

苏氨酸不同分子构型太赫兹吸收峰的量子化学指认

李 伟, 燕 芳*, 王志春, 刘成毫

内蒙古科技大学信息工程学院, 内蒙古 包头 014010

摘 要 与红外、紫外和拉曼光谱相比, 太赫兹光谱能量低, 在待测物质中不会出现有害光致电离现象, 伴随太赫兹技术的不断成熟, 太赫兹波已经成为常用的无损检测用波。很多生物大分子在高频光波探测下具有指纹性, 太赫兹时域光谱技术是对生物大分子无损检测的最佳手段。同时, 不同生物分子在太赫兹吸收谱中呈现出各不相同的吸收峰, 获得待测物质的太赫兹吸收谱后, 与标准谱进行对照可以为待测物质做出定性辨识。在此基础上, 结合最小二乘法、支持向量机等数据处理技术还可以实现基于太赫兹时域光谱对待测物质的定量分析。量子化学分析方法应用了量子力学的基本原理和方法, 其中电子分析理论从电子角度出发, 在分析大分子或原子个数众多的体系时近似误差较小, 并且密度泛函理论不依赖实验数据和先验知识的支撑。通过量子化学计算方法计算氨基酸分子的太赫兹吸收谱, 可以为氨基酸分子的太赫兹吸收峰匹配分子振动模式, 对氨基酸定性分析有一定的参考性, 并为实验获取的样品太赫兹时域光谱提供理论支撑, 在实验获得太赫兹吸收谱的基础上进行量子化学计算, 能验证实验结果的准确性。首先利用太赫兹时域光谱系统获取进口苏氨酸样品的太赫兹吸收谱, 其次分别构建苏氨酸样品在实物中以两性离子形式存在的单分子、二聚体和晶胞三种构型, 并利用量子化学计算方法完成了每种构型的结构优化, 最后计算三种苏氨酸分子构型的太赫兹吸收谱。结果表明, 单分子构型和二聚体构型的太赫兹计算谱与实验谱差异较大, 但在高频段计算谱与实验谱的吸收峰峰位基本吻合, 而较为全面反映分子间氢键及范德华力作用的晶胞构型计算谱与实验谱则较为吻合。同时表明, 与样品结构较为一致的、保持苏氨酸物理性质的最小结构为晶胞。

关键词 太赫兹吸收谱; 苏氨酸; 量子化学; 吸收峰

中图分类号: TP391.9 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)07-2054-05

引 言

太赫兹 (THz) 辐射指波长在 $30 \mu\text{m} \sim 3 \text{ mm}$ 之间的电磁波。太赫兹波段不仅是高效的信息载体, 且太赫兹波光子能量极低, 能穿透非金属和非极性材料, 因此在对物体内部缺陷的探测和对隐藏物的检测时具有很大的优势。目前, 太赫兹光谱技术在无损检测技术^[1]、生物医学研究、生物化学研究、安全检查、军事应用^[2-3]等领域都展现出巨大的潜力。许多生物分子的集体振转模式均位于太赫兹波段, 故不同生物分子在太赫兹吸收谱中呈现出不同的吸收峰, 即生物分子在太赫兹波段具有“指纹性”, 获得待测物质的太赫兹吸收谱后, 与标准指纹谱库进行对照可以实现对待测物质的定性分析。

同时含有氨基和羧基的有机化合物统称为氨基酸, 是构成蛋白质大分子的基础结构, 几乎一切生命活动都与之相关。因太赫兹波可以穿透氨基酸分子, 与可见光和红外光谱技术相比, 太赫兹散射较小, 并且不存在有害光致电离, 与紫外光谱技术相比, 不会引起待测样品变性, 能够安全无损地检测氨基酸样品。因此作为红外、紫外等光谱技术重要补充手段的太赫兹光谱技术可为氨基酸材料鉴别提供新方法。

1 实验部分

采用透射式太赫兹时域光谱系统进行苏氨酸样品实验谱的测定。图 1 为本实验用透射式太赫兹时域光谱系统原理图, 光谱仪及飞秒激光器等主要装置的性能参数如表 1 所示。

收稿日期: 2019-06-28, 修订日期: 2019-10-24

基金项目: 国家自然科学基金项目(61463041), 内蒙古自然科学基金项目(2018MS06020)资助

作者简介: 李 伟, 1992 年生, 内蒙古科技大学信息工程学院硕士研究生 e-mail: yjsumu@163.com

* 通讯联系人 e-mail: 0472yanfang@163.com

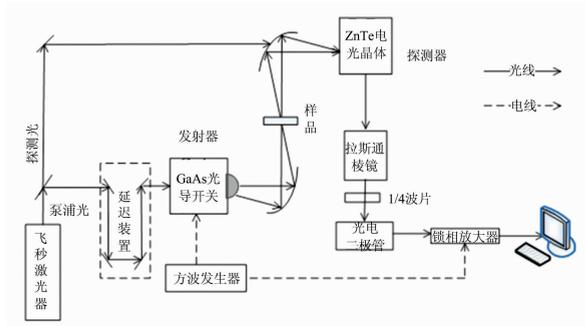


图 1 透射式太赫兹时域光谱系统原理图

Fig. 1 Principle diagram of transmission terahertz time-domain spectroscopy system

表 1 太赫兹时域光谱系统主要装置的性能参数
Table 1 Performance parameters of main devices in Terahertz time domain spectroscopy system

主要装置	品牌	性能参数
太赫兹时域光谱仪	EKASPLA	
飞秒激光器	Verdi-V5	脉冲宽度 10 fs, 重复频 75.42 MHz, 平均功率 300 mW
发射晶体		LT-GaAs
探测晶体		<110>碲化锌

为了减少空气中的水分对太赫兹波的吸收, 实验前在太赫兹光路中充入干燥的高纯度氮气, 确保光路密闭系统在实验时保持在室温 25 °C, 湿度控制在 4% 以下。

在实验过程中太赫兹波作用简图如图 2 所示, 太赫兹波垂直打在样品上, 其中黑色部分为样品。对参考波形和待测样品波形做傅立叶变换, 参考光谱和信号光谱依次是

$$E_{\text{ref}}(\omega) \equiv A_r(\omega) \exp[-i\phi_r(\omega)] \quad (1)$$

$$E_{\text{sam}}(\omega) \equiv A_s(\omega) \exp[-i\phi_s(\omega)] \quad (2)$$

表 2 样品配比表(湿度<4%, 温度 25 °C)

Table 2 Sample ratio table (Humidity <4%, Temperature 25 °C)

Z	编号	聚乙烯质量 /mg	样品质量 /mg	样片质量 /mg	样片直径 /mm	样片厚度 /mm	制片数量 /片
苏氨酸	1	118.25	30.41	120.58	13.46	0.98	1
	2	117.59	30.12	119.99	13.05	1.00	1
	3	118.01	30.33	121.42	13.31	1.03	1
	4	118.21	29.98	120.15	13.09	0.99	1

实验中, 为了减小太赫兹时域光谱系统测量样品信号和参考信号的误差, 每组样品测量三次后取其平均值, 由式(6)得到进口苏氨酸样品的太赫兹吸收谱如图 3 所示。

2 量子化学模拟计算

量子化学建模就是绘制分子结构, 分子构型可以在 GaussView 中直接绘制, 也可以在 ChemDraw 中通过构建分

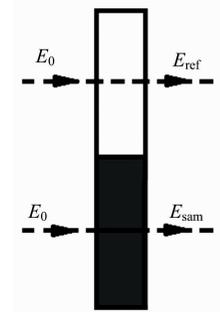


图 2 太赫兹波通过样片和空气的模型图

Fig. 2 Model of THz wave passing through the sample and air

式(1)和式(2)中 $A(\omega)$ 为太赫兹辐射的幅值, $\phi(\omega)$ 为太赫兹辐射的相位。

取 $E_{\text{ref}}(\omega)$ 和 $E_{\text{sam}}(\omega)$ 比值的幅频特性和相频特性[见式(3)]

$$\left| \frac{E_{\text{sam}}(\omega)}{E_{\text{ref}}(\omega)} \right| = \frac{4n}{(1+n)^2} \exp\left(-\frac{ad}{2}\right) \quad (3)$$

$$\phi(\omega) = -(n-1) \frac{\omega d}{c} \quad (4)$$

折射率和吸收系数等光学常数是表明待测样品宏观光学性质的基本物理量。由式(4)得折射率[见式(5)]

$$n(\omega) = 1 - \frac{c}{\omega d} \phi(\omega) \quad (5)$$

由折射率关于频率的波形得样品的吸收系数

$$\alpha(\omega) = -\frac{d}{2} \ln \left\{ \frac{|E_{\text{sam}}(\omega)|}{|E_{\text{ref}}(\omega)|} \frac{[n(\omega)+1]^2}{4n(\omega)} \right\} \quad (6)$$

式(6)中, $\alpha(\omega)$ 为吸收系数。

为了降低颗粒度差异引起的太赫兹波散射, 实验所用进口苏氨酸样品经充分研磨后加入聚乙烯粉末混合, 聚乙烯粉末颗粒直径约为 53~75 μm , 购买于美国 Sigma-Aldrich 公司, 以 10 MPa 压力冲压成片, 表 2 为样品配比表。

子式的方法转出分子构型。利用 GaussView 绘制分子构型需有原子坐标、键长、键角等先验知识, 而 ChemDraw 转出的分子构型直接具备原子坐标、键长、键角等信息。

结构优化在红外光谱、振动、偶极矩等分析中格外重要, 优化的目的是将绘制或编写的待计算结构的原子大小、键长、键角等更合理, 这种“合理”的标准实际上是要找到体系总能量最低的一种符合客观实际的待计算化学结构。

Gaussian03 和 MaterialsStudio 示出的红外吸收光谱为

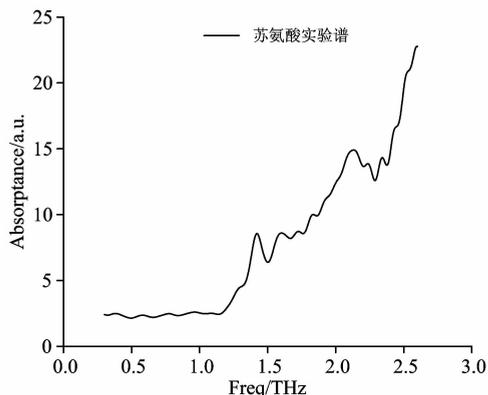


图 3 苏氨酸样品的太赫兹吸收谱

Fig. 3 THz absorption spectra of threonine samples

波数-能量($n_\lambda - E$)数据,波数(n_λ)为波长的倒数,为了使吸收谱曲线更为直观,通常将横轴的波数转换为频率,波长与频率之间的关系为

$$\lambda = \frac{c}{f} \quad (7)$$

式(7)中,取真空光速 $c = 2.998 \times 10^8 \text{ m} \cdot \text{s}^{-1}$,则频率为

$$f = 0.02998 \times n_\lambda \quad (8)$$

氨基酸在固体样品中以两性离子形式存在,为了模拟得到与实际样品最为吻合的结果,对苏氨酸单分子的两性离子进行计算分析。分子构型在 ChemDraw 中绘制,导出分子结构的笛卡尔坐标并导入 Gaussian03 进行结构优化和振动光谱计算:因两性离子带正负两性电荷,而选用 6-311G++(d,p) 基组,同时采用基于 B3LYP 杂化泛函的密度泛函理论做构型优化,最后进行红外振动及光谱(IR)计算。由于太赫兹波处于远红外波段,故太赫兹吸收光谱需在红外吸收光谱的远红外段提取。如图 4 所示为苏氨酸两性离子形式存在的单分子构型,右上角为两性离子形式的苏氨酸单分子结构式。

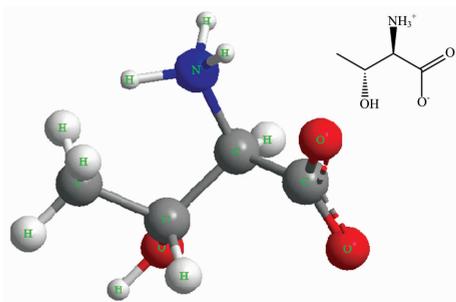


图 4 苏氨酸样品的单分子构型

Fig. 4 Single molecular configuration of threonine samples

聚合体理论指出:带强电正性粒子(如 N 原子和 N^{3+} 离子)易与带强电负性粒子(如 O 原子和 O^{2-} 离子)通过氢键作用形成聚合体。在苏氨酸中,羧基中氢氧根离子(OH^-)的氧原子带一个单位强电负性,而氨基中的氮原子带一个单位强电正性,故在氮氧之间易形成不稳定化学键——氢键,使苏氨酸的氨基酸官能团形成聚合体内环。特别是,在 n 个苏氨酸

酸分子中能形成 n 个氢键,并且 m 个苏氨酸分子形成的聚合体称为 m 聚苏氨酸。二聚体为苏氨酸聚合体的最小构型,计算二聚体构型的太赫兹吸收谱,能考虑分子间氢键作用对太赫兹的影响。图 5 为苏氨酸二聚体构型,右上角为两性离子形式苏氨酸二聚体结构式。

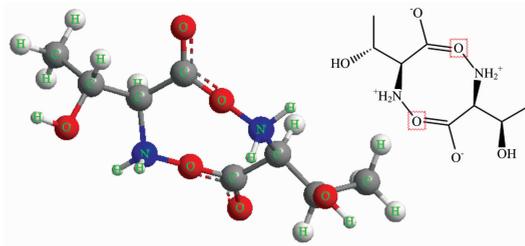


图 5 苏氨酸二聚体构型

Fig. 5 Threonine dimer configuration

一些物质保持其物理性质的是晶胞,晶胞由两个及两个以上构成该物质的同种分子构成。苏氨酸的物理性质由晶胞保持。通过英国剑桥晶胞数据库(CCDC)查得苏氨酸晶格参数为 $a = 13.630 \text{ \AA}$, $b = 7.753 \text{ \AA}$, $c = 5.162 \text{ \AA}$, $V = 545.486 \text{ \AA}^3$,晶胞需要在 Materials Studio 中的 CASETUP 模块中运算,CASTEUP 是基于固相平面波赝势密度泛函的计算方法。苏氨酸晶胞构型的太赫兹声子振动选取平面波赝势截断能 750 eV 进行结构优化,结果存在过度优化,进而选择 GGD 基组对 PBE 方法的改进方法 PBEsol 进行再优化,然后进行声子振动分析得到苏氨酸晶胞的太赫兹吸收谱。图 6 为苏氨酸晶胞结构,从图中可以看出,保持苏氨酸的物理性质最少需要四个苏氨酸分子。

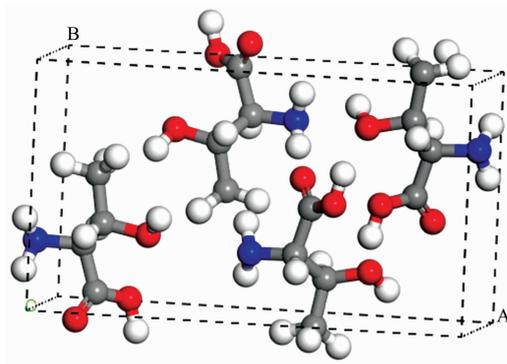


图 6 苏氨酸晶胞结构

Fig. 6 Threonine cell structure

上述苏氨酸单分子、二聚体、晶胞构型的量子化学计算均在配置有数据处理器 2.6 GHz 主频,单核四线程,缓存数据存储器 128G 固态硬盘(SSD)的计算机中。计算机时与收

表 3 计算机时及收敛迭代次数

Table 3 Computer time and convergence iteration number			
编号	名称	计算机时/h	收敛迭代次数/次
1	单分子	17	70
2	二聚体	42	115
3	晶胞	146	252

收敛迭代次数限于原子个数、分子空间结构和数据处理器配置等,表 3 为各结构的计算机时和收敛迭代次数。

3 结果与讨论

图 7 为苏氨酸单分子构型、二聚体构型计算谱与实验谱的对比,图中实线为实验谱,虚线为单分子构型计算谱,点划线为二聚体构型计算谱。

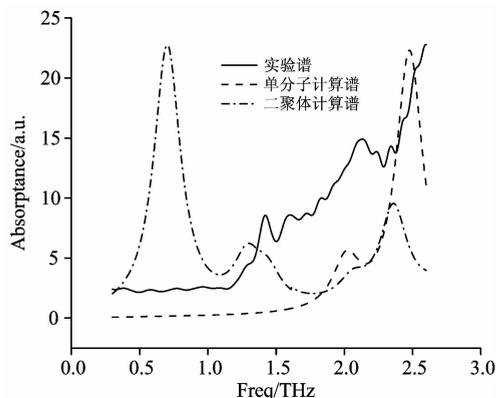


图 7 苏氨酸实验谱与单分子、二聚体计算谱的对比

Fig. 7 Comparison of threonine experimental spectrum with single molecule and dimer calculated spectrum

结果显示,考虑了分子间氢键作用的二聚体构型的太赫兹吸收谱已具备与实验谱一致的吸收峰个数,但除 2.2~2.6 THz 段外的峰位均有较大红移,同时与单分子构型计算谱相比,单分子构型计算谱的基线较好。为了研究能否获得吸收峰个数和峰位均基本一致的实验谱与计算谱,需要考虑分子间范德华力对太赫兹谱的影响。

图 8 为苏氨酸晶胞构型计算谱与实验谱的对比,图中实

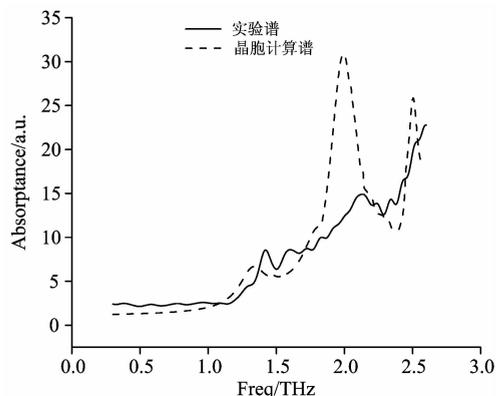


图 8 苏氨酸晶胞构型计算谱与实验谱的对比

Fig. 8 Comparison of calculated and experimental spectra of threonine cell configuration

线为实验谱,虚线为计算谱。

结果显示,苏氨酸晶胞构型计算谱无论是吸收峰个数还是峰位都与实验谱基本吻合,晶胞构型兼顾了分子间氢键作用与范德华力作用,在太赫兹辐射作用下也与实际样品结构较为吻合。表 4 列出了晶胞构型计算谱与实验谱的对比,计算谱与实验谱的相对误差均小于 10% (0.26 THz),“-”号表示计算谱结果较实验谱红移。

表 4 晶胞构型实验谱与计算谱的对比

Table 4 Comparison of experimental and computational spectra of cell configurations

项目	实验谱特征吸收峰 对应频率/THz	计算谱特征吸收峰 对应频率/THz	误差 /THz
1	1.42	1.34	-0.08
2	2.13	1.99	-0.14
3	2.60	2.50	-0.10

值得指明,计算谱吸收峰峰位存在明显红移,造成差异的原因归结为以下两点:

(1) 计算基组选择不同对体系总能的量子化学计算精度不同,并且无论选取何种基组进行计算,量子化学的计算法都存在近似误差;

(2) 结构优化的迭代收敛标准设定不同,优化后结构的无扰体系总能量不同。越严格的收敛标准能够得到越低的体系总能量,但迭代次数和需要的数据存储空间越大。

4 结论

以实验法获取了苏氨酸的太赫兹吸收谱,构建了不同的苏氨酸分子构型,以量子化学方法模拟计算不同分子构型的太赫兹吸收谱,并对计算谱进行了讨论。

研究表明,苏氨酸单分子构型的计算谱与实验谱差异较大,但单分子构型在高频段吸收峰峰位基本吻合。二聚体构型的计算谱虽然吸收峰个数与实验谱一致,但吸收峰位出现红移,说明考虑了分子间氢键的二聚体构型仍与实际样品结构不符。晶胞构型的计算谱无论是吸收峰个数还是吸收峰峰位均与实验谱较吻合,表明样品的太赫兹吸收谱是分子内原子及原子团作用、分子间氢键及范德华力作用的集中体现。

基于太赫兹时域光谱技术的氨基酸定性检测需要像“查字典”一样依赖于标准吸收谱,标准吸收谱收录于标准图库,这些吸收谱都是由实验法获得的,但实验谱的准确性依赖于实验仪器的精度、数据收集及处理的舍入误差等,基于量子化学模拟计算得到的计算谱可以为实验谱提供理论支持。而某些化学分析需要从微观入手,这需要找到保持待分析物质理化性质的最小结构,用实验谱指认计算谱,可以得到保持苏氨酸最小物理性质的构型为晶胞。

References

- [1] Kaori Fukunaga, Yuichi Ogawa, Yuichi Ogawa, et al. *IEICE Electronics Express*, 2007, 4(8): 258.
- [2] Peiponen K E, Zeitler A, Kuwata Gonokamim. *Terahertz Spectroscopy and Imaging*. Atlanta: Springer Series in Optical Sciences, 2003.
- [3] WANG Feng-xia, ZHANG Zhuo-yong, ZHANG Cun-lin(王凤霞, 张卓勇, 张存林). *Chinese Journal of Analytical Chemistry(分析化学)*, 2006, (4): 576.

Quantum Chemical Identification of Terahertz Absorption Peaks of Threonine with Different Molecular Configurations

LI Wei, YAN Fang*, WANG Zhi-chun, LIU Cheng-hao

Information Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, China

Abstract Compared with infrared, ultraviolet and Raman spectra, terahertz spectra have low energy and no harmful photoionization phenomena in the substances to be measured. With the maturation of terahertz technology, the terahertz wave has become a common wave for non-destructive testing. Many biological macromolecules have fingerprints under high-frequency light detection. THz time-domain spectroscopy is the best method for nondestructive detection of biological macromolecules. At the same time, different biological molecules show different absorption peaks in Terahertz Absorption spectra. After obtaining Terahertz Absorption Spectra of the substance to be measured, compared with standard spectra, qualitative identification of the substance to be measured can be made. On this basis, combined with data processing techniques such as least squares method and support vector machine, the quantitative analysis of measured substances based on terahertz time-domain spectroscopy can also be realized. The basic principles and methods of quantum mechanics are applied in the quantum chemical analysis method. From the electronic point of view, the approximation error of the electronic analysis theory in the analysis of systems with large molecules or atoms is small, and the density functional theory does not depend on the support of experimental data and prior knowledge. The Terahertz Absorption Spectra of amino acids can be calculated by a quantum chemistry calculation method, which can match the molecular vibration mode of terahertz absorption peaks of amino acids, provide certain reference and directivity for qualitative analysis of amino acids, and provide theoretical support for terahertz time domain spectra of samples obtained from experiments. Quantum chemistry calculation is carried out on the basis of the terahertz absorption spectra obtained from experiments. It can further verify the accuracy of the experimental results. In this paper, the Terahertz Absorption Spectra of imported threonine samples were obtained by the terahertz time domain spectroscopy system. Then, three configurations of threonine samples in the form of zwitterionic ions were constructed, and the structure optimization of each configuration was completed by quantum chemical calculation method. Finally, the Terahertz Absorption Spectra of three threonine molecular configurations were calculated. The results show that the terahertz calculation spectra of the monomer and dimer configurations are quite different from the experimental spectra, but in the high frequency band, the absorption peaks of the calculated spectra are basically in agreement with the experimental spectra, while the lattice configuration calculation spectra of the more comprehensive reaction of intermolecular hydrogen bond and van der Waals force are in good agreement with the experimental spectra. At the same time, the smallest structure, which is consistent with the sample structure and keeps the physical properties of threonine is the cell.

Keywords Terahertz absorption spectrum; Threonine; Quantum chemistry; Absorption peak

(Received Jun. 28, 2019; accepted Oct. 24, 2019)

* Corresponding author