

# Three-Step One-Way Model in Terahertz Biomedical Detection

Yan Peng (✉ [py@usst.edu.cn](mailto:py@usst.edu.cn))

University of Shanghai for Science and Technology <https://orcid.org/0000-0002-8273-5566>

Jieli Huang

Tongji Hospital Department of Hematology

Jie Luo

University of Shanghai for Science and Technology

Zhangfan Yang

University of Shanghai for Science and Technology

Liping Wang

University of Shanghai for Science and Technology

Xu Wu

University of Shanghai for Science and Technology

Chen Yu

Tongji Hospital Department of Nephrology

Xiaofei Zang

University of Shanghai for Science and Technology

Qing Hu

Massachusetts Institute of Technology

Min Gu

University of Shanghai for Science and Technology University of Shanghai for Science and Technology

Xicheng Zhang

University of Rochester

Yiming Zhu

University of Shanghai for Science and Technology

Songlin Zhuang

University of Shanghai for Science and Technology

---

## Research Article

**Keywords:** Terahertz spectroscopy, renal fibrosis, biomedical detection, biomarker, mixture interference, individual difference

**Posted Date:** April 28th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-433019/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at PhotoniX on July 23rd, 2021. See the published version at <https://doi.org/10.1186/s43074-021-00034-0>.

# Three-step one-way model in terahertz biomedical detection

Yan Peng<sup>1,#</sup>, Jieli Huang<sup>2,#</sup>, Jie Luo<sup>1</sup>, Zhangfan Yang<sup>1</sup>, Liping Wang<sup>1</sup>, Xu Wu<sup>1</sup>, Chen Yu<sup>2,†</sup>, Xiaofei Zang<sup>1</sup>, Qing Hu<sup>1,3</sup>, Min Gu<sup>1,4</sup>, Xicheng Zhang<sup>5</sup>, Yiming Zhu<sup>1,\*</sup>, Songlin Zhuang<sup>1</sup>

<sup>1</sup>Terahertz Technology Innovation Research Institute, Shanghai Key Lab of Modern Optical System, and Shanghai Cooperation Innovation Centre of Terahertz Spectroscopy and Imaging Technology, University of Shanghai for Science and Technology, Shanghai 200093, P. R. China.

<sup>2</sup>Department of Nephrology, Tongji Hospital, Tongji University School of Medicine, Shanghai 200065, China.

<sup>3</sup>Department of Electrical Engineering and Computer Science and Research Laboratory of Electronics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.

<sup>4</sup>Laboratory of Artificial-Intelligence Nanophotonics, University of Shanghai for Science and Technology, Shanghai 200093, P. R. China.

<sup>5</sup>The Institute of Optics, University of Rochester, Rochester, New York 14627, USA.

<sup>#</sup>The first coauthor.

Corresponding author: \*ymzhu@usst.edu.cn, †yuchen@tongji.edu.cn

## Abstract

Terahertz technology has broad application prospects in biomedical detection. However, the mixed characteristics of actual samples make the terahertz spectrum complex and difficult to distinguish, and there is no practical terahertz detection method for medical clinics. Here, we propose a three-step one-way terahertz model, presenting a detailed flow analysis of terahertz technology in the biomedical detection of renal fibrosis as an example: **1) biomarker determination**: screening disease biomarkers and establishing the terahertz spectrum and concentration gradient; **2) mixture interference removal**: clearing the interfering signals in the mixture for the biomarker in the animal model and evaluating and retaining the effective characteristic peaks; and **3) individual difference removal**: excluding individual interference differences and confirming the final effective terahertz parameters in the human sample. The root mean square error of our model is three orders of magnitude lower than that of the gold standard, with profound implications for the rapid, accurate and early detection of diseases.

## Keywords

31 Terahertz spectroscopy, renal fibrosis, biomedical detection, biomarker, mixture interference,  
32 individual difference.

33

## 34 **Introduction**

35 The early detection, early diagnosis and early treatment of diseases directly affect the quality of life  
36 and survival rate of patients. Pathological diagnosis is the gold standard for the diagnosis of many  
37 diseases. The corresponding technologies include light microscopic morphological detection,  
38 immunohistochemical enzyme labelling, fluorescence in situ hybridization and gene rearrangement  
39 detection [1-5]. Light microscopic morphological detection can judge the nature of the disease by  
40 observing the morphological changes in cells but is subjective because this approach is influenced  
41 by the observer's experience and intuition. Immunochemical enzyme labelling and fluorescence in  
42 situ hybridization (FISH) both use the fluorescent labelling of specific proteins or nucleic acids for  
43 localization and quantitative labelling, and gene rearrangement detection technology is used to  
44 detect and diagnose diseases by gene sequencing of lesion tissue. However, the sample processing  
45 of these techniques is cumbersome and time-consuming and requires large amounts of reagents and  
46 dyes, so the samples cannot be reused, and the judgement of results also has subjective interference.  
47 Other spectroscopic methods, such as fluorescence spectroscopy and Raman spectroscopy, can also  
48 be used for sample detection. Fluorescence spectrometry mainly uses certain fluorescent dyes to  
49 form complexes with substances that do not ordinarily emit fluorescence but are induced to emit  
50 fluorescence before determination [6-9]. The procedure is complicated, the fluorescence intensity  
51 is affected by the wavelength and temperature of the stimulated luminescence, and the data stability  
52 is poor. Raman spectroscopy exposes the sample to a strong laser and collects the scattering signal  
53 for material identification [10, 11]. However, the high-power laser easily damages the samples,  
54 which is not conducive to secondary detection, and the detection error at low concentrations is large,

55 so the method cannot be applied to the early quantitative analysis of diseases. Therefore, there is an  
56 urgent need for rapid, label-free, simply operated, low-sample-loss, and low-cost detection  
57 technology for the early diagnosis of diseases.

58 Terahertz (THz) waves (0.1–10 THz) lie between the millimetre wave and infrared band,  
59 which is of significant importance to the biological sciences due to providing information  
60 complementary to traditional spectroscopic measurements of low-frequency bond vibrations,  
61 hydrogen bond stretching, and bond torsions in liquids and gases [12]. Therefore, the collective  
62 behaviour (vibration and rotation) characteristics of biomolecules make THz spectroscopy a  
63 promising sensing modality for clinical diagnosis [13]. Because of these properties and its non-  
64 destructiveness, accuracy, rapidity and good penetrability [14], THz spectroscopy also has many  
65 other potential applications in several research fields, including physics [15, 16], biology [17-21],  
66 chemistry [22, 23], and medicine [24, 25]. At present, substance identification with THz technology  
67 is mainly based on the fingerprint peaks of compounds. Some studies have indicated that the  
68 approach can be combined with chips, algorithms, reagents and other auxiliary means to improve  
69 the recognition accuracy [18, 26-30]. However, THz technology has not been effectively applied in  
70 actual medical diagnosis or evaluation, because human tissue is a mixture of many substances, and  
71 the THz fingerprint spectra of these substances overlap with each other, which interferes with the  
72 quantitative analysis of disease markers or key substances. More importantly, there are individual  
73 differences between human bodies, and the proportions of various components are continually  
74 changing, which leads to a rapid decrease in the effectiveness of auxiliary means such as algorithms.  
75 To date, the application of THz technology to biomedical sample detection is still in the process of  
76 exploration, and there is no standardized process or model to assist the accurate diagnosis of  
77 diseases.

78 In this paper, we propose a three-step one-way THz model to analyse the application steps and  
79 processes of using THz spectroscopy to detect real diseases. We hope that we can apply this model  
80 to improve the application of THz technology in research on related diseases and to realize the rapid  
81 and effective detection of pathological specimens of various diseases by using THz technology.

82 The THz three-step one-way model is divided into three steps, namely, biomarker  
83 determination, mixture interference removal, and individual difference removal. In the first step,  
84 we need to screen and determine the biomarkers or key substances (of which there can be multiple),  
85 test the THz fingerprint spectrum, and establish the gradient relationship between the biomarker  
86 concentration and THz parameters to lay the foundation for subsequent calibration of the disease  
87 pathological stage. Second, to verify the possibility of applying THz technology to recognize  
88 biomarkers in biological tissues (mixtures), we initially used animal models to predict and quantify  
89 these markers in accordance with scientific research ethics. It is possible to analyse whether the  
90 biomarkers of rat pathological samples can be identified by observing the pathological samples of  
91 rats. The third step is mainly to determine whether the individual differences in the human body  
92 will affect the qualitative identification and quantitative analysis of the biomarkers by THz  
93 spectroscopy and to verify the validity and accuracy of the biomarker parameters. The above three  
94 steps can be implemented only in one direction; otherwise, it is easy to reach a misdiagnosis or  
95 incorrect stage identification due to individual differences or mixture identification errors.

96 To assess the feasibility of the THz three-step one-way model, we used renal fibrosis, a  
97 common chronic disease, as an example. Renal fibrosis is a common pathological manifestation of  
98 end-stage renal disease; its main pathological features are glomerulosclerosis, tubular atrophy and  
99 increased extracellular matrix deposition [31]. Studies have shown that the assessment of renal  
100 fibrosis can guide the treatment and prognosis of autologous kidneys [32] and transplanted kidneys  
101 [33]. There are three methods for the diagnosis of fibrosis: histological evaluation, imaging

102 evaluation and biomarker detection. 1) Histology is the gold standard for the diagnosis of fibrosis  
103 and can directly observe the degree of fibrosis, but its identification depends on the subjective  
104 judgement of pathologists, and there are individual differences. In addition, in some cases, there are  
105 sampling errors in the tissues obtained from core biopsy [34]. Moreover, there are time limitations  
106 in histopathological detection: sample preparation takes a long time (approximately 15 hours),  
107 preventing rapid determination of the degree of fibrosis. 2) Imaging evaluation methods, including  
108 ultrasound and functional magnetic resonance imaging (MRI), are new noninvasive techniques.  
109 Renal fibrosis is reflected by observing the renal elasticity, oxygen content and blood perfusion.  
110 However, the identification of renal fibrosis is easily affected by many factors, such as blood  
111 pressure, body weight, respiratory movement and differences in subjective judgement among  
112 observers, resulting in low detection accuracy [35]. 3) The use of biomarkers in haematuria as a  
113 noninvasive detection method is expected to be usable to monitor the progression of renal fibrosis  
114 [36]. However, some promising biomarkers, such as microRNAs [37], still suffer from instability  
115 or lack of regularity as the disease progresses. Therefore, we expect to solve these problems through  
116 the THz three-step one-way model.

117

## 118 **Materials and Methods**

### 119 *Terahertz time-domain spectroscopy (THz-TDS)*

120 The experimental setup used here is THz time-domain spectroscopy. In our experiments, we used  
121 an 800 nm femtosecond laser with a pulse duration of 100 fs, a repetition rate of 76 MHz, and an  
122 average power of 150 mW. The emitted laser beam was split into a pump beam and probe beam by  
123 a 50/50 beam splitter. A pump beam modulated by an optical chopper was focused on a gallium  
124 arsenide (GaAs) photoconductive emitter of THz waves. The diverging THz beam was collected  
125 and focused by four off-axis paraboloid mirrors to pass through samples, and then, the probe beam

126 was used to detect the THz wave by photoconductive antennas. For our experimental system, the  
127 effective bandwidth for the measured signals ranged from 0.2 to 2.0 THz, the spectrum resolution  
128 was ~15 GHz, and the signal-to-noise ratio (SNR) was greater than 1000:1. All the spectra were  
129 averaged 256 times to ensure a high SNR. The sample absorbance  $\alpha(\omega)$  was calculated by using the  
130 following equation:  $\alpha(\omega)=\log (I_{\text{ref}}(\omega)/I_{\text{sam}}(\omega))/d$ , where  $d=0.1$  mm is the thickness of the sample,  
131  $I_{\text{sam}}(\omega)=E_{\text{sam}}(\omega)\times E_{\text{sam}}(\omega)^*$  is the power spectrum of the sample, and  $I_{\text{ref}}(\omega)=E_{\text{ref}}(\omega)\times E_{\text{ref}}(\omega)^*$  is the  
132 power spectrum of the reference signal.

### 133 *Preparation of pure L-hydroxyproline samples*

134 Pure L-hydroxyproline was purchased in powder form (Sigma-Aldrich) at the highest purity  
135 available ( $\geq 99\%$ ). The tablets were prepared with 30 mg of L-hydroxyproline and 120 mg of high-  
136 density polyethylene (powders, 40-48  $\mu\text{m}$  particle size), which were then pressed into discs under  
137 a pressure of 4 tons. The thickness of the sample tablet was approximately 1 mm, and the diameter  
138 was 13 mm. The entire THz spectroscopy system was enclosed in a sealed box filled with dry air  
139 (humidity  $< 2\%$ ) to reduce water vapour absorption.

### 140 *Variable-temperature liquid nitrogen cryostat*

141 A variable-temperature liquid nitrogen cryostat (Optistat DN2, Oxford Instruments) equipped with  
142 high-density polyethylene windows is a commercial product designed for optical spectroscopy. The  
143 cryostat temperature can continuously vary between 77 K and 500 K, with an accuracy of  $\pm 0.5$  K.  
144 In our experiments, the cryostat was placed at the THz radiation focal point between the off-axis  
145 parabolic mirrors of the THz-TDS system. The sample was placed inside in the cryostat, and THz  
146 radiation was passed through the sample to obtain the fingerprint spectrum.

### 147 *Preparation of animal model samples*

148 Male Sprague-Dawley (SD) rats aged 8-12 weeks with body weights of 180-200 g were provided  
149 by the Jiesijie Animal Experimental Center (Shanghai, China). All rats were nurtured and  
150 maintained according to the 'Care and Use of Laboratory Animals' guidelines published by the  
151 National Institute of Health. The Experimental Animal Ethics Committee of Shanghai Tongji  
152 Hospital (Shanghai, China, document No. 2020-DW-002) approved the research protocol. A total  
153 of 20 male SD rats were randomly divided into the following four groups: the sham surgery group  
154 (sham, n=5), the UUO d3 group (d3, n=5), the UUO d5 group (d5, n=5), the UUO d7 group (d7,  
155 n=5), and the UUO d14 group (d14, n=5). After anaesthesia by using 3% intraperitoneal  
156 pentobarbital (2 ml/kg), the UUO groups received left ureteral ligation in the sham group, and the  
157 left ureter was separated, but not ligated. On days 3, 7, and 14 after surgery, the rats were sacrificed  
158 using euthanasia. Their left kidneys were removed and homogenized and then placed into sterile  
159 tubes that were labelled. All sterile tubes were kept in a 200 K freezer. Then, rat left kidney  
160 homogenates were carefully mounted between two flat quartz plates with a groove thickness of 480  
161  $\mu\text{m}$ .

### 162 *Masson staining*

163 After the kidneys were harvested, they were sliced axially into 3-mm-thick sections, fixed in 10%  
164 buffered formalin, embedded in paraffin, and deparaffinized by submersing the slides in 4 changes  
165 of xylene for 2 minutes each, then dehydrated in 2 changes of absolute alcohol for 2 minutes, 1  
166 change of 95% alcohol for 2 minutes, and 1 change of 70% alcohol for 2 minutes. The slides were  
167 then mordanted by submersion in Bouin's fixative for 1 hour at 50-60°C, washed in gently running  
168 tap water until the yellow colour completely disappeared, rinsed in distilled 3 times, stained by  
169 submersion in Weigert's iron haematoxylin for 10 minutes, rinsed in warm tap water for 15 minutes,  
170 stained by submersion in Biebrich scarlet-acid fuchsin for 15 minutes, immersed in  
171 phosphomolybdic-phosphotungstic acid for 15 minutes, and immersed in Aniline blue solution for

172 15 minutes. Next, the slides were rinsed with distilled water 3 times, immersed in 1% acetic water  
173 for 1 to 3 minutes, rinsed in distilled water 3 times, dehydrated by submersion in 4 absolute alcohol  
174 changes for 2 minutes each, and cleared by submersion in 4 xylene changes for 2 minutes each.  
175 Finally, the coverslip was mounted in resinous mounting media.

176 ***Preparation of human kidney samples***

177 Fourteen human kidney samples were obtained from Tongji Hospital. The processing method of  
178 human kidneys was the same as that of rat kidneys. The trial was conducted in accordance with the  
179 Declaration of Helsinki and good clinical practice guidelines. The ethics committees of Tongji  
180 Hospital (Shanghai, China, document No. K-W-2021-001) approved the research protocol. All  
181 patients provided written informed consent before entering the trial.

182 ***Alkaline hydrolysis (AH)***

183 First, kidney tissue (40 mg wet weight) was sufficiently mixed with the digest (1 mL). After the  
184 test tube was covered by a stopper, the mixture was heated at 96°C for 20 minutes, and the pH was  
185 adjusted to 6.0. Second, distilled water was added to the test tube until the total volume reached 10  
186 mL. Next, diluted lysate (4 mL) and quantum active carbon (30 mg) were added to the test tube and  
187 mixed sufficiently. The mixture was then centrifuged at 3500 rpm for 10 minutes. The supernatant  
188 (1 mL) was transferred into a clean tube. The remainder of the experimental process is described in  
189 Table 1.

190 **Table 1.** Experimental process of alkaline hydrolysis (AH).

	Blank tube	Standard tube	Sample tube
Distilled water (mL)	1.0		
5 µg/mL standard working solution (mL)		1.0	
Supernatant			1.0

Dichloramine-T	0.5	0.5	0.5
Mix sufficiently and let stand for 10 minutes.			
Perchloric acid	0.5	0.5	0.5
Mix sufficiently and let stand for 10 minutes.			
Dimethylaminobenzaldehyde	0.5	0.5	0.5
The reagents were mixed well, and the mixture was placed in a 60°C water bath for 15 minutes, cooled to room temperature, and centrifuged at 3500 rpm for 10 minutes. The supernatant was taken and transferred into the spectrometer (at 550 nm) to measure the absorption (OD values) of all tubes (distilled water was used as a blank).			

191

192 Finally, the content of L-hydroxyproline can be calculated according to the following formula:

193 Hydroxyproline content  $\left(\frac{\mu\text{g}}{\text{mL}}\right)$

194 
$$= \frac{\text{OD}_{\text{Standard}} - \text{OD}_{\text{Blank}}}{\text{OD}_{\text{Standard}} - \text{OD}_{\text{Blank}}} \times \text{Standard content (5 } \mu\text{g/mL)}$$

195 
$$\times \frac{\text{Total volume of lysate (10 mL)}}{\text{Tissue wet weight (mg)}}$$

196 ***Liquid chromatography-mass spectrometry (LC-MS)***

197 Kidney tissue (10 mg) was cut into slices and then transferred into a screw vial containing 6 mol/L  
 198 HCl (1 mL). After the vial cap was screwed on tightly, the mixture was heated for 6 hours at 110°C.  
 199 The hydrolysed mixture was then centrifuged at 12000 rpm for 10 minutes at 4°C. The supernatant  
 200 (100 µL) was transferred into a clean Eppendorf tube containing 10 µL of internal standard (10  
 201 µg/mL proline-13C in water) and acetonitrile (ACN, 480 µL). After being mixed on a vortexer at  
 202 1400 rpm for 5 minutes, the mixture was centrifuged at 12000 rpm for 5 minutes at 4°C. The  
 203 supernatant (100 µL) was transferred into an Ultrafree centrifugal filter. After centrifugation at  
 204 12000 rpm for 5 minutes at 4°C, the filtrate (50 µL) was transferred into a 150 µL insert fixed in a  
 205 2 mL vial for analysis.

206 The prepared sample was separated on an Agilent 1260 Infinity liquid chromatography  
207 instrument equipped with a 2.1 × 100 mm Venusil ASB-C18 3.0 μm column (Agela Technologies,  
208 Tianjin, China) held at 25°C. The solvent system consisted of water/ACN/formic acid (v/v,  
209 949/50/3, solvent A) and methanol/formic acid (v/v, 998/2; solvent B). The gradient program  
210 started at 99% solvent A for 2.5 minutes, which was decreased linearly to 83% solvent A in 1.5  
211 minutes, then decreased linearly to 5% solvent A in 1 minute and maintained for 5 minutes, returned  
212 linearly to 99% solvent A in the following 0.5 minutes, and finally equilibrated for 2.0 minutes until  
213 the next analysis. The flow rate was 0.3 mL/min. The injection volume was 5 μL, and the samples  
214 were kept at 10°C in the autosampler. Analytes were monitored by positive mode electrospray  
215 ionization tandem mass spectrometry in MRM mode on an AB Sciex QTrap6500 Mass  
216 Spectrometer (AB Sciex, Framingham, MA). By direct infusion of the single-compound solution  
217 at a concentration of 1 μM, the declustering potential (DP), collision energy (CE), and collision cell  
218 exit potential (CXP) for each analyte and internal standard were optimized for the acquisition of  
219 the maximum intensity of the precursor and production ions, respectively. The precursor and  
220 dominant daughter ions used to set up the transition monitored in MRM mode are presented in  
221 Table 2. The dwell time was set to 40 ms for each compound. The accuracy of the analyte ranged  
222 from 90~110%, with an RSD less than 10%.

223 **Table 2.** The MRM transitions and associated parameters for the analyte and internal standard.

Compounds	Precursor ion (m z <sup>-1</sup> )	Daughter ion (m z <sup>-1</sup> )	DP (V)	CE (V)	CXP (V)
L-proline-1- <sup>13</sup> C	116.9	40.0	53.0	21.0	11.0
hydroxyproline	132.0	67.9	22.0	26.8	10.1

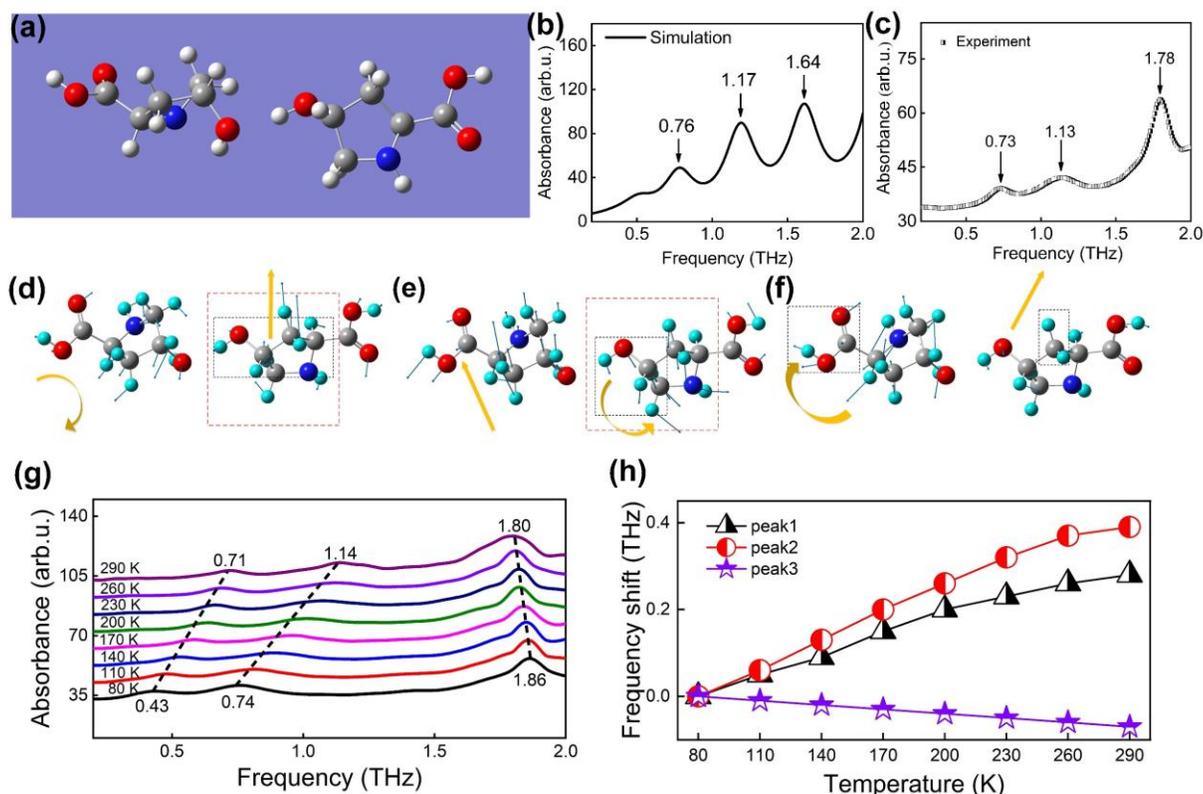
224

## 225 Results

### 226 STEP 1: Biomarker determination

227 By analysing the biological process and characteristics of fibrosis, we know that collagen is the  
228 main component of the extracellular matrix. The accumulation of extracellular matrix in renal tissue  
229 increases with the progression of fibrosis, and the content of collagen, as its main component, also  
230 increases. However, collagen itself is a mixture of a variety of proteins and substances, so it is not  
231 suitable for use as a biomarker for THz technology. Regarding L-hydroxyproline, the main  
232 component of collagen, 1) its concentration is relatively fixed in mammalian collagen  
233 (approximately 14.4%), which provides reliability for stage evaluation; 2) its molecular weight is  
234 low (131.1), and it may have clear characteristic peaks in the THz band; and 3) it is stable and  
235 difficult to react with other substances. Therefore, L-hydroxyproline has the potential to be used to  
236 identify the concentration of collagen and the developmental stage of renal fibrosis, and it was  
237 chosen as the biomarker here.

238 Before the experimental measurements, we needed to confirm whether the absorption peaks  
239 of the biomarker were located within the spectral range of the THz time-domain spectroscopy (THz-  
240 TDS) system. *Ab initio* calculations based on density functional theory (DFT) were performed to  
241 study the vibrational and rotational modes as well as the intermolecular interactions. The BP86  
242 functional and 6-31++G(d,p) basis set were applied. We employed a dimer model for the molecule,  
243 where the signal molecule structure of L-hydroxyproline was obtained from the ChemSpider  
244 website. Our DFT-based calculations and analyses were implemented in the Gaussian09 package  
245 [38].



**Fig. 1** Molecular simulation and THz spectrum of L-hydroxyproline. **a** Molecular model of L-hydroxyproline. **b** DFT calculated results of L-hydroxyproline. **c** Absorption spectrum of pure L-hydroxyproline measured experimentally by THz-TDS at room temperature. **d** to **f** Several vibration modes corresponding to the absorption peaks of L-hydroxyproline. **d** 0.76 THz **e** 1.17 THz **f** 1.64 THz. White, grey, red, and blue balls represent H, C, O, and N atoms, respectively. Yellow arrows indicate the vibrational direction of the atom. Black dotted and red dashed boxes refer to the dominant functional groups in the vibration modes and the entire molecular vibration of the ring, respectively. **g** Spectrum of L-hydroxyproline changes with the detection temperature from 290 K to 80 K at 30 K intervals. **h** Frequency shift of L-hydroxyproline absorption peaks as a function of temperature.

Fig. 1a&1b present the molecular model and the DFT results of L-hydroxyproline, where three absorption peaks are observed at 0.76, 1.17, and 1.64 THz, located within the spectral window of the THz-TDS system. The corresponding origins of these absorption peaks can be identified by DFT analysis, as shown in Table 3.

**Table 3.** Calculated absorption peaks of L-hydroxyproline and vibration model analysis.

Absorption peak	DFT results	Dominant vibration model
1	0.76 THz	The vibration of the propyl alcohol group of the right pyrrolidine ring. Under the intermolecular hydrogen bond, the left molecule vibrates back and forth following the right molecule [see Fig. 1d].

2	1.17 THz	The torsion and vibration of the right pyrrolidine ring. Under the intermolecular hydrogen bond, the up-and-down vibration of the left L-hydroxyproline molecule causes the absorption peak at 1.17 THz [see Fig. 1)].
3	1.64 THz	The combination of the right-left vibration of the right molecule and the strong contractions of the left molecule [see Fig. 1f].

261

262 After confirming that the absorption peaks were located within the THz spectral window, we  
 263 measured the THz absorption spectrum of pure L-hydroxyproline. The corresponding experimental  
 264 results are shown in Fig. 1c, revealing clear and distinctive absorption peaks at 0.73, 1.13, and 1.78  
 265 THz, which agree reasonably well with the DFT results. Additionally, considering that real tissue  
 266 samples provided by the hospital will be frozen and that the characteristic peaks are dependent on  
 267 the test temperature, we also tested pure L-hydroxyproline at different temperatures (290, 260, 230,  
 268 200, 170, 140, 110, and 80 K) by using a variable-temperature liquid nitrogen cryostat (see  
 269 Methods). The corresponding results are shown in Fig. 1g. For convenience, all spectra are offset  
 270 at a fixed interval. These peaks have different frequency shifts as the temperature increases: peak 3  
 271 presents the classical redshift, while peak 1 and peak 2 exhibit unusual blueshifts. For example, at  
 272 80 K, the absorption peaks are located at 0.64, 1.02, and 1.82 THz. To show this process more  
 273 clearly, we plot the frequency shift of absorption peaks as a function of temperature in Fig. 1h. The  
 274 corresponding frequency values are presented in Table 4.

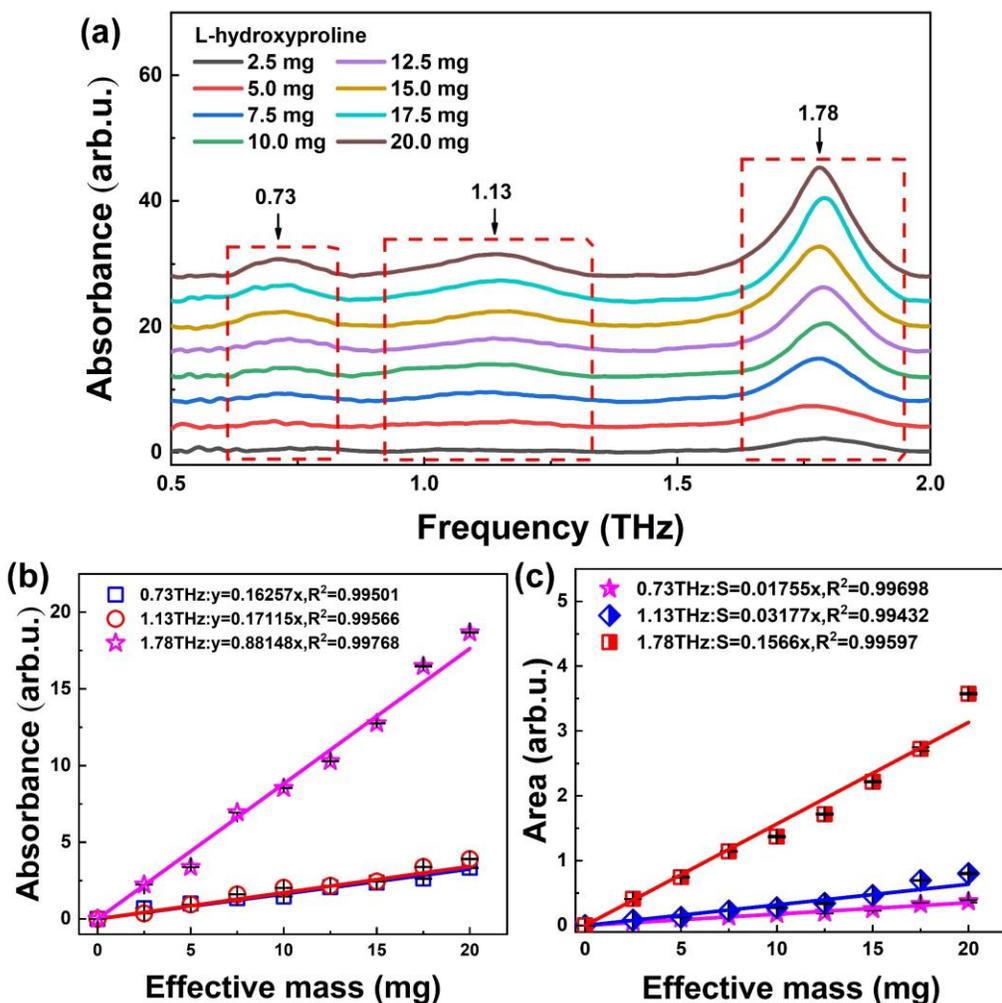
275 **Table 4.** The frequency values (THz) of pure L-hydroxyproline absorption peaks at different  
 276 temperatures.

Temperature	80 K	110 K	140 K	170 K	200 K	230 K	260 K	290 K
Peak 1	0.44	0.49	0.53	0.59	0.64	0.67	0.70	0.72
Peak 2	0.76	0.82	0.89	0.96	1.02	1.08	1.13	1.15
Peak 3	1.86	1.85	1.84	1.83	1.82	1.81	1.80	1.79

277

278 Similar to the classical case, peak 3 undergoes a redshift as the temperature increases. This  
279 temperature dependence of the spectra can be explained by the anharmonicity of the vibrational  
280 potentials [39]. Interestingly, peak 1 and peak 2 show different behaviours, and they exhibit a  
281 blueshift as the temperature increases. Based on the DFT analysis, we deduce that the blueshift is  
282 due to competition between the intramolecular and intermolecular hydrogen bonds. As the  
283 temperature increases, thermal expansion of the intermolecular hydrogen bond partially restores the  
284 electron density of intramolecular hydrogen bonds [40], resulting in the intramolecular hydrogen  
285 bond being “harder” and its coefficient of stiffness being “stronger”. Because the potential curve of  
286 the hydrogen bond near the local minimum is described by the harmonic oscillator potential, the  
287 vibration frequency is proportional to the stiffness coefficient. Therefore, as the temperature  
288 increases, the vibration frequency tends to increase, and the two absorption peaks undergo a  
289 blueshift.

290 For subsequent analysis of the relationship between the THz characteristic information and the  
291 evolution stage of renal fibrosis, we tested the THz spectrum of L-hydroxyproline at different  
292 concentrations (see Methods). As shown in Fig. 2a, when the content of L-hydroxyproline was  
293 varied from 2.5 mg to 20 mg, the corresponding THz absorption peaks exhibited a regular amplitude  
294 increase. For the convenience of identification, we extracted the amplitude and area of three  
295 characteristic peaks for further analysis.



296

297 **Fig. 2 Concentration gradient change of the L-hydroxyproline spectrum.** a THz absorption spectra of L-  
 298 hydroxyproline samples at different concentrations measured experimentally. b Absorbance at 0.73 THz (square), 1.13  
 299 THz (circle) and 1.78 THz (pentagram) for different concentrations of L-hydroxyproline samples. c The area under the  
 300 absorption peaks of all L-hydroxyproline samples; error bars have been labelled on each data point.

301 As shown in Fig. 2b and 2c, linear fittings on the data were performed according to the Beer-

302 Lambert law. The corresponding linear fitting function expressions are:

303 
$$y_{0.73\text{THz}} = 0.16257x, R^2 = 0.99501 \quad (1)$$

304 
$$y_{1.13\text{THz}} = 0.17115x, R^2 = 0.99566 \quad (2)$$

305 
$$y_{1.78\text{THz}} = 0.88148x, R^2 = 0.99768 \quad (3)$$

306 
$$S_{0.73\text{THz}} = 0.01755x, R^2 = 0.99698 \quad (4)$$

307 
$$S_{1.13\text{THz}} = 0.03177x, R^2 = 0.99432 \quad (5)$$

$$S_{1.78\text{THz}} = 0.1566x, R^2 = 0.99597 \quad (6)$$

x is the sample concentration, y is the amplitude S is the area in the expression, and  $R^2$  is the correlation coefficient. Eq. (1)-(3) correspond to the peak amplitudes at 0.73, 1.13 and 1.78 THz, and Eq. (4)-(6) correspond to the areas of these three peaks. The corresponding  $R^2$  values show that the amplitude and area of the three peaks of L-hydroxyproline all have a very good linear relationship with the concentration. Therefore, the concentration of L-hydroxyproline can be effectively calibrated and identified by its THz characteristic peak amplitude and area, which can be used to distinguish different stages of renal fibrosis.

## **STEP 2: Mixture interference removal**

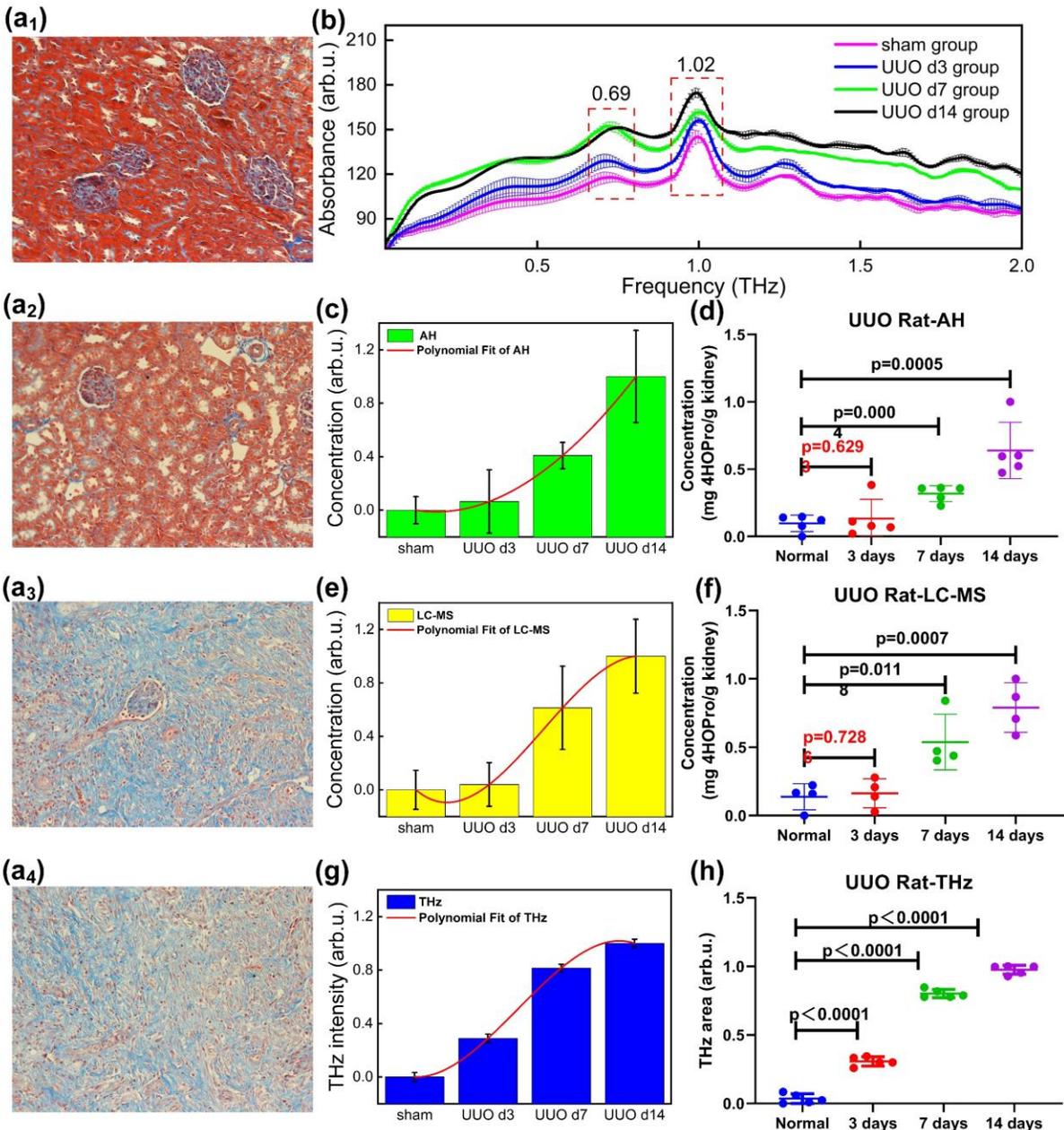
After confirming that L-hydroxyproline has a clear THz characteristic peak that changes linearly with concentration, we started the rat test to observe whether this biomarker can be recognized in tissue samples. The unilateral ureteral obstruction (UUO) rat model was used to study interstitial fibrosis (see Methods). A total of 20 male Sprague-Dawley rats were randomly divided into the following four groups: the sham group (sham) represented normal kidneys, the UUO day 3 group (D3) represented early injury, the UUO day 7 group (D7) represented moderate fibrosis, and the UUO day 14 group represented severe fibrosis (D14). The corresponding pathologic images and THz spectra of the four different groups are shown in Fig. 3.

Fig. 3a<sub>1</sub>-3a<sub>4</sub> shows the pathological features of renal tissue in rats. In the sham group, the renal tissue structure was normal. Three days after the UUO operation, the renal tubules were slightly dilated, the renal interstitium was oedematous, the interstitium was widened, and the “back-to-back” characteristics were lost. There were scattered infiltrating inflammatory cells in the tubulointerstitium. Seven days after UUO, some proximal tubules showed vacuolar degeneration, epithelial cells were seen in the renal tubules, the renal tubules at the junction of the medulla and

332 dermis medulla were obviously expanded, and a large number of inflammatory cells had infiltrated.  
333 On the 14th day after UUO, some tubules had disappeared, the collecting duct and distal renal  
334 tubule had expanded into cysts, inflammatory cell infiltration and fibroblast proliferation were  
335 obvious, and the cortex was thinner. These microscopic images are common evidence of clinical  
336 prognosis, but the accuracy of early diagnosis is low. For the early detection of fibrosis, alkaline  
337 hydrolysis (AH) or liquid chromatography-mass spectrometry (LC-MS) could be used to quantify  
338 the hydroxyproline contents and complete the diagnosis. Although these two methods are expensive  
339 and time-consuming, there is no better alternative in clinical practice. For the convenience of  
340 comparison, we prepared 5 rats for each stage of the disease. Each UUO kidney homogenate was  
341 divided into three separate parts for AH, LC-MS and THz detection.

342 Fig. 3b shows the THz absorption spectra of rat renal tissue at different stages after UUO  
343 operation. It can be clearly seen that the spectral shape of rat kidney tissue had changed compared  
344 with that in the spectrum of the biomarker L-hydroxyproline. In addition to the mixture causing  
345 other characteristic peaks (such as those at 0.15, 0.41, 1.27, and 1.63 THz), the relative intensity  
346 and discrimination of the peaks were changed. Fortunately, the absorption peaks at 0.69 and 1.02  
347 THz still had high amplitude and could be effectively distinguished from the background signal.  
348 However, the characteristic peak at 1.82 THz had undergone interference by the absorption lines of  
349 other substances in the mixture, making it difficult to distinguish effectively. Therefore, after animal  
350 model validation, only two of the three characteristic peaks were effective, and the absorption peak  
351 at 1.82 THz was discarded. Other characteristic peaks, such as those at 0.41 and 1.27 THz, showed  
352 significant changes, but they cannot be used for qualitative diagnosis and quantitative analysis of  
353 diseases because their correlation with the disease cannot be determined. Based on the two  
354 characteristic peaks at 0.69 and 1.02 THz, we can see that the overall amplitude of the characteristic  
355 peaks increases regularly with the evolution of pathological stage, which indicates that the

356 quantitative detection of L-hydroxyproline in rat kidney tissue by the THz-TDS system is expected  
 357 to provide a basis for accurate medical diagnosis.



358  
 359 **Fig. 3 Medical and THz test results in a UUO rat model.** a<sub>1</sub> - a<sub>4</sub> Kidney histologic changes in the UUO rat model.  
 360 a<sub>1</sub> Sham a<sub>2</sub> UUO d3 group a<sub>3</sub> UUO d7 group a<sub>4</sub> UUO d14 group (Masson staining, 200 times magnification). During  
 361 each stage, five rats were tested. b THz absorption spectra of rat kidneys at different stages of renal fibrosis. c - f  
 362 Hydroxyproline expression in different groups of UUO rat models by alkaline hydrolysis (c - d) and LC-MS/MS  
 363 methods (e - f). g - h THz-TDS detection in UUO kidneys by THz intensity g and THz area h.

364

365 For quantitative analysis, we calculated the area under the peak of these two absorption peaks  
366 of each spectrum; that is, by integrating the spectral area of 0.59-0.84 THz and 0.84-1.14 THz, the  
367 corresponding values were used as the judgement criteria of different fibrosis stages. The  
368 corresponding results are shown in Fig. 3c & 3e & 3g. Considering the different data units of the  
369 three methods, we presented the normalized values and the corresponding error bars to compare the  
370 accuracy and sensitivity in disease classification. It can be clearly seen that the measured average  
371 values of the AH and LC-MS methods are very close to each other. The difference between the  
372 second and first stages is very small, and the third and fourth stages are significantly different. In  
373 addition, the AH and LC-MS methods require a multi-step chemical pretreatment [see Methods]  
374 and are sensitive to the operator and operation process, so the error bars are large. The THz-TDS  
375 method presents much higher values at UUO d3 and d7, and the corresponding error bars are much  
376 smaller than those of AH and LC-MS because THz-TDS is a physical spectral detection method  
377 that requires no pre-processing, reducing sample loss and maximizing the effective signal. In  
378 addition, the non-destructiveness and repeatability of THz-TDS measurements greatly decrease the  
379 error rate. To help clarify the standard deviation, these data are polynomially fitted as follows:

380 LC-MS:  $y = 1.22 - 2.08x + 0.99x^2 - 0.12x^3$  (7)

381 AH:  $y = 0.25 - 0.41x + 0.17x^2 - 0.01x^3$  (8)

382 THz-TDS:  $y = 0.53 - 1.13x + 0.70x^2 - 0.10x^3$  (9)

383 where  $x$  represents the UUO days, and  $y$  represents the signal intensity detected by different  
384 methods. Mathematically, the THz fitting curve has the highest slope at low concentrations among  
385 all the methods; i.e., the THz method has higher sensitivity for lower-concentration samples. On  
386 the other hand, the 0.675% error bars of THz measurements (VS the 11.155% and 11.914% error  
387 bars of AH and LC-MS, respectively) obtained from the multiple measurements also indicate its

388 high repeatability in identifying the presence and concentration of the biomarker L-hydroxyproline  
 389 in real tissues.

390 Furthermore, to quantitatively compare the differences between these three test methods of  
 391 AH, LC-MS, and THz-TDS, we calculated the standard deviation and root mean square error for  
 392 the four stages of the UUO rat model, as presented in Table 5. The differences in magnitude ( $S_{THz}$   
 393 :  $S_{AH}$ :  $S_{LC-MS} = 10^{-2}$ :  $10^{-1}$ :  $10^{-1}$ ,  $RMSE_{THz}$ :  $RMSE_{AH}$ :  $RMSE_{LC-MS} = 10^{-5}$ :  $10^{-2}$ :  $10^{-2}$ ) directly prove  
 394 that the THz method has a higher accuracy and sensitivity than AH and LC-MS. These results are  
 395 highly useful in the early identification of the biomarker L-hydroxyproline with greater confidence.

396 **Table 5.** The standard deviation and root mean square error for the four stages of the UUO rat  
 397 model.

UUO rat kidney samples	Standard deviation (S)			Root mean square error (RMSE)		
	AH	LC-MS	THz	AH	LC-MS	THz
sham	1.01E-01	1.38E-01	3.38E-02	1.06E-02	1.62E-02	8.26E-05
UUO d3	2.37E-01	1.54E-01	3.20E-02	5.14E-02	1.86E-02	6.39E-05
UUO d7	9.79E-02	2.93E-01	2.86E-02	4.97E-03	2.80E-02	3.97E-05
UUO d14	3.45E-01	2.60E-01	3.10E-02	3.03E-02	1.44E-02	4.31E-05
<b>Average</b>	<b>1.95E-01</b>	<b>2.11E-01</b>	<b>3.14E-02</b>	<b>2.43E-02</b>	<b>1.93E-02</b>	<b>5.73E-05</b>

398  
 399 We analysed the results of the LC-MS, AH and THz tests by using the one-way analysis of  
 400 variance (ANOVA) method, as shown in Fig. 3d&3f&3h. It can be clearly seen that for both the  
 401 LC-MS and AH methods, the P values were all greater than 0.05 ( $P_{AH}=0.6293$ ,  $P_{MS}=0.7286$ ) in the  
 402 second stage (Fig. 3d&1f), which indicated that the early stage of fibrosis could not be effectively  
 403 distinguished. By comparison, THz detection showed good differentiation (Fig. 3h). For the third  
 404 and fourth stages, all three methods effectively identified fibrosis, but the THz method had a smaller

405 error and higher discrimination. Therefore, the accuracy of early diagnosis for the THz method is  
406 much higher than that of the other two conventional methods.

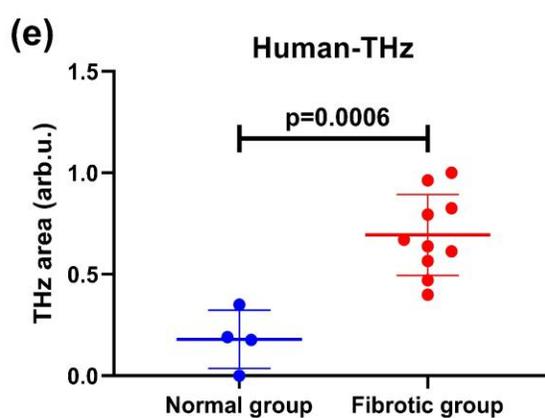
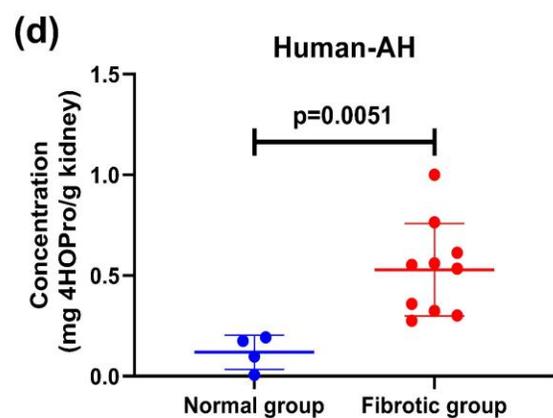
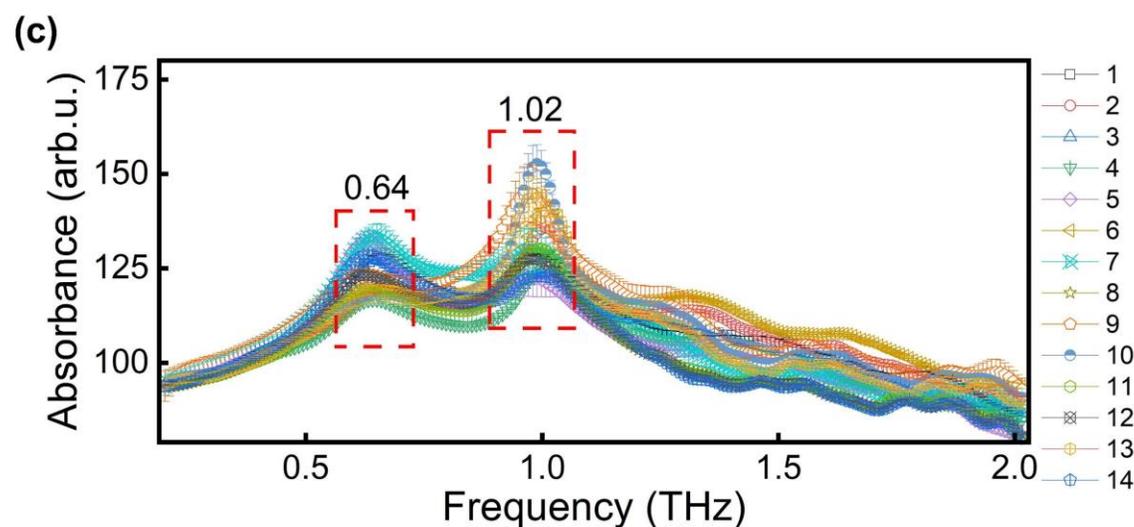
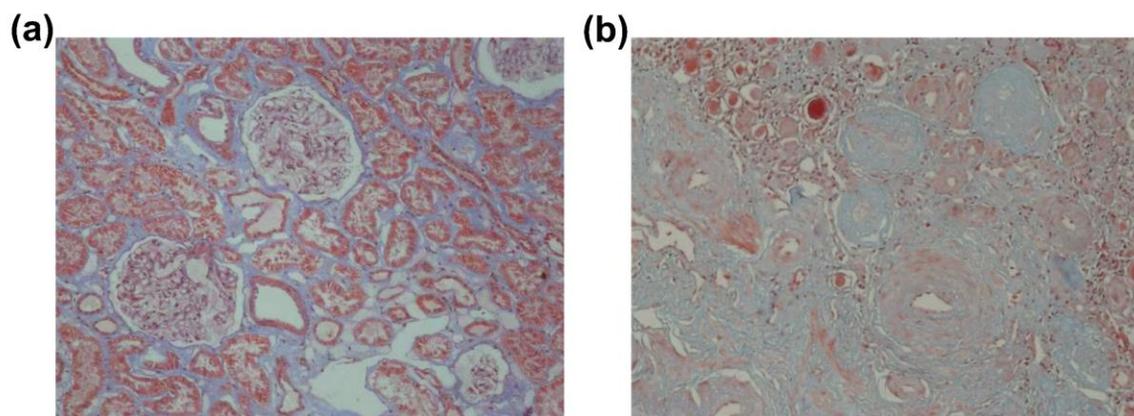
407  
408 **STEP 3: Individual difference removal**

409 After confirming the feasibility of tissue detection in a rat model, we carried out an experiment with  
410 real human kidney tissue. Because samples of human renal fibrosis tissue are very rare and the total  
411 amount of each sample is small, we first adopted the test mode of non-destructive detection with  
412 the THz method and then evaluated the degree of fibrosis by histopathology and the AH method  
413 (the amount of single tissue was not enough to support both the AH and LC-MS methods, so we  
414 chose only the conventional AH method). According to the area of renal fibrosis, the patients were  
415 divided into a normal group and a fibrosis group. The general information on the two groups is  
416 shown in Table 6.

417 **Table 6.** Clinical characteristics of 14 patients.

Groups	Sex (M:F)	Age (yr)	eGFR (ml/min)	Fibrotic area (%)
Normal	4:0	50.00±9.41	66.75±12.55	4.00±0.41
fibrosis	6:4	57.00±3.25	63.80±28.97	30.40±8.50

418



**Fig. 4 Medical and THz test results of human kidney tissues.** **a** Normal tissue image. **b** Fibrosis tissue image. **c** THz absorption spectra of 14 samples of human kidneys. Lines with different colours represent different human kidney samples, and the corresponding error bars are labelled. The red dotted box shows the THz spectra of the biomarker L-hydroxyproline. The numbers 1-14 represent the 14 human renal fibrosis samples. **d-e** The difference between human normal and fibrotic tissues by the AH and THz-TDS methods. All data are normalized to their maximum.

According to Masson staining, 14 renal tissues were divided into a normal group (n = 4) and a fibrosis group (n = 10). The average age of the normal group was  $50.00 \pm 9.41$  years (mean eGFR

428  $66.75 \pm 6.28$  mL/min), and the fibrotic area was  $4.00 \pm 0.82\%$ ; the average age of the fibrosis group  
429 was  $57.00 \pm 3.25$  years, the average eGFR was  $65.36 \pm 8.30$  mL/min, and the fibrotic area was  
430  $30.40 \pm 8.50\%$ . Fig. 4a shows a pathological image of normal human kidney tissue. The glomerular  
431 capillary loop is well opened, and the renal tubules are back-to-back. No obvious tubular atrophy,  
432 interstitial fibrosis or inflammatory cell infiltration was found. Fig. 4b presents a pathological image  
433 of a human fibrotic kidney with obvious glomerulosclerosis, tubular atrophy with interstitial  
434 fibrosis and inflammatory cell infiltration.

435 The THz absorption spectra of human samples are shown in Fig. 4c. The spectra of human  
436 kidney tissue changed again compared to those of the rat model. The central frequency of the first  
437 characteristic peak at 0.69 THz in the rat sample redshifted slightly to 0.64 THz, and the central  
438 frequency of the characteristic peak at 1.02 THz did not change; the biomarker-unrelated peaks at  
439 0.15, 0.41, 1.27, and 1.63 THz in Fig. 3b no longer exhibit obvious or regular changes in Fig. 4c.  
440 We speculate that this is because the rats in the UUO model were fed uniformly and injected  
441 quantitatively. However, the individual differences among human bodies (the proportion and  
442 content of each component varies with individual age, sex, physique, nutrition, exercise and other  
443 factors) will change the centre frequency of the overlapping spectrum, which causes the spectrum  
444 to no longer have obvious regular changes. However, the content of hydroxyproline, a common  
445 characteristic of renal fibrosis patients, increased significantly, making its characteristic peak still  
446 more significant and able to be effectively identified.

447 As in UUO, we integrated the area under the peak of the two characteristic peaks at 0.64 and  
448 1.02 THz, corresponding to the frequency range of 0.50-0.86 THz and 0.86-1.19 THz, respectively,  
449 as shown in the red box in Fig. 4c. The relevant data and AH test results were divided into two  
450 groups, normal and pathological, as shown in Fig. 4d and 4e. The results showed that the content  
451 of hydroxyproline and the area under the peaks by THz-TDS in renal fibrotic and normal tissues

452 were significantly increased, and there was a significant difference between the two groups by THz  
453 intensity and THz peak area assessment. Considering these results and the rapidity, simple operation  
454 and losslessness of the THz method, we conclude that the overall effect of the THz method is much  
455 better than that of the AH method.

## 456

### 457 **Discussion**

458 Now, we can deduce that these series of experiments directly prove the high sensitivity, accuracy  
459 and early recognition rate of THz methods for renal fibrosis. This sensitivity to different stages of  
460 renal fibrosis can be directly used for the dynamic monitoring of chronic kidney disease. This good  
461 performance may be due to four factors: (1) the high directionality of the THz beam yields a high  
462 collection efficiency; (2) the transmission-type detection results in a strong absorption of THz  
463 waves by each biomarker molecule in the sample; (3) the detection process for the THz method has  
464 only one step—inserting the sample into the system—which is quick and has no mass loss or  
465 denaturation; and (4) the non-destructive detection of the THz method allows repeated  
466 measurements so that we can obtain error bars to quantify the reliability of the result (a single THz  
467 measurement takes only a few seconds). Finally, the corresponding spectrum contains all the useful  
468 information. In contrast, the AH and LC-MS methods both use multiple and complex chemical  
469 reactions followed by spectral analyses. The more complex the workflows are, the more operation  
470 steps there are, which inevitably results in a reduced amount of sample and thereby a decreased  
471 signal.

472 So far, we have completed the three-step one-way model of the THz method: first, we screened  
473 the biomarkers of diseases to confirm whether the THz characteristic peaks exist in the effective  
474 detection range; second, by observing the interference degree of the mixture on the characteristic  
475 peaks of the biomarker by rat tissues, we confirmed the effective and recognizable characteristic

476 peaks; and third, we excluded the influence of individual differences in human tissues and  
477 confirmed the most effective and recognizable characteristic peaks. Finally, the THz characteristic  
478 peak can be used effectively, allowing accurate comparison and evaluation. This one-way process  
479 gradually eliminates the interference of uncertain factors and other non-disease-related substances  
480 and establishes a standardized process for the rapid and noninvasive diagnosis of clinical diseases  
481 by the THz method.

## 483 **Conclusion**

484 In summary, we provide a “THz three-step one-way” model for disease detection, which details the  
485 process of THz dynamic monitoring of disease evolution and the method of eliminating  
486 interference. Compared with AH and LC-MS, THz-TDS can improve the accuracy, sensitivity and  
487 detection speed of diagnosis. Its high sensitivity, high accuracy and high speed highlight its  
488 potential in the early diagnosis, staging evaluation and disease monitoring of biological diseases.  
489 In the future, this kind of dynamic monitoring is expected to achieve real-time imaging during  
490 operation, which can greatly improve the accuracy of surgery.

## 492 **Acknowledgements**

493 The authors are grateful to Professor Bo Song for fruitful discussions.

## 495 **Author contributions**

496 YP, YZ, JH and CY conceived the study and designed the experiments. JL and ZY performed the  
497 fibrosis experiments. LW, XW and XZ designed and performed theoretical simulations. JH and CY  
498 provided all samples, provided expertise for fibrosis progression, and performed the AH and LC-  
499 MS experiments in the hospital. QH, MG and SZ provided useful advice. YP and JH wrote the  
500 manuscript.

501 **Funding**

502 This work is financially supported by NSFC (61922059, 81961138014, 61805140, 81873609,  
503 61771314) and the Terahertz Science and Technology Frontier Project (6198802).

504

505 **Availability of data and materials**

506 The calculation and experiment data that support the works of this study are available from the  
507 corresponding authors on reasonable request.

508

509 **Competing interests**

510 The authors declare no competing interests.

511

512 **Authors' information**

513 <sup>1</sup>Terahertz Technology Innovation Research Institute, Shanghai Key Lab of Modern Optical  
514 System, and Shanghai Cooperation Innovation Centre of Terahertz Spectroscopy and Imaging  
515 Technology, University of Shanghai for Science and Technology, Shanghai 200093, P. R. China.

516 <sup>2</sup>Department of Nephrology, Tongji Hospital, Tongji University School of Medicine, Shanghai  
517 200065, China.

518 <sup>3</sup>Department of Electrical Engineering and Computer Science and Research Laboratory of  
519 Electronics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge,  
520 Massachusetts 02139, USA.

521 <sup>4</sup>Laboratory of Artificial-Intelligence Nanophotonics, University of Shanghai for Science and  
522 Technology, Shanghai 200093, P. R. China.

523 <sup>5</sup>The Institute of Optics, University of Rochester, Rochester, New York 14627, USA.

524

525 **References**

- 526 1. Lucchinetti CF, Popescu BFG, Bunyan RF, Moll NM, Roemer SF, Lassmann H, Brück W,  
527 Parisi JE, Scheithauer BW, Giannini C, Weigand SD, Mandrekar J, Ransohoff RM,  
528 Inflammatory Cortical Demyelination in Early Multiple Sclerosis. *N. Engl. J. Med.*  
529 2011,**365**:2188-97.
- 530 2. Longerich T, Endris V, Neumann O, Rempel E, Kirchner M, Abadi Z, Uhrig S, Kriegsman  
531 M, Weiss KH, Breuhahn K, Mehrabi A, Weber TF, Wilkens L, Straub BK, Rosenwald A,  
532 Schulze F, Brors B, Froehling S, Pellegrino R, Budczies J, Schirmacher P, Stenzinger A,  
533 RSPO2 gene rearrangement: a powerful driver of  $\beta$ -catenin activation in liver tumours. *Gut*  
534 2019, 68:317632.
- 535 3. Balázs M, Koroknai V, Szász I, Ecsedi S, "Detection of CCND1 Locus Amplification by  
536 Fluorescence In Situ Hybridization" in *The Retinoblastoma Protein*, P. G. Santiago-  
537 Cardona, Ed. (Springer New York, New York, NY, ), 2018:85-100.
- 538 4. Worlikar T, Mendiratta-Lala M, Vlaisavljevich E, Hubbard R, Shi J, Hall TL, Cho CS, Lee  
539 FT, Greve J, Xu Z, Effects of Histotripsy on Local Tumor Progression in vivo Orthotopic  
540 Rodent Liver Tumor Model. *BME Frontiers* 2020:9830304.
- 541 5. Stephens PJ, McBride DJ, Lin ML, Varela I, Pleasance ED, Simpson JT, Stebbings LA,  
542 Leroy C, Edkins S, Mudie LJ, Greenman CD, Jia M, Latimer C, Teague JW, Lau K W,  
543 Burton J, Quail MA, Swerdlow H, Churcher C, Natrajan R, Sieuwerts AM, Martens JWM,  
544 Silver DP, Langerød A, Russnes HEG, Foekens A, Reis-Filho JS, Veer L van 't, Richardson  
545 AL, Børresen-Dale AL, Campbell PJ, Futreal PA, Stratton MR, Complex landscapes of  
546 somatic rearrangement in human breast cancer genomes. *Nature*, 2009,462:1005-10.
- 547 6. Philip J, Potts AT, Ellis, Peter, Kregsamer, Christina, Strelj, X-ray fluorescence  
548 spectrometry. *J. Anal. At. Spectrom.*,1999.
- 549 7. Höhne J, Hohenberger C, Proescholdt M, Riemenschneider MJ, Wendl C, Brawanski A,  
550 Schebesch KM, Fluorescein sodium-guided resection of cerebral metastases—an update.  
551 *Acta Neurochir.* 2017,159:363-7.
- 552 8. Stummer W, Tonn JC, Goetz C, Ullrich W, Stepp H, Bink A, Pietsch T, Pichlmeier U, 5-  
553 Aminolevulinic Acid-derived Tumor Fluorescence: The Diagnostic Accuracy of Visible  
554 Fluorescence Qualities as Corroborated by Spectrometry and Histology and Postoperative  
555 Imaging. *Neurosurgery*, 2013,74, 310-20.
- 556 9. Ishizawa T, Fukushima N, Shibahara J, Masuda K, Tamura S, Aoki T, Hasegawa K, Beck  
557 Y, Fukayama M, Kokudo N, Real-time identification of liver cancers by using indocyanine  
558 green fluorescent imaging. *Cancer*, 2009, 115, 2491-504.
- 559 10. Zhu Y, Zheng W, Wang W. et al. Raman tensor of layered black phosphorus. *PhotoniX*,  
560 2020 ,1:17.
- 561 11. You A Y F, Bergholt MS, St-Pierre JP, Kit-Anan W, Pence IJ, Chester AH, Yacoub MH,  
562 Bertazzo S, Stevens MM, Raman spectroscopy imaging reveals interplay between  
563 atherosclerosis and medial calcification in the human aorta. *Sci. Adv.* 2017, 3: e1701156.

- 564 12. Peng Y, Shi C, Wu X, Zhu Y, Zhuang S, Terahertz Imaging and Spectroscopy in Cancer  
565 Diagnostics: A Technical Review. *BME Frontiers* 2020: 2547609.
- 566 13. Peng Y, Shi C, Zhu Y, Gu M, Zhuang S, Terahertz spectroscopy in biomedical field: a  
567 review on signal-to-noise ratio improvement. *Photonix*, 2020,1.
- 568 14. Ge H, Jiang Y, Lian F, Zhang Y, Xia S, Quantitative determination of aflatoxin B1  
569 concentration in acetonitrile by chemometric methods using terahertz spectroscopy. *Food*  
570 *Chem.* 2016, 209: 286-92.
- 571 15. Wang Z, Wu J, Yang W, Bera AK, Kamenskyi D, Islam ATMN, Xu S, Law JM, Lake B,  
572 Wu C, Loidl A, Experimental observation of Bethe strings. *Nature*, 2018, 554:219-23.
- 573 16. Chanana A, Liu X, Zhang C, Vardeny ZV, Nahata A, Ultrafast frequency-agile terahertz  
574 devices using methylammonium lead halide perovskites. *Sci. Adv.* 2018, 4: eaar7353.
- 575 17. Lu S, Zhang X, Zhang Z, Yang Y, Xiang Y, Quantitative measurements of binary amino  
576 acids mixtures in yellow foxtail millet by terahertz time domain spectroscopy. *Food Chem.*  
577 2016, 211:494-501.
- 578 18. Peng Y, Shi C, Xu M, Kou T, Wu X, Song B, Ma H, Guo S, Liu L, Zhu Y, Qualitative and  
579 Quantitative Identification of Components in Mixture by Terahertz Spectroscopy. *IEEE*  
580 *Trans. Terahertz Sci. Technol.* 2018, 8:696-701.
- 581 19. Wu X, Dai Y, Wang L, Peng Y, Lu L, Zhu Y, Shi Y, Zhuang S, Diagnosis of methylglyoxal  
582 in blood by using far-infrared spectroscopy and o-phenylenediamine derivation. *Biomed.*  
583 *Opt. Express*, 2020, 11:963-70.
- 584 20. Stantchev RI, Sun B, Hornett SM, Hobson PA, Gibson GM, Padgett MJ, Hendry E,  
585 Noninvasive, near-field terahertz imaging of hidden objects using a single-pixel detector.  
586 *Sci. Adv.* 2016, 2: e1600190.
- 587 21. Lee SH, Lee D, Choi MH, Son JH, Seo M, Highly Sensitive and Selective Detection of  
588 Steroid Hormones Using Terahertz Molecule-Specific Sensors. *Anal. Chem.* 2019, 91,  
589 6844-9.
- 590 22. Elgabarty H, Kampfrath T, Bonthuis DJ, Balos V, Kaliannan NK, Loche P, Netz RR, Wolf  
591 M, Kühne TD, Sajadi M, Energy transfer within the hydrogen bonding network of water  
592 following resonant terahertz excitation. *Sci. Adv.* 2020, 6: eaay7074.
- 593 23. Haddad J El, Miollis F de, Sleiman J Bou, Canioni L, Mounaix P, Bousquet B,  
594 Chemometrics Applied to Quantitative Analysis of Ternary Mixtures by Terahertz  
595 Spectroscopy. *Anal. Chem.* 2014, 86: 4927-33.
- 596 24. Wu X, Wang L, Peng Y, Wu F, Cao J, Chen X, Wu W, Yang H, Xing M, Zhu Y, Shi Y,  
597 Zhuang S, Quantitative analysis of direct oral anticoagulant rivaroxaban by terahertz  
598 spectroscopy. *Analyst*, 2020, 145: 3909-15.
- 599 25. Du SQ, Li H, Xie L, Chen L, Peng Y, Zhu YM, Li H, Dong P, Wang JT, Vibrational  
600 frequencies of anti-diabetic drug studied by terahertz time-domain spectroscopy. *Appl.*  
601 *Phys. Lett.* 2012, 100: 143702.

- 602 26. Sommer S, Koch M, Adams A, Terahertz Time-Domain Spectroscopy of Plasticized  
603 Poly(vinyl chloride). *Anal. Chem.* 2018, 90: 2409-13.
- 604 27. Gu H, Shi C, Wu X, Peng Y, Molecular methylation detection based on terahertz  
605 metamaterial technology. *Analyst*, 2020, 145: 6705-12.
- 606 28. Peng Y, Yuan X, Zou X, Chen W, Huang H, Zhao H, Song B, Chen L, Zhu Y, Terahertz  
607 identification and quantification of neurotransmitter and neurotrophin mixture. *Biomed. Opt.*  
608 *Express* 7, 2016,7: 4472-9.
- 609 29. Li T, Ma H, Peng Y, Chen X, Zhu Z, Wu X, Kou T, Song B, Guo S, Liu L, Zhu Y, Gaussian  
610 numerical analysis and terahertz spectroscopic measurement of homocysteine. *Biomed.*  
611 *Opt. Express*, 2018 9: 5467-76.
- 612 30. Ajito K, Ueno Y, Kim JY, Sumikama T, Capturing the Freeze-Drying Dynamics of NaCl  
613 Nanoparticles Using THz Spectroscopy. *J. Am. Chem. Soc.* 2018, 140, 13793-7.
- 614 31. Zhou D, Liu Y, Understanding the mechanisms of kidney fibrosis. *Nat. Rev. Nephrol.* 2016,  
615 12: 68-70.
- 616 32. Malvar A, Pirruccio P, Alberton V, Lococo B, Recalde C, Fazini B, Nagaraja H, Indrakanti  
617 D, Rovin BH, Histologic versus clinical remission in proliferative lupus nephritis. *Nephrol.,*  
618 *Dial., Transplant.* 2017, 32: 1338-44.
- 619 33. Mengel M, Chapman JR, Cosio FG, Cavaille-Coll MW, Haller H, Halloran PF, Kirk AD,  
620 Mihatsch MJ, Nankivell BJ, Racusen LC, Roberts IS, Rush DN, Schwarz A, Seron D,  
621 Stegall MD, Colvin RB, Protocol biopsies in renal transplantation: insights into patient  
622 management and pathogenesis. *Am. J. Transplant.* 2007, 7: 512-7.
- 623 34. Zhuo L, Wang H, Chen D, Lu H, Zou G, Li W, Alternative renal biopsies: past and present.  
624 *Int. Urol. Nephrol.* 2018, 50: 475-9.
- 625 35. Torres HR, Queirós S, Morais P, Oliveira B, Fonseca JC, Vilaça JL, Kidney segmentation  
626 in ultrasound, magnetic resonance and computed tomography images: A systematic review.  
627 *Comput. Meth. Prog. Bio.* 2018, 157: 49-67.
- 628 36. Zhang X, Li X, Zhou W, Liu X, Huang J, Zhang Y, Lindholm B, Yu C, Serum Lysyl  
629 Oxidase Is a Potential Diagnostic Biomarker for Kidney Fibrosis. *Am. J. Nephrol.* 2020, 51:  
630 907-18.
- 631 37. Nakhjavani M, Etemadi J, Poulak T, Mirhosaini Z, Zununi Vahed S, Abediazar S, Plasma  
632 levels of miR-21, miR-150, miR-423 in patients with lupus nephritis. *Iran. J. Kidney. Dis.*  
633 2019, 13: 198-206.
- 634 38. Frisch MJ, Trucks G, Schlegel HB, Scuseria GE, Robb MA, Cheeseman J, Scalmani G,  
635 Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP,  
636 Izmaylov AF, Bloino J, Zheng G, Sonnenberg J, Hada M, Fox D, Gaussian 09 Revision  
637 A.1. Gaussian Inc. 2009.
- 638 39. Joseph J, Jemmis ED, Red-, Blue-, or No-Shift in Hydrogen Bonds: A Unified Explanation.  
639 *J. Am. Chem. Soc.* 2007, 129: 4620-32.

640 40 Shen YC, Upadhyay PC, Linfield EH, Davies AG, Temperature-dependent low-frequency  
641 vibrational spectra of purine and adenine. *Appl. Phys. Lett.* 2003, 82: 2350-2.  
642

# Figures

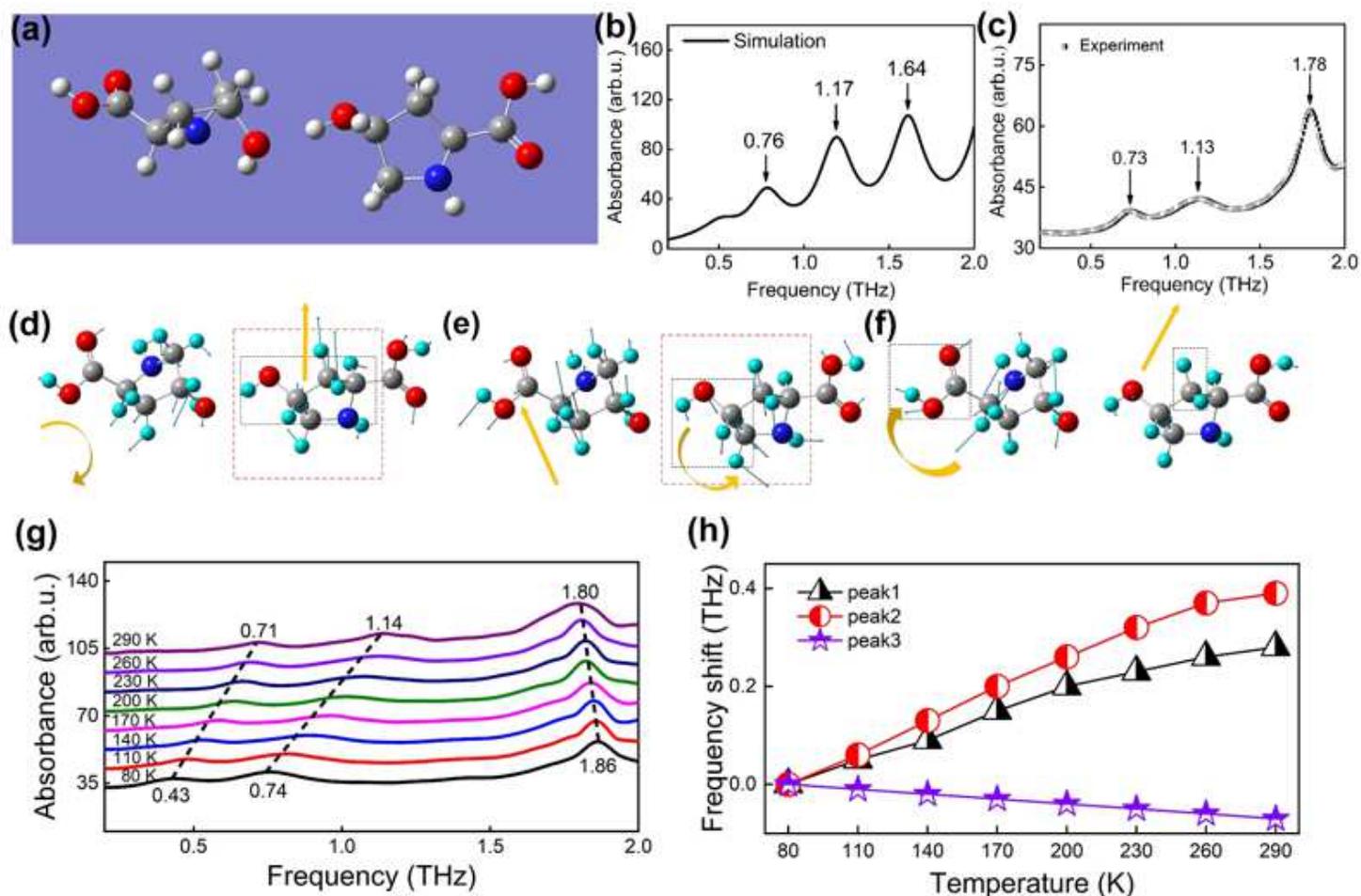


Figure 1

Molecular simulation and THz spectrum of L-hydroxyproline. a Molecular model of L-hydroxyproline. b DFT calculated results of L-hydroxyproline. c Absorption spectrum of pure L-hydroxyproline measured experimentally by THz-TDS at room temperature. d to f Several vibration modes corresponding to the absorption peaks of L-hydroxyproline. d 0.76 THz e 1.17 THz f 1.64 THz. White, grey, red, and blue balls represent H, C, O, and N atoms, respectively. Yellow arrows indicate the vibrational direction of the atom. Black dotted and red dashed boxes refer to the dominant functional groups in the vibration modes and the entire molecular vibration of the ring, respectively. g Spectrum of L-hydroxyproline changes with the detection temperature from 290 K to 80 K at 30 K intervals. h Frequency shift of L-hydroxyproline absorption peaks as a function of temperature.

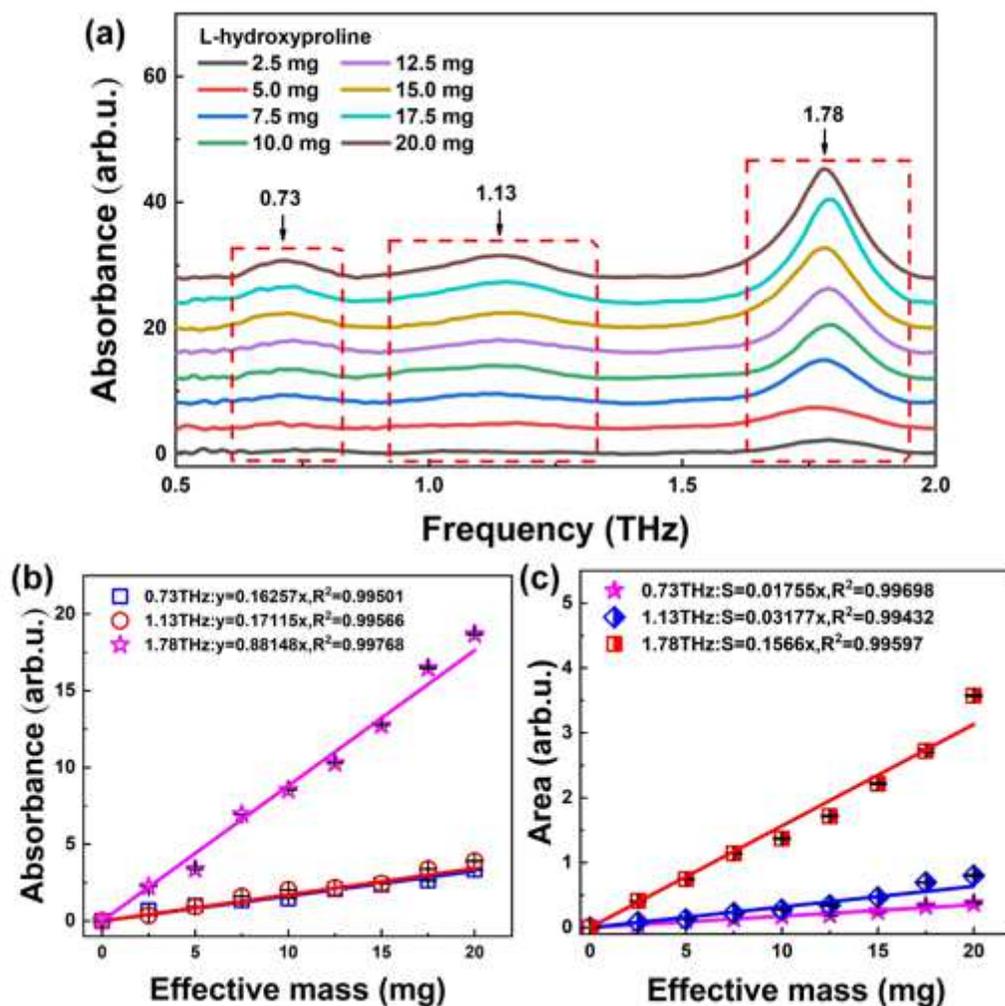
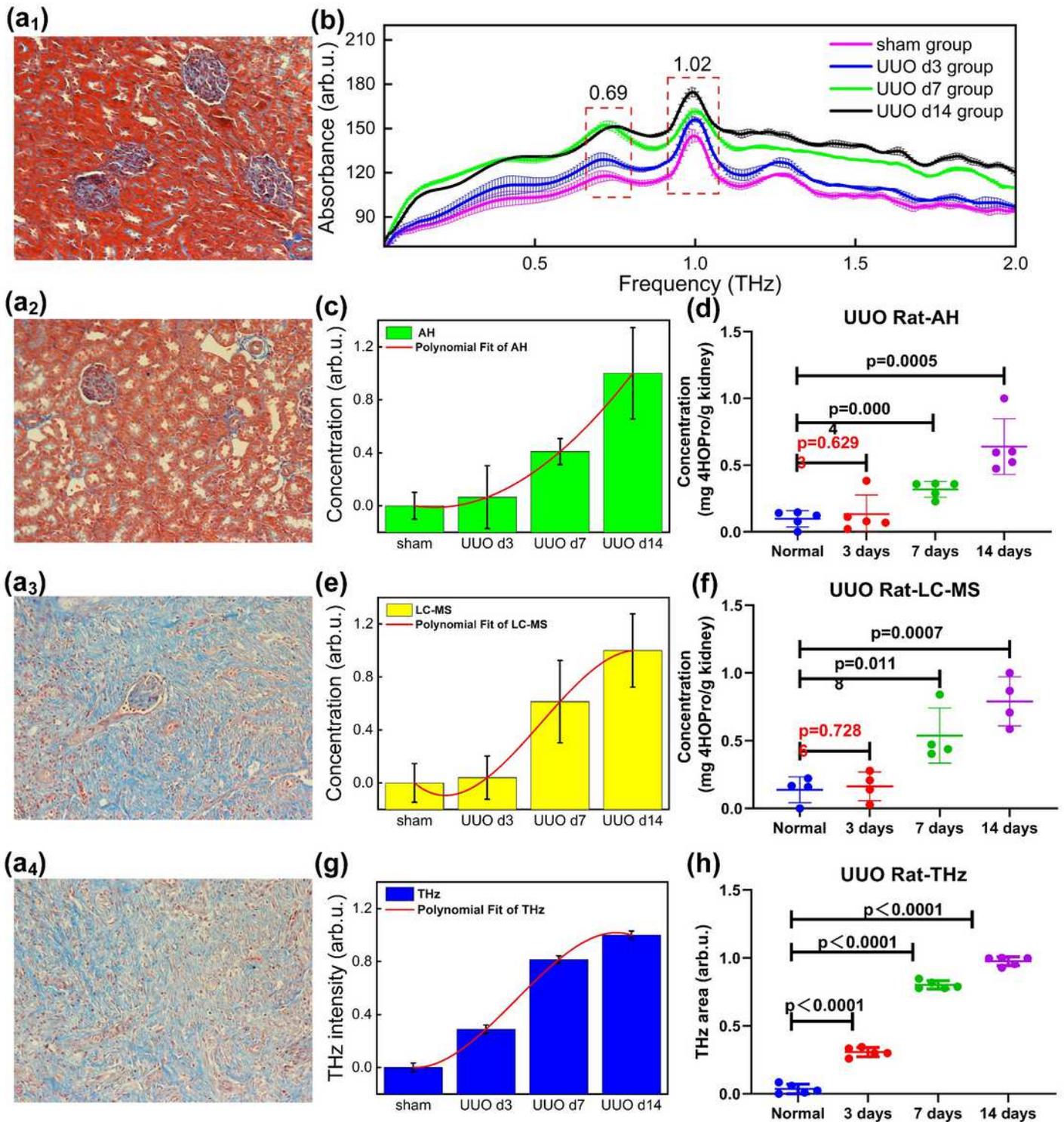


Figure 2

Concentration gradient change of the L-hydroxyproline spectrum. a THz absorption spectra of L-hydroxyproline samples at different concentrations measured experimentally. b Absorbance at 0.73 THz (square), 1.13 THz (circle) and 1.78 THz (pentagram) for different concentrations of L-hydroxyproline samples. c The area under the absorption peaks of all L-hydroxyproline samples; error bars have been labelled on each data point.



**Figure 3**

Medical and THz test results in a UUO rat model. a1 - a4 Kidney histologic changes in the UUO rat model. a1 Sham a2 UUO d3 group a3 UUO d7 group a4 UUO d14 group (Masson staining, 200 times magnification). During each stage, five rats were tested. b THz absorption spectra of rat kidneys at different stages of renal fibrosis. c - f Hydroxyproline expression in different groups of UUO rat models by

alkaline hydrolysis (c - d) and LC-MS/MS methods (e - f). g - h THz-TDS detection in UUO kidneys by THz intensity g and THz area h.

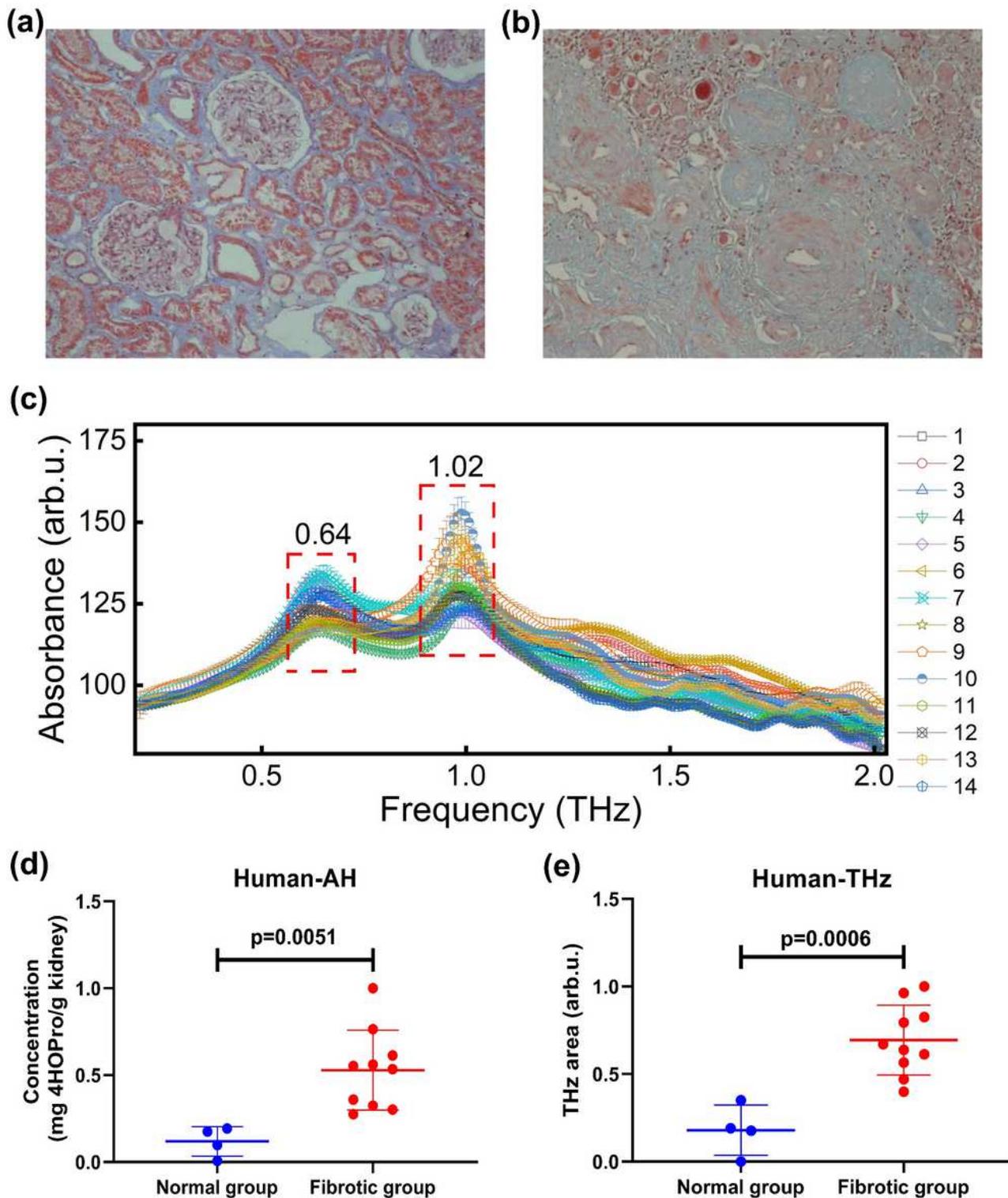


Figure 4

Medical and THz test results of human kidney tissues. a Normal tissue image. b Fibrosis tissue image. c THz absorption spectra of 14 samples of human kidneys. Lines with different colours represent different human kidney samples, and the corresponding error bars are labelled. The red dotted box shows the THz

spectra of the biomarker L-hydroxyproline. The numbers 1-14 represent the 14 human renal fibrosis samples. d-e The difference between human normal and fibrotic tissues by the AH and THz-TDS methods. All data are normalized to their maximum.