, and the scanning times is 256.

D. Sample preparation and spectrum collection

We set our target progesterone concentrations to cover the three specifications of progesterone in sterilized oil solution as per the 2020 edition of Pharmacopoeia of The People's Republic of China. According to the Pharmacopoeia, there are three specifications of progesterone sterilized oil solution, including 1ml: 5mg, 1ml: 10mg and 1ml: 20mg. Each 1ml ampoule bottle contains different specifications of progesterone, including 5mg, 10mg and 20mg. We first calculated the concentration of progesterone in the sterilized oil solution. The density of sterile soybean oil is about 0.916-0.922mg/ml according to the record, and the molecular weight of progesterone is 314.Then we can have the corresponding progesterone concentrations of the three solutions, including 0.00796mol/L,0.0318mol/L, and 0.0637mol/L. After that, we set up a concentration gradient based on the three concentrations, including 0.00796mol/l, 0.0159mol/l, 0.0251mol/l, 0.0318mol/l, 0.0477mol/l, 0.0637mol/l and 0.0796mol/l. These seven concentrations respectively correspond to the seven specifications of 2.5mg/ml, 5mg/ml, 7.5mg/ml, 10mg/ml, 15mg/ml, 20mg/ml, 25mg/ml.

Firstly, put the steel ball with the diameter of 2mm and the pure progesterone into the grinding test tube, and grind the progesterone with a grinder for 190s. We can obtain the formula for calculating the concentration of progesterone (mol/L) from Ueno's research:

$$C = \frac{w}{MV_{pellet}} \qquad (1)$$

$$V_{pellet} = \pi r^2 d \qquad (2)$$

In the formula, w represents the mass of progesterone powder; M represents relative molecular mass; d represents thickness of the tablet; r represents the radius of the tablet; and V_{pellet} is the volume of the tablet made by the mixture of progesterone and PE powder [19]. According to this formula, we controlled the total mass of the tablets at 150mg, and prepared the tablets with the mass ratio of pure progesterone to PE at 1:38, 1:19, 1:12, 1:9, 1:6, 1:4, and 1:3, which corresponds to the above seven concentrations of progesterone. Meanwhile, we made sure that the mass loss was controlled within 1% during the experimental operation. Then we put the mixed powder into the mold and pressed them for two minutes with a pressure of 3 tons to form a tablet of 1 mm thick. Three progesterone samples of each concentration are also prepared. During the collection process, we collected 3 different positions for each sample tablet, while each point was collected five times and then averaged. Since water vapor absorbs the terahertz signal, excessive water vapor will interfere with the signal reception. As a result, we

dried the inside of the experimental device before the experiment to ensure that the whole terahertz signal detection could in right condition, where the humidity was less than 3% [17].

Furthermore, this experiment has also compared the pros and cons of terahertz spectroscopy by using Raman technology. Laser Confocal Micro-Raman Spectroscopy was adopted, of which the optical efficiency is greater than 30%, the spectral resolution is greater than 0.4 wavenumber, the horizontal resolution is better than 1 micron, and the longitudinal resolution is better than 2 microns. We used a 532nm laser as the Raman laser light source as well as a 2.5% filter, the acq time of which is 5s, and RTD time of which 1s.

E. Data processing and model establishment

Influenced by PE, the reflection peaks appeared when the spectral signal was collected. We used Matlab to find their position in the time-domain spectroscopy of the signal and replaced them by zero-padding, retaining the original main signal and eliminating the spectral vibration caused by the reflection peak. The UnscramblerX 10.1 was used to perform the adjacent average smoothing of every 15 data points, and correct baseline to improve the signal-to-noise ratio.

Under certain wave number, the absorption of THz is proportional to the concentration of the molecule, generally following the Beer-Lambert law:

$$A_v = \varepsilon_v bc \qquad (3)$$

In this formula, A_v represents the absorbance at the wave number v, ε_v represents the molar absorption coefficient of the substance, b represents the optical path length of the sample, and c represents the concentration of the substance in the sample. The linear fitting was performed according to the spectral absorbance value and the concentration we obtained, and the model was evaluated by the coefficient of determination (R^2) and the Root Mean Squared Error ($RMSE$).

F. Multivariate curve resolution-alternating least squares (MCR-ALS)

Multivariate curve resolution-alternating least squares(MCR-ALS)can decompose the unknown mixed system in physical and chemical meaning, distinguish and determine the curves (such as the spectrum, the PH curve, the system curve, etc.) or relative concentration distribution corresponding to the main components, with no need for types and combination information of the unknown, which is an effective method in chemometrics [26-27].The two-dimensional spectral data matrix D $(m*n)$ can be expressed as:

$$D = CS^T + E \qquad (4)$$

In this formula, $C(m*p)$ and $S^T(p*n)$ are the concentration distribution matrix of the pure component and the spectrum matrix respectively, E is the measurement error matrix, the dimension is $m*n$—m represents the actual measured points of PH or degradation time, and n represents the points of wavelength, P is the number of components in the analysis system. The MCR-ALS is used to obtain the C and the S with physical meaning, and iteratively calculate the initial values of the approximate spectral matrix or the concentration distribution matrix to get a true concentration distribution curve and a spectral curve. In this experiment, we used

progesterone capsules on the market. The progesterone in a capsule is 50mg. The excipients are four different substances in a total of 110mg, and the mass ratio is 5 to 11. We prepared four samples in which the mass fraction of progesterone remained the same, the total mass of the excipients remained the same, and the ratio of the excipients was changed. Each one was sampled three points by a terahertz time domain spectrometer, each point for four times.

III.   RESULTS AND DISCUSSION

A.   Simulation of progesterone molecular

We visited the molecular structure of progesterone in the protein database (pdb ID: 2 aa5). Density functional theory (DFT) were used to calculate and study the absorption of progesterone by Gaussion-09 package with B3LYP hybrid function and 6-31G basis set in the terahertz frequency range. At the same time, the source of the absorption peak of the progesterone molecule was analyzed with the help of a visual window [28].
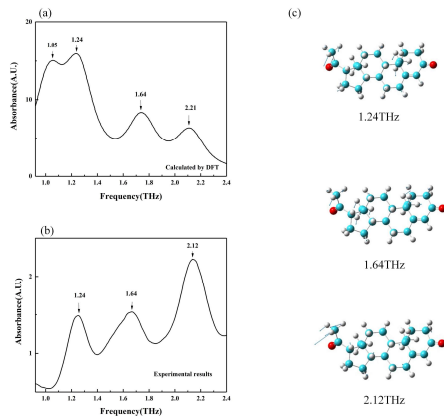

Fig.1 (a) The spectral curve of progesterone calculated by DFT;(b) The spectral curve of progesterone obtained through experiment;(c) The vibrational modes of progesterone at 1.24 THz,1.64 THz and 2.12 THz calculated by DFT.

Fig.1(a) shows the absorption spectrum of the progesterone molecule obtained by theoretical calculation, and Fig.1(b) shows the absorption spectrum of the standard progesterone obtained through experiment. It can be concluded from Fig.1(a) that there are four absorption peaks of the progesterone molecule in theoretical calculation and they are located at 1.05 THz, 1.24 THz, 1.64 THz and 2.21 THz respectively, of which the vibration positions at 1.05 THz and 1.24 THz are close. And the experiment shows that there are three absorption peaks of the progesterone and they are located at 1.24 THz, 1.64 THz, and 2.12 THz, of which the characteristic peak at 1.64 THz is the widest while its overall intensity is the weakest, and the characteristic peak at 2.12 THz has the highest absorption intensity. Comparing the theoretical and experimental calculations, we found that the 1.24 THz and 1.05 THz obtained by the theoretical calculation correspond to the absorption peak at 1.24 THz obtained by experiment; the 1.64 THz calculated by theory corresponds to the absorption peak at 1.64 THz obtained by experiment; the 2.21 THz calculated by theory corresponds to the absorption

peak at 2.12 THz obtained in the experiment. Through the visual window of Gaussian, it is found that the vibration mode at 1.24 THz is caused by the molecular collective torsion dominated by the acetyl functional group at the end of the progesterone molecule; the vibration modes at 1.64 and 2.21 THz are both caused by the overall vibration of the progesterone molecule, yet they belong to the molecular curling and in-plane rocking vibration respectively, as shown in Fig. 1(c). It is worth noting that the absorption peak at 1.05 THz in the theory is not reflected in the experiment. After the analysis, it is believed that due to the close frequency of 1.05 and 1.24 THz, the temperature in the laboratory causes the absorption peak to drift. This is similar to the research results on the drift of the absorption peak of progesterone molecules in different temperatures by Smirnova et al., 2012 [23]. What is calculated according to theory is the vibration mode formed by the force within a single molecule, while the experimental object is the crystal structure, in which the absorption peak obtained is influenced by not only the intramolecular force but also the intermolecular force. Therefore, absorption envelopes are formed at some characteristic peaks, which broaden the absorption frequency, causing two close absorption peaks to form one absorption peak. From the above analysis, it can be seen that the three characteristic absorption peaks of progesterone in the frequency band of 1.24,1.64 and 2.12 THz are more obvious, which can be used as the basis for qualitative detection of the progesterone.

B.   Modeling of progesterone concentration based on terahertz spectroscopy

In order to verify that the terahertz spectroscopy can be used to detect the concentration of progesterone, the absorption spectrum of the progesterone-PE mixed tablet was collected in the experiment. The spectrum contains seven concentration samples. As shown in Fig.2, we analyzed the 1-2.4 THz frequency band. As the concentration increases, the intensity of the absorption peak shows an increasing trend. The higher the concentration value is, the sharper the characteristic peak at 1.64 THz is, so does the situation at 1.24THz.
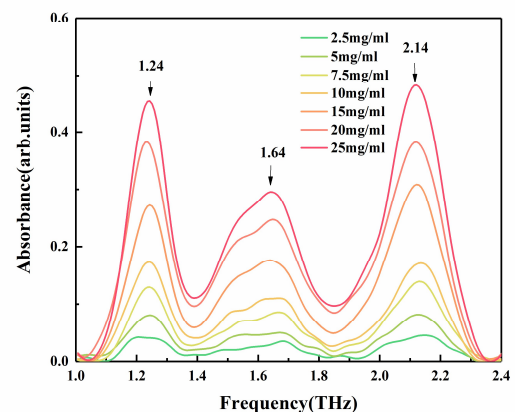

Fig.2 The terahertz spectral curve of seven concentrations of progesterone.

Furthermore, the relationship between the spectral data and the concentration of progesterone is modeled and analyzed by linear fitting. The intensity and peak area of the characteristic

absorption peaks were used to model the three characteristic peaks. The three peak heights of 1.24, 1.64 and 2.12 THz are taken as the absorption intensity, and the three integral areas of 1-1.4 THz,1.4-1.85 THz and 1.85-2.4 THz are taken as the peak areas. The results of the fitting model are shown in Fig.3 and Fig.4 respectively.
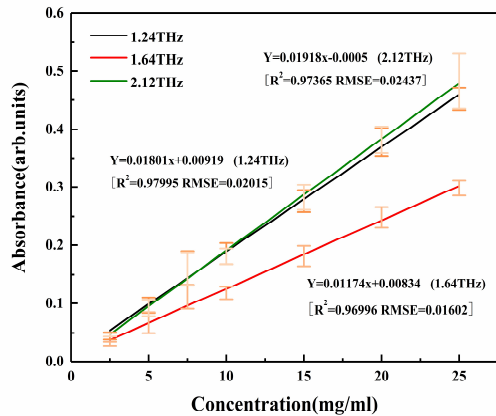

Fig.3 Linear fitting results of the absorption intensity of characteristic peak and the progesterone concentration at 1.24, 1.64, 2.12 THz.
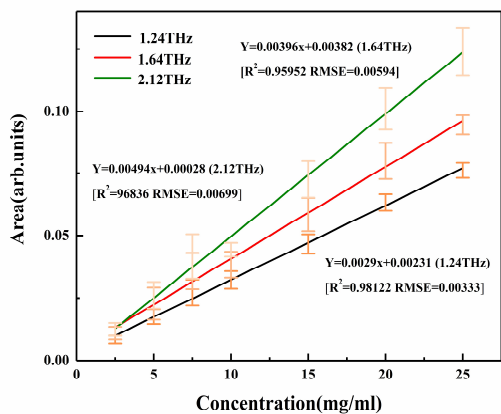

Fig.4 Linear fitting results of the characteristic peak area and the progesterone concentration at 1.24, 1.64 and 2.12 THz.

$$Y=0.01801x+0.00919 \quad \left[R^2 = 0.97995, RMSE = 0.02015 \right] \quad (5)$$

$$Y=0.01174x+0.00834 \quad \left[R^2 = 0.96966, RMSE = 0.01602 \right] \quad (6)$$

$$Y=0.01918x+0.0005 \quad \left[R^2 = 0.97365, RMSE = 0.02437 \right] \quad (7)$$

$$Y=0.0029x+0.00231 \quad \left[R^2 = 0.98122, RMSE = 0.00333 \right] \quad (8)$$

$$Y=0.00369x+0.00382 \quad \left[R^2 = 0.95952, RMSE = 0.00594 \right] \quad (9)$$

$$Y=0.00494x+0.00028 \quad \left[R^2 = 0.96836, RMSE = 0.00699 \right] \quad (10)$$

In the formula, x is the sample concentration, Y is the absorption intensity or area, R^2 is the coefficient of determination, and $RMSE$ is root mean squared error. Formulas (5), (6) and (7) correspond to the absorption intensity at 1.24, 1.64, and 2.12 THz; formulas (8), (9) and (10) correspond to the peak area at 1.24, 1.64, and 2.12 THz.

The modeling shows that both the absorption intensity and peak area can reach the result that R^2 is greater than 0.95,

indicating the feasibility of terahertz spectroscopy for the identification of progesterone concentration.

After using the terahertz spectroscopy to quantitatively analyze the concentration of progesterone, we repeated the above experiment by using Raman spectroscopy, trying to compare the quantitative accuracy for the concentration of progesterone between these two technologies. First, we collected the Raman spectrum of pure progesterone tablets. The Raman signal intensity of progesterone at $2881cm^{-1}$ is the most obvious in $2750-3000cm^{-1}$, so this characteristic peak was used for Raman data analysis. We then performed the adjacent average processing on 15 points in the collected Raman spectrum. As shown in Fig.5(a), the Raman spectrum curves show a growing trend of absorption intensity as the concentration increases. Next, we modeled the Raman absorption intensity and progesterone concentration, and compared with the terahertz model (8). The analysis result is shown in Fig.5(b). The coefficient of determination (R^2) of the model, is 0.71, which has a certain gap compared with the terahertz fitting model. Therefore, the terahertz technology has obvious advantages in the quantitative analysis of progesterone concentration.
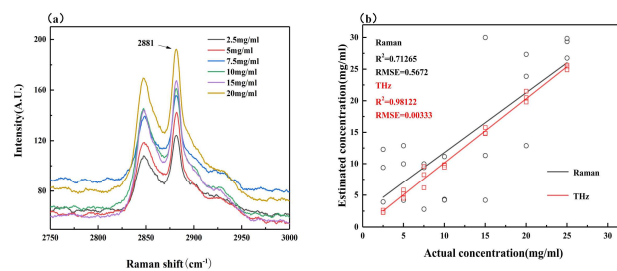

Fig.5 (a) Raman spectrum curves of different concentrations of progesterone;(b) Comparison of actual concentration and estimated concentration.

C. Study on the effect of different excipient ratio on concentration prediction of progesterone in capsules based on the MCR-ALS

The experiment used the MCR-ALS to detect and research on the progesterone content of drug samples with progesterone capsules of the same ratio on the market (The proportion of progesterone and excipients is the same and the proportion of progesterone by mass is 31.25%).We used the Unscrambler to smooth the collected terahertz spectrum curve with 15 points by the adjacent average method, and used the collected spectrum of pure progesterone as the limited initial condition for the MCR-ALS calculation. After 300 iterations, and comparing with the progesterone spectra respectively calculated by theory and detected by experiment, the composition of progesterone is analyzed. In Fig.6, component 1 is the spectrogram of progesterone obtained by analysis and calculation, and component 2 is the mixed spectral information of four different excipients. At the same time the MCR-ALS can analyze the concentration proportion of progesterone in each sample. Table 1 shows the mass percentage concentration of the samples with four different excipient ratios and the actual mass percentage concentration. It can be seen that the progesterone concentration obtained by analysis is very close to the actual concentration (31.25%),

and the error values are all less than 5%. Therefore, we believe that it is feasible to analyze and calculate the concentration of progesterone in progesterone capsules using MCR-ALS.
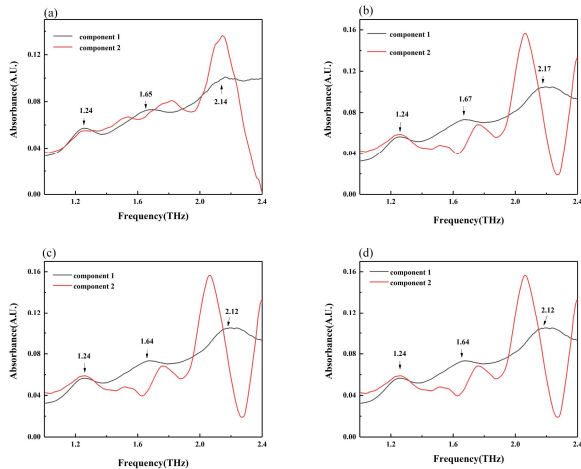

Fig.6 (a), (b), (c) and (d) are the component diagrams of the collected terahertz spectra with four different excipients ratios after the MCR analysis.

TABLE.1
THE PREDICTED CONCENTRATION AND ERROR RATE OF PROGESTERONE WITH FOUR DIFFERENT EXCIPIENTS RATIOS AFTER THE MCR-ALS ANALYSIS.

Sample	a	b	c	d
Actual value	31.25%	31.25%	31.25%	31.25%
Predicted value	32.69%	31.81%	32.36%	31.83%
Error value	4.61%	1.79%	3.55%	1.86%

IV. CONCLUSION

In this study, we calculated and verified the feasibility of the rapid and non-destructive detection of progesterone concentration based on the terahertz spectroscopy technology. First, based on the calculation of the density functional theory in Gaussian, we obtained the molecular calculation spectrum and vibration model of progesterone, and determined that fingerprint peaks of progesterone in the terahertz spectrum are 1.24, 1.64, and 2.12 THz. Then we collected their terahertz spectra and used the peak heights and peak areas of the above characteristic peaks to establish a linear correlation model with the concentration of progesterone. The result shows that R^2, coefficient of determination of the model, is greater than 0.95. At the same time, Raman spectroscopy was used for model comparison, and the result shows that the terahertz spectroscopy technology ($R^2 > 0.95$) is better than the Raman spectroscopy technology ($R^2 = 0.71$) in predicting the concentration of progesterone. In addition, the experiment collected the terahertz spectrum information of the drug samples with progesterone of the same ratio capsules (four samples with the same ratio of progesterone to excipients), and analyzed the progesterone content of the drugs by the multivariate curve resolution-alternating least squares (MCR-

ALS). The result shows that the error value between the predicted and actual concentration of progesterone is less than 5%. This experiment provides an idea for the use of terahertz spectroscopy to detect the content of progesterone drugs.

V. ACKNOWLEDGMENT

VI. REFERENCES

1. R. C. Tuckey, "Progesterone synthesis by the human placenta," *Placenta*, vol. 26, no.4, pp. 273-281, 2005.
2. D. Lengel, J. W. Huh, and J. R. Barson, "Progesterone treatment following traumatic brain injury in the 11-day-old rat attenuates cognitive deficits and neuronal hyperexcitability in adolescence," *Experimental Neurology*, vol. 330, pp. 113329, 2019.
3. S. L. Gonzalez, "Progesterone for the treatment of central nervous system disorders: the many signaling roads for a single molecule," *Neural Regeneration Research*, vol. 15, no.10, pp. 1846-1847, 2020.
4. M. Schumacher, R Guennoun, A. Ghoumari, C. Massaad, F. Robert, M. El-Etr, Y. Akwa, K. Rajkowski, E. Baulieu, "Novel Perspectives for Progesterone in Hormone Replacement Therapy, with Special Reference to the Nervous System" *Endocrine Reviews*, vol.28, no.4, pp.387-439,2007.
5. Ayad. A, Iguer. Ouada. M, Benbarek. H, "Electrochemiluminescence immunoassay for progesterone by using a heterologous system in plasma bovine" *Veterinary World*, vol.7, no.8, pp.610-613, 2014.
6. F. U. Erko, S. Ozsar, B. Gven, G. Kalkandelen, E. Uğrar, "High-performance liquid chromatographic analysis of steroid hormones" *Journal of Chromatographic Science*, vol.27, no.2, pp86-90, 1989.
7. F. Petr, D. Jing, W. James L, "Simultaneous identification of estrogen and progesterone receptors by HPLC using a double isotope assay" *Journal of Steroid Biochemistry and Molecular Biology*, vol.42, 1992.
8. J. V. Samsonova, A. P. Osipov, S. E. Kondakov, "Strip-dried whole milk sampling technique for progesterone detection in cows by ELISA" *Talanta*, vol.175, pp143-149, 2017.
9. L. Wu, C. Xu, C. Xia, Y. Duan, C. Xu, H. Zhang, J. Bao, "Development and application of an ELISA kit for the detection of milk progesterone in dairy cows" *Monoclonal Antibodies in Immunodiagnosis and Immunotherapy*, vol.33, no.5, pp330-333, 2014.
10. A. Kasal, M. Budesinsky, W. J. Griffiths, "Spectroscopic Methods of Steroid Analysis" *Steroid Analysis*, Springer, Dordrecht, 2010.
11. F. Omeis, A. F. S. Seica, R. Bernard, "Following the Chemical Immobilization of Membrane Proteins on Plasmonic Nanoantennas Using Infrared Spectroscopy," *ACS Sensors*, vol. 5, no. 7, pp. 2191-2197, 2020.
12. A. Nawaf, H. Mohammed, "Visible/Near Infrared (VIS/NIR) spectroscopy as an optical sensor for evaluating

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TTHZ.2021.3082286, IEEE Transactions on Terahertz Science and Technology

> REPLACE THIS LINE WITH YOUR PAPER IDENTIFICATION NUMBER (DOUBLE-CLICK HERE TO EDIT) <   7

olive oil quality," *Computers & Electronics in Agriculture*, vol. 173, 2020.

13. Feng. H, G. M. Xu, D. X. Liu, "Simultaneous Determination of Several Oestrogens and Progestogens by Near Infrared Spectroscopy," *Chinese Journal of Analytical Chemistry*, vol. 29, no. 2, pp. 175-177, 2001.

14. X. Y. Wang, "Solvent Effect in the Infrared Spectra of Progesterone," *Chinese Journal of Analytical Chemistry*, vol. 30, no. 10, pp. 1206-1209, 2002.

15. F. Wang, J. A. Wachter, "An Investigation of Solvent-Mediated Polymorphic Transformation of Progesterone Using in Situ Raman Spectroscopy," *Organic Process Research and Development*, vol. 4, no. 5, pp. 391-395, 2000.

16. S. Minah, P. Hyeong-Ryeol, "Terahertz Biochemical Molecule-Specific Sensors," *Advanced Optical Materials*, vol. 8, no.3, 2020.

17. X. M. Zheng, "Origin broadband terahertz sources and sensors," *Journal of Nanoelectronics and Optoelectronics*, vol. 2, no. 1, pp.58-76, 2007.

18. S. A. Yoon, S. H. Cha, S. W. Jun, S. J. Park, "Identifying different types of microorganisms with terahertz spectroscopy," *Biomedical optics express*, vol. 11, no. 1, pp. 406-416, 2020.

19. Y. Peng, C. J. Shi, Y. M. Zhu, "Terahertz spectroscopy in biomedical field: a review on signal-to-noise ratio improvement," *PhotoniX*, vol. 1, no. 1, 2020.

20. Y. Q. Cao, P. J. Huang, W. T. Ge, "Qualitative and quantitative detection of liver injury with terahertz time-domain spectroscopy," *Biomedical optics express*, vol. 11, no. 2, pp. 982-993, 2020.

21. Y. Xiang, Z. Xiang, Y. Ke, "Biomedical Applications of Terahertz Spectroscopy and Imaging," *Trends in Biotechnology*, vol. 34, no. 10, pp. 810-824, 2016.

22. Y. J. Son, D. K. Lee, J. H. Son, "Identification of interpolymorph transformations of progesterone by terahertz time-domain spectroscopy," *Current Applied Physics*, vol. 16, no. 1, pp. 45-50, 2016.

23. I. N. Smirnova, D. A. Sapozhnikov, A. V. Kargovsky, V. A. Volodin, O. P. Cherkasova, "Lowest-lying vibrational signatures in corticosteroids studied by terahertz time-domain and Raman spectroscopies," *Vibrational Spectroscopy*, vol. 62, pp. 238-247, 2012.

24. S. H. Lee, D. Lee, M. H. Choi, J. H. Son, M. Seo, "Highly Sensitive and Selective Detection of Steroid Hormones Using Terahertz Molecule-Specific Sensors," *Analytical Chemistry*, vol. 91, no. 10, pp. 6844-6849, 2019.

25. Y. Peng, X. R. Yuan, "Terahertz identification and quantification of neurotransmitter and neurotrophy mixture," *Biomedical Optics Express*, vol. 7, no. 11, pp. 4472-4479, 2016.

26. R. Tauler, K. Bruce, F. Sydney, "Multivariate Curve Resolution Applied to Spectral Data from Multiple Runs of an Industrial Process[J]," *Analytical Chemistry*, vol. 65, no. 15, pp. 2040-2047, 1993.

27. R. Tauler, S. Lacorte, D. Barcelo. "Application of multivariate self-modeling curve resolution to the quantitation of trace levels of organophosphorus pesticides in natural waters from interlaboratory studies[J]," *Journal of Chromatography A*, vol. 730, no. 1, pp. 177-183, 1996.

28. M. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. Petersson, Gaussian 09; Gaussian, Inc.: Wallington, CT, 2009.
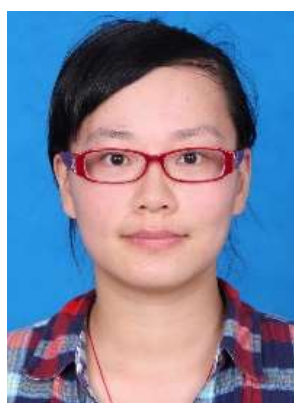
Yongni Shao was born in Cixi, Zhejiang, China, in 1983. She received the Ph.D. degree from the Zhejiang University, Zhejiang, China, in 2010. She is currently an Associate Professor with the Shanghai Key Lab of Modern Optical System, University of Shanghai for Science and Technology, Shanghai, China. Her current research involves biological application based on Terahertz technology.

Yutian Wang was born in Wuhu, China. He is currently working toward the master's degree in optical engineering at the Shanghai Key Lab of Modern Optical System, University of Shanghai for Science and Technology.

Zhi Zhu was born in Chenzhou, China, in 1988. He received the Ph.D. degree from the University of Chinese Academy of Sciences, China, in 2017. He is currently a research associate in University of Shanghai for Science and Technology, Shanghai, China. His current research interest involves biophysics and spectroscopy of soft materials, including terahertz spectroscopy.

Jialing Liu was born in Sichuan, China, in 1995.She received the B.S. degree from the University of Shanghai for Science and Technology, Shanghai, China, in 2018. She is currently working toward the M.S. degree at the University of Shanghai for Science and Technology, Shanghai, China.

Xiongxin was born in Hubei, China, in 1994. She received the B.S. degree from Hubei Engineering University, Shanghai, China, in 2018. She is currently working toward the M.S. degree at the University of Shanghai for Science and Technology, Shanghai, China.

Yiming Zhu received the B.S. and M.S. degrees in applied physics from Shanghai Jiaotong University, Shanghai, China, in 2002 and 2004, respectively, and the Ph.D. degree in electronics engineering from the University of Tokyo, Tokyo, Japan, in 2008. He is currently a Professor with the University of Shanghai for Science and Technology, Shanghai, China, and the Vice Director of the Shanghai Key Lab of Modern Optical System.

Ruoqing Zhai was born in Anhui, China, in 1997. She received the B.A. degree from Anhui University, Anhui, China, in 2018. She is currently working toward the M.T.I. degree at Tongji University, Shanghai, China.

Ming Gu was born in Shanghai, China, in 1997.She received the B.A. degree from Donghua University, Shanghai, China, in 2018.She is currently working toward the M.T.I. degree at Tongji University, Shanghai, China.

Tian Zhengan, graduated from Chinese Academy of Preventive Medicine majoring in epidemiology and health statistics in 2000, is the chief physician of Shanghai Customs, deputy director of Shanghai International Travel Medical Center, director of State key laboratory of hygienic supervision and inspection at frontier port. His current research interest involves surveillance and emergency treatment of infectious diseases at frontier.

Yan Peng studied at the Department of Physics, East China Normal University, Shanghai, China, from 2004 to 2009. In 2009, she joined the Shanghai Key Lab of Modern Optical System, University of Shanghai for Science and Technology, Shanghai. In 2013, she joined the Optical Research Center, University of Rochester, New York, NY, USA, as a Scientist. In 2014, she joined the University of Pennsylvania, as an Adjunct Professor. Since 2016, she has been working as a Professor with the University of Shanghai for Science and Technology. Her research interests include generation and regulation of terahertz waves and application of terahertz waves in biomedicine.