

# DNA, RNA 碱基和稀有碱基的太赫兹光谱检测

汤明杰<sup>1,2</sup>, 颜识涵<sup>1</sup>, 张明焜<sup>1</sup>, 魏东山<sup>1</sup>, 杜春雷<sup>1\*</sup>, 崔洪亮<sup>1</sup>

1. 中国科学院重庆绿色智能技术研究院, 重庆市跨尺度制造重点实验室, 重庆 400714  
2. 中国科学院大学, 北京 100049

**摘要** 核酸分子碱基的光谱特征分析对生物遗传学研究具有重要意义。探讨太赫兹时域光谱 (THz-TDS) 技术和拉曼光谱 (Raman spectroscopy) 技术用于 DNA 和 RNA 碱基和稀有碱基的光谱检测的可行性。分析了 7 种固体核酸碱基的太赫兹振动光谱和拉曼光谱, 对 THz-TDS 技术和拉曼光谱技术进行了对比研究。实验结果表明, 在 THz-TDS 实验中, 胞嘧啶 (cytosine, C)、鸟嘌呤 (guanine, G)、腺嘌呤 (adenine, A)、胸腺嘧啶 (thymine, T, DNA 专有) 和尿嘧啶 (uracil, U, RNA 专有) 和稀有碱基 [5-甲基胞嘧啶, 5-methylcytosine ( $m^5C$ )、1-甲基腺嘌呤, 1-methyladenine ( $m^1A$ )] 在 0.2~2.0 THz 频段内特征吸收峰和吸收强度差异显著, 可以直观地识别出 7 种碱基的差异; 拉曼光谱中, 7 种碱基也表现出很多明显的不同特征峰, 而正是由于拉曼光谱的特征峰杂而多, 故不能直观的识别多种物质。且其吸收强度的差异与粉末的厚度, 粒度和光聚焦的深度有关, 样品的荧光还会给拉曼光谱带来干扰, 同时激光还可能会损伤生物样品。这说明此两种技术均能够识别 7 种常见和稀有碱基, THz-TDS 技术在识别这 7 种碱基的能力上优于拉曼光谱技术, 表现出较为简洁, 快速和无损的检测性能。THz-TDS 技术不仅为 DNA 和 RNA 碱基和稀有碱基的识别提供了一种快速和准确的检测方法, 也为生物遗传学研究奠定实验基础。

**关键词** 碱基; 稀有碱基; 太赫兹时域光谱; 拉曼光谱

**中图分类号:** O657.6 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)08-2388-04

## 引言

基因是携带了遗传信息的 DNA 或 RNA 序列, 是控制生物性状的基本遗传单位。碱基是核酸分子 (包括 DNA 和 RNA) 的重要组成部分, 此外, 核酸中还有一些含量较少的稀有碱基, 大部分是碱基的甲基衍生物。碱基是核酸信息存储和传递的基础, 是生物遗传、进化和变异的根本原因, 所以研究核酸分子碱基对研究生物活性和生物行为具有重要意义<sup>[1-5]</sup>。因此, 研究核酸分子碱基的光谱特征对于表征核酸分子结构的改变和核酸分子与小分子的相互作用提供参考价值。

太赫兹 (THz) 辐射介于微波和红外之间, 频率范围大概在 0.1~10 THz。虽然有些太赫兹辐射会影响 DNA 的动态过程, 进而影响 DNA 的复制和基因表达<sup>[6-7]</sup>, 甚至高强度的太赫兹脉冲会造成 DNA 损伤<sup>[8]</sup>, 但是辐射强度低的太赫兹并不会影响 DNA 分子的结构和功能<sup>[9-10]</sup>。近年来, 太赫兹时

域光谱技术 (Terahertz time-domain spectroscopy, THz-TDS) 已被广泛应用于化学检测、材料鉴定和生物医学<sup>[11-13]</sup>。太赫兹光谱对分子间氢键的相互作用非常敏感, 许多低频的振动模式 (例如骨架振动) 位于 THz 波段。

采用 THz-TDS 检测了 DNA、RNA 碱基 (胞嘧啶 (C)、鸟嘌呤 (G)、腺嘌呤 (A)、胸腺嘧啶 (T, DNA 专有) 和尿嘧啶 (U, RNA 专有) 和稀有碱基 [5-甲基胞嘧啶 ( $m^5C$ )、1-甲基腺嘌呤 ( $m^1A$ )] 的太赫兹光谱, 同时与拉曼光谱技术对 7 种碱基的检测能力作了对比, 探讨了两种光谱技术检测 7 种核酸碱基的可行性。

## 1 实验部分

### 1.1 样品制备

7 种核酸碱基的分子结构式如图 1 所示。制备每种碱基样品时, 先称取 0.12 g 左右粉末放入直径为 13 mm 的压片磨具内, 轻微震荡使其分布均匀, 使用压片机在 2 MPa 的压

收稿日期: 2019-07-16, 修订日期: 2019-11-09

基金项目: 国家自然科学基金项目 (61605206), 重庆市自然科学基金项目 (cstc2019jcyj-msxmX0654) 资助

作者简介: 汤明杰, 女, 1985 年生, 中国科学院重庆绿色智能技术研究院工程师 e-mail: mjtang@cigit.ac.cn

\* 通讯联系人 e-mail: cldu@cigit.ac.cn

力下做成厚度为 0.6~0.9 mm 的圆盘状薄片进行太赫兹时域光谱和拉曼光谱实验。实验中采用了一个自行设计的样品夹具将样品夹在带有小孔的两个靠磁性吸附在一起的薄片之间, 样品尺寸略大于小孔(如图 2 所示)。



图 1 7 种核酸碱基的分子结构式

Fig. 1 7 Molecular structural formulas of nucleic acid bases

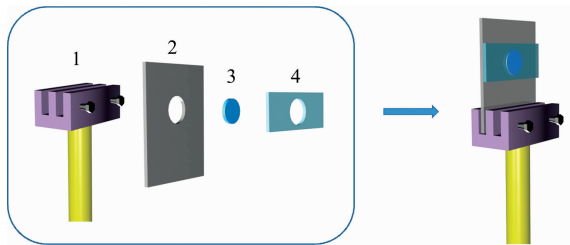


图 2 自制样品夹具示意图

1: 夹具底座; 2: 带孔铁片; 3: 样品片; 4: 带孔的橡胶薄片

Fig. 2 The schematic diagram of self-made sample fixture

1: Fixture base; 2: Iron sheet with holes;  
3: Sample sheet; 4: Rubber sheet with holes

### 1.2 碱基的太赫兹时域光谱检测

采用 T-Ray 5000 THz-TDS 系统 (Advanced Photonix Inc., 美国)。实验在室温(约 294 K)下进行, 为了避免空气中水分对 THz 波的影响, 实验装置放在充有氮气的密封箱内, 保证样品在测量时的湿度被控制在 3% 左右。在分析碱基的太赫兹时域光谱时, 首先对太赫兹时域光谱进行傅里叶变换求取频域光谱, 然后根据如下物理模型求取样品的吸收系数

$$n(\omega) = \frac{\varphi(\omega)c}{\omega d} + 1$$

$$a(\omega) = \frac{2}{d} \ln \frac{4n(\omega)}{\rho(\omega)(n(\omega) + 1)^2}$$

其中,  $n$  为样品折射率,  $\varphi(\omega)$  为样品信号和参考信号的相位差,  $\omega$  为角频率,  $d$  为样品的厚度,  $a$  为样品吸收系数,  $\rho(\omega)$  为样品信号和参考信号的振幅之比。

### 1.3 碱基的拉曼光谱检测

测试样品拉曼光谱所采用的设备为英国 Renishaw inVia Reflex 拉曼光谱仪, 直接对粉末进行测量。选用 532 nm 激光器。实验时, 首先用洁净的硅片进行校准, 校准硅片的拉曼峰值位置在  $520 \text{ cm}^{-1}$ 。在分析碱基的拉曼光谱时, 将得到的碱基拉曼光谱分别进行基线校准, 平滑和归一化处理。

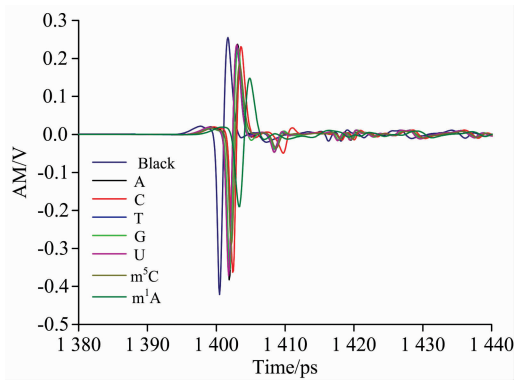


图 3 DNA、RNA 碱基(胞嘧啶(C)、鸟嘌呤(G)、腺嘌呤(A)、胸腺嘧啶(T, DNA 专有)和尿嘧啶(U, RNA 专有)和稀有碱基[5-甲基胞嘧啶(m<sup>5</sup>C)、1-甲基腺嘌呤(m<sup>1</sup>A)]和背景样品池的太赫兹时域光谱

Fig. 3 THz time-domain spectra of DNA, RNA bases (cytosine (C), guanine (G), adenine (A), thymine (T, DNA-specific) and uracil (U, RNA-specific) and rare bases (5-methylcytosine (m<sup>5</sup>C), 1-methyladenine (m<sup>1</sup>A) and background sample cell

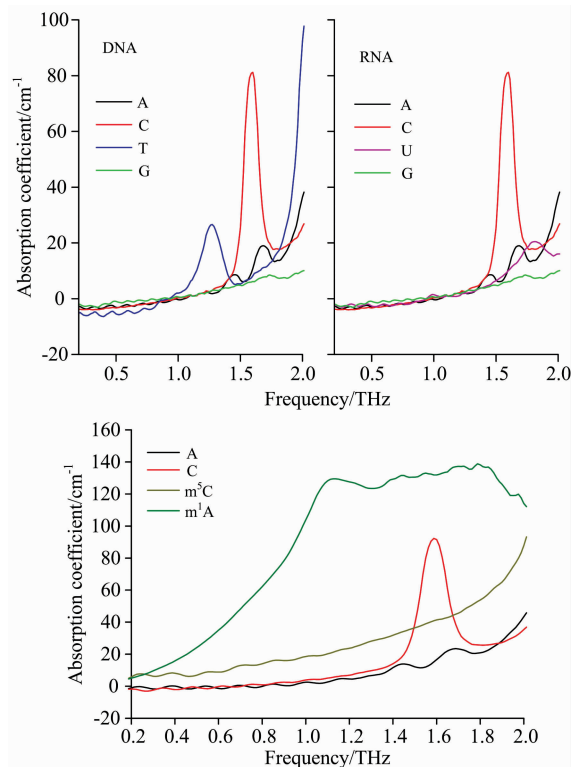


图 4 DNA、RNA 碱基(胞嘧啶(C)、鸟嘌呤(G)、腺嘌呤(A)、胸腺嘧啶(T, DNA 专有)和尿嘧啶(U, RNA 专有)和稀有碱基(5-甲基胞嘧啶(m<sup>5</sup>C)、1-甲基腺嘌呤(m<sup>1</sup>A))在 0.2~2.0 THz 的太赫兹吸收谱

Fig. 4 Terahertz absorption spectra of DNA, RNA bases (cytosine (C), guanine (G), adenine (A), thymine (T, DNA-specific) and uracil (U, RNA-specific) and rare bases (5-methylcytosine (m<sup>5</sup>C), 1-methyladenine (m<sup>1</sup>A)) at 0.2 to 2.0 THz

## 2 结果与讨论

### 2.1 碱基的太赫兹时域光谱分析

图 3 为 DNA 和 RNA 碱基(胞嘧啶(C)、鸟嘌呤(G)、腺嘌呤(A)、胸腺嘧啶(T, DNA 专有)和尿嘧啶(U, RNA 专有)和稀有碱基[5-甲基胞嘧啶( $m^5C$ )、1-甲基腺嘌呤( $m^1A$ )]和背景样品池的太赫兹时域光谱。图 4 为 DNA、RNA 碱基(胞嘧啶(C)、鸟嘌呤(G)、腺嘌呤(A)、胸腺嘧啶(T, DNA 专有)和尿嘧啶(U, RNA 专有)和稀有碱基(5-甲基胞嘧啶( $m^5C$ )、1-甲基腺嘌呤( $m^1A$ ))在 0.2~2.0 THz 的太赫兹吸收谱。从图中可以看出, DNA 的 4 种碱基差异明显, 只有 G 没有明显吸收峰, A 有两个吸收峰, T 和 C 各有一个吸收峰, 峰位不同; RNA 的 4 种碱基差异也很明显, G 没有明显吸收峰, A 有两个吸收峰, U 和 C 各有一个吸收峰, 峰位不同; 碱基 C 和稀有碱基  $m^5C$  也有很大差异, 稀有碱基  $m^5C$  没有明显的特征峰, 但吸收系数较高, 随着频率的增加,  $m^5C$  吸收谱线缓慢上升; 碱基 A 和稀有碱基  $m^1A$  也有很大差异, 稀有碱基  $m^1A$  有一个较为明显的特征峰, 且吸收系数明显

高于其他 6 种碱基; 7 种碱基在太赫兹特征峰和吸收强度上差异显著。表 1 列出了 7 种核酸碱基的太赫兹特征吸收峰。

表 1 7 种核酸碱基的太赫兹特征吸收峰 (THz)

Table 1 Terahertz characteristic absorption peaks of 7 nucleic acid bases (THz)

碱基	特征吸收峰 1	特征吸收峰 2	碱基	特征吸收峰 1	特征吸收峰 2
A	1.45	1.69	C	1.59	—
T	1.30	—	G	—	—
U	1.81	—	$m^5C$	—	—
$m^1A$	1.14	—			

### 2.2 碱基的拉曼光谱分析

图 5 为 DNA、RNA 碱基(胞嘧啶(C)、鸟嘌呤(G)、腺嘌呤(A)、胸腺嘧啶(T, DNA 专有)和尿嘧啶(U, RNA 专有)和稀有碱基(5-甲基胞嘧啶( $m^5C$ )、1-甲基腺嘌呤( $m^1A$ ))在 600~1750  $cm^{-1}$  的拉曼光谱。从图 5 可以看出, 7 种碱基也表现出很多明显的不同的拉曼谱峰, 但其特征峰杂而多, 较难直观识别。

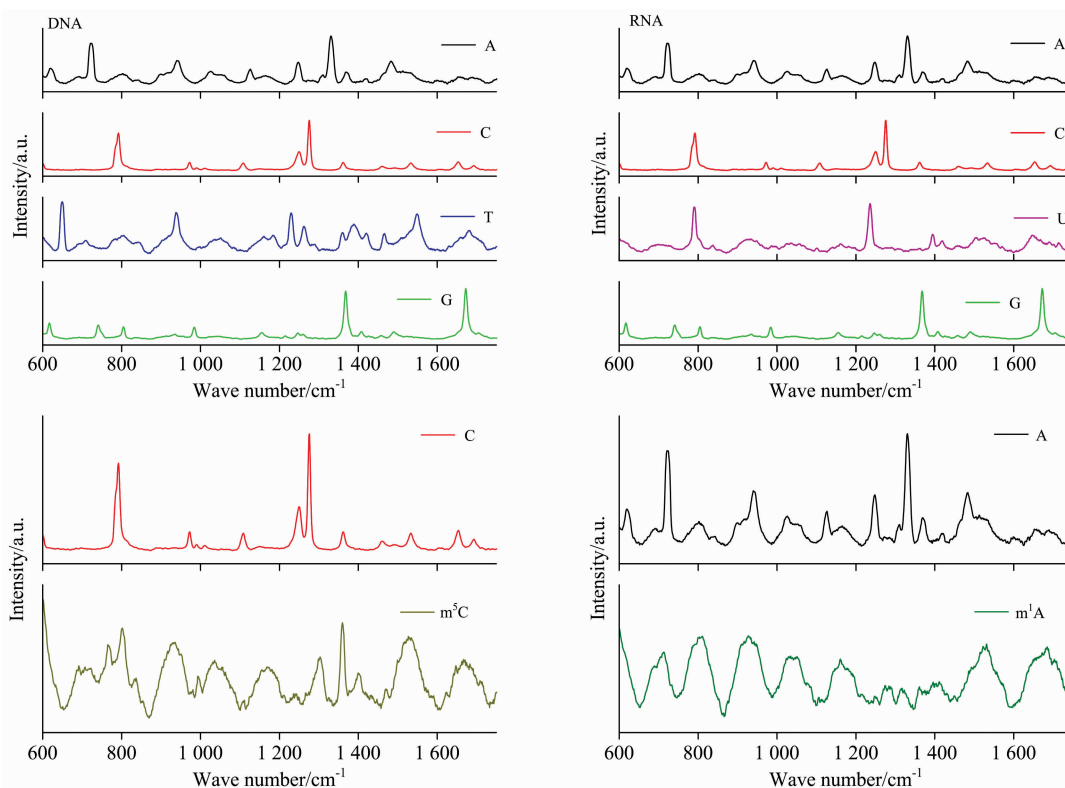


图 5 DNA、RNA 碱基(胞嘧啶(C)、鸟嘌呤(G)、腺嘌呤(A)、胸腺嘧啶(T, DNA 专有)和尿嘧啶(U, RNA 专有)和稀有碱基(5-甲基胞嘧啶( $m^5C$ )、1-甲基腺嘌呤( $m^1A$ ))在 600~1750  $cm^{-1}$  的拉曼光谱

Fig. 5 Raman spectra of DNA, RNA bases (cytosine (C), guanine (G), adenine (A), thymine (T, DNA-specific) and uracil (U, RNA-specific) and rare bases (5-methylcytosine ( $m^5C$ ), 1-methyladenine ( $m^1A$ )) from 600 to 1750  $cm^{-1}$

### 3 结 论

针对 DNA 和 RNA 碱基和稀有碱基样品,应用 THz-TDS 技术和拉曼光谱技术进行了对比研究。结果表明,在 THz-TDS 实验中,胞嘧啶(C)、鸟嘌呤(G)、腺嘌呤(A)、胸腺嘧啶(T, DNA 专有)和尿嘧啶(U, RNA 专有)和稀有碱基[5-甲基胞嘧啶( $m^5C$ )、1-甲基腺嘌呤( $m^1A$ )]在 0.2~2.0 THz 频段内特征吸收峰和吸收强度差异显著,可以直观地识

别出 7 种碱基的差异;拉曼光谱中,7 种碱基也表现出很多明显的不同特征峰,而正是由于拉曼光谱的特征峰杂而多,故不能直观地识别多种物质。且其吸收强度的差异与粉末的厚度,粒度和光聚焦的深度有关,样品的荧光还会给拉曼光谱带来干扰,同时激光还可能会损伤生物样品。这说明 THz-TDS 技术在识别这 7 种碱基的能力上优于拉曼光谱技术。THz-TDS 技术不仅为 DNA 和 RNA 碱基和稀有碱基的识别提供了一种快速和准确的检测方法,也为生物遗传学研究奠定实验基础。

### References

- [1] Wang F, Zhao D B, Dong H, et al. *Spectrochim. Acta A*, 2017, 179: 255.
- [2] Singh J S. *Spectrochim. Acta A*, 2015, 137: 625.
- [3] Wu L, Li F, Jin Z Y, et al. *Prog. Biochem. Biophys.*, 2016, 43: 281.
- [4] Bagraev N T, Chernev A L, Klyachkin L E, et al. *Semiconductors*, 2016, 50: 1208.
- [5] Tang M J, Huang Q, Wei D S, et al. *J. Biomed. Opt.*, 2015, 20: 095009.
- [6] Echchgadda I, Cerna C Z, Sloan M A, et al. *Proceedings of SPIE*, 2015, 9321: 93210Q.
- [7] Romanenko S, Begley R, Harvey A R, et al. *J. R. Soc. Interface*, 2017, 14: 20170585.
- [8] Titova L V, Ayesheshim A K, Golubov A, et al. *Biomed. Opt. Express*, 2013, 4: 559.
- [9] Yaekashiwa N, Yoshida H, Otsuki S, et al. *Photonics*, 2019, 6: 33.
- [10] Bogomazova A N, Vassina E M, Goryachkovskaya T N, et al. *Scientific Reports*, 2015, 5: 7749.
- [11] Cheon H, Paik J H, Choi M, et al. *Scientific Reports*, 2019, 9: 6413.
- [12] Son J H, Oh S J, Cheon H. *J. Appl. Phys.*, 2019, 125: 190901.
- [13] Cheon H, Yang H J, Lee S H, et al. *Scientific Reports*, 2016, 6: 37103.

## Detection of DNA, RNA Bases and Rare Bases by Terahertz Spectroscopy

TANG Ming-jie<sup>1, 2</sup>, YAN Shi-han<sup>1</sup>, ZHANG Ming-kun<sup>1</sup>, WEI Dong-shan<sup>1</sup>, DU Chun-lei<sup>1\*</sup>, CUI Hong-liang<sup>1</sup>

1. Chongqing Key Laboratory of Multi-Scale Manufacturing Technology, Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing 400714, China
2. University of Chinese Academy of Sciences, Beijing 100049, China

**Abstract** Spectral analysis of nucleic acid molecular bases is of great significance to the study of biogenetics. In this paper, the feasibility of terahertz time domain spectroscopy (THz-TDS) and Raman spectroscopy for the spectroscopic detection of DNA, RNA bases and rare bases is discussed. The terahertz and Raman spectra of seven solid nucleic acid bases were analyzed, and the THz-TDS and Raman spectroscopy techniques were compared. In THz-TDS experiments, cytosine (C), guanine (G), adenine (A), thymine (T, DNA-specific) and uracil (U, RNA-specific) and rare bases (5-methylcytosine ( $m^5C$ ), 1-methyladenine ( $m^1A$ )) were identified. Signature absorption peaks and absorption intensities differ significantly in the 0.2~2.0 THz, and the differences of seven bases can be identified intuitively; In Raman spectroscopy, the seven bases also show many distinct characteristic peaks, however the characteristic peaks of Raman spectroscopy are complex and many, so it is not intuitive to identify many substances. The difference of absorption intensity is related to the thickness of powder, particle size and the depth of light focusing. The fluorescence of samples also interfere with Raman spectra. The laser may also damage biological samples. The experimental results show that these two technologies can identify seven common and rare bases. THz-TDS technology is superior to Raman spectroscopy in the ability to identify these seven bases, showing a relatively concise, fast and non-destructive detection performance. THz-TDS technology not only provides a fast and accurate method for the identification of DNA, RNA and rare bases, but also lays an experimental foundation for the study of biogenetics.

**Keywords** Base; Rare base; Terahertz spectroscopy; Raman spectroscopy

\* Corresponding author

(Received Jul. 16, 2019; accepted Nov. 9, 2019)