

# RNA 组装体对太赫兹外场的频率选择性响应\*

张琴<sup>1)</sup> 王凯程<sup>1)</sup> 孙远昆<sup>1)</sup> 王少萌<sup>1)</sup> 宫玉彬<sup>1†)</sup>

1) 电子科技大学电子科学与工程学院；真空电子器件技术全国重点实验室；四川省太赫兹辐射与应用重点实验室，成都 611731)

## 摘 要

GGGGCC (G4C2) 六核苷酸重复序列的异常扩增是肌萎缩侧索硬化症和额颞叶痴呆最常见的遗传病因。该序列转录出的 RNA 可通过多重碱基配对自组装为多链四级结构，这种自组装行为是其发生液-液相分离、形成病理性凝聚体的分子基础。本文以 5xG4C2 RNA 组装体为模型，应用分子动力学模拟，系统考察了 0.5 V/nm 太赫兹 (Terahertz, THz) 波扰动对链间相互作用与溶剂界面耦合关系的差异性影响。结果表明，不同频率选择性揭示了组装体在频率维度上的本征响应：3.0 THz 和 8.0 THz 诱导了全局力学紧致化，溶剂可及表面积 (SASA) 显著下降而链间接触的置换效率未变；11.9 THz 处链间接触的 SASA 掩埋和水分子置换效率同时显著提高，呈现协同开关特征；20.5 THz 和 37.3 THz 则分别选择性增强了链间紧致度和氢键驱动的水分子释放通道。所有频率下组装体的链间氢键网络和主链堆积均保持完整。上述结果表明，不同的太赫兹频率可选择性揭示 RNA 组装体在不同物理维度上的本征响应特征。

**关键词：**GGGGCC RNA 重复序列，太赫兹波，分子动力学模拟，RNA 组装体

**PACS:** 87.14.gn, 87.15.ap, 87.50.uj, 87.15.hg

**基金：**国家自然科学基金（批准号：62527804, T2241002, 62427806, 62588201）资助的课题。

† 通信作者.E-mail: ybgong@uestc.edu.cn

第一作者.E-mail: 202111022406@std.uestc.edu.cn

## 1 引言

GGGGCC (G4C2) 六核苷酸重复序列在 C9orf72 基因非编码区的异常扩增是肌萎缩侧索硬化症 (amyotrophic lateral sclerosis, ALS) 和额颞叶痴呆 (frontotemporal dementia, FTD) 最常见的遗传病因<sup>[1-4]</sup>，其转录出的 RNA 具有独特的物理化学性质和自组装能力。在分子结构层面，Reddy 等通过紫外熔解和圆二色谱证明， $r(\text{GGGGCC})_n$  可形成热稳定性极高的平行 G-四链体，且既可为单分

子内结构，也可在较高浓度或较长重复数下形成分子间四链体<sup>[5]</sup>。在物理化学层面，Jain 和 Vale 表明，重复 RNA 在体外存在一个临界重复数阈值，超过该阈值后发生溶胶-凝胶相变形成微米级凝聚体，且该阈值与患者中的病理重复数高度吻合<sup>[6]</sup>。在细胞层面，Haeusler 等进一步发现 G4C2 重复序列的结构多态性可触发转录异常、核仁蛋白错误定位和核糖体 RNA 加工受损等一系列致病分子级联<sup>[7]</sup>。上述研究从 G-四链体结构、物理相变到细胞内致病级联三个递进层次共同揭示：G4C2 RNA 具有通过 G-四链体进行多链自组装的固有物理倾向，而这种自组装行为构成其致病功能的分子基础。

然而，现有研究主要集中在 G4C2 体系的相行为层面<sup>[6, 8, 9]</sup>，即 RNA 在给定条件下凝聚还是分散，而对于组装体在已形成稳定四级拓扑的前提下，其内部相互作用与外部溶剂响应之间的耦合关系尚缺乏系统研究。澄清这一关系，将有助于从分子层级理解组装体自身的基本物理性质。

太赫兹电磁波因其独特的频段位置，为解析这类耦合关系提供了天然工具。太赫兹光子能量横跨从弱非键相互作用（氢键、范德华力、 $\pi$  堆积）到骨架局域振动（磷酸基团伸缩、核糖环呼吸）的多个能量尺度<sup>[10-13]</sup>，使不同频率可能选择性地耦合到核酸大分子体系的不同分子自由度。根据 Heyden 等的第一性原理分子动力学模拟，液态水的太赫兹吸收具有显著的空间分辨特征，~2.4 THz (~80  $\text{cm}^{-1}$ ) 处的吸收由第二溶剂层的氢键四面体伞状集体运动主导，~6 THz (~200  $\text{cm}^{-1}$ ) 处则由第一溶剂层的氢键伸缩振动贡献<sup>[14]</sup>。Xu 和 Havenith 进一步指出，太赫兹光谱可探测界面水的低频集体动力学，而在生物分子识别过程中，这种动力学的扰动程度与生物活性直接相关<sup>[15]</sup>。这些研究表明，太赫兹波并非单一的调控"开关"，而是一个可连续调谐的物理变量——不同频率或许能选择性地耦合到体系中不同的分子自由度和空间尺度，从而揭示体系在不同物理维度上的固有响

应特征。然而，目前太赫兹频率外场作用下系统解析核酸组装体链间-界面耦合关系的研究尚属空白。

此前我们采用 Martini 3 粗粒度分子动力学模拟，在凝聚体层级研究了 12 条 5×G4C2 RNA 在 10 THz 和 37.3 THz 振荡场下的相行为，发现太赫兹场可通过调控抗衡离子分布、并间接影响水合层动力学，从而促进链间接触<sup>[8]</sup>。然而，粗粒化模型存在固有局限，无法分辨碱基层面的氢键类型、核糖构象涨落以及第一溶剂化层水分子与 RNA 表面之间的特异性相互作用。而这些正是全原子分辨率下频率差异可能编码的分子物理信息。为兼顾统计功效与计算可行性并与前序粗粒化研究序列保持一致<sup>[8]</sup>，本文以完整四级拓扑的五链 5×G4C2 RNA 组装体为模拟体系，在 3-37.3 THz 范围内选取 7 个离散频率（3.0 THz、8.0 THz、11.9 THz、20.5 THz、25.5 THz、32.2 THz 和 37.3 THz），进行全原子显式溶剂分子动力学模拟，旨在全原子分辨率下揭示粗粒化模型无法分辨的问题：不同的太赫兹外场，是否选择性地作用于组装体的不同物理维度，链间接触的置换效率、SASA 的紧致化响应幅度以及界面水合密度，从而使太赫兹频率成为解析 RNA 组装体链间-界面耦合关系的一个可调物理变量。

## 2 方法

### 2.1 模拟体系与太赫兹辐照方案

本研究以五链 5×G4C2 重复序列 RNA 组装体为模型体系，采用 GROMACS 2020.3 软件进行全原子显式溶剂分子动力学模拟<sup>[16]</sup>。模拟体系仅含 145 个 K<sup>+</sup> 用于中和 RNA 磷酸骨架的负电荷，未额外添加盐，模拟盒尺寸约 10.1×10.1×10.1 nm<sup>3</sup>，包含约 10 万个原子，RNA 分子内及分子间相互作用采用 Amber14SB-OL15 力场结合  $\chi$ OL3 RNA 补丁描述<sup>[17, 18]</sup>。模拟在三维周期性边界条件（PBC）下进行—长程静电相互作用采用 PME 方法处理，短程范德华相互作用采用截断方案，二者截断

半径均设为1.0 nm，并对范德华项施加长程色散校正（EnerPres）。温度控制采用V-rescale恒温器<sup>[19]</sup>，分别耦合溶剂（SOL，TIP3P水分子）、离子（Ion，K<sup>+</sup>）及RNA分子，参考温度298.15 K，耦合时间常数0.002 ps；压力控制采用Parrinello-Rahman恒压器<sup>[20, 21]</sup>，各向同性耦合，参考压力1 bar，耦合时间常数0.5 ps。含氢键的键长采用LINCS算法约束。模拟时间步长为2 fs。

已有研究计算表明当太赫兹电磁波与带电粒子相互作用时，其磁场分量比电场分量小几个数量级，而电场起着主导作用<sup>[22, 23]</sup>，在本文分子动力学模拟中，对每个原子仅施加电场力，不含洛伦兹力。因此，以下讨论均基于纯电场模拟。太赫兹电场以沿y方向的正弦电场形式施加： $E_y(t) = E_0(2\pi ft + \phi_0)$ ，场强幅值 $E_0=0.5$  V/nm，初始相位 $\phi_0 = 0$ 。选取8个频率：0.0 THz（对照组，无外场）、3.0 THz、8.0 THz、11.9 THz、20.5 THz、25.5 THz、32.2 THz和37.3 THz。各频率条件下独立运行20个副本，初始速度由Maxwell-Boltzmann分布独立采样，每个副本模拟时长200 ns，分析均取平衡后轨迹。

我们太赫兹扰动频率的选择依据如下：3.0-8.0 THz（ $\sim 100$ -267  $\text{cm}^{-1}$ ）位于核酸大分子由弱非键相互作用和碱基对间氢键集体运动所主导的低频振动区间<sup>[24, 25]</sup>。11.9 THz（ $\sim 397 \text{ cm}^{-1}$ ）精准落入核糖环的集体呼吸振动频段<sup>[26]</sup>。20.5-32.2 THz（ $\sim 684$ -1074  $\text{cm}^{-1}$ ）覆盖了从核糖-磷酸耦合振动到磷酸对称伸缩振动的骨架振动过渡区<sup>[13]</sup>。37.3 THz（ $\sim 1244 \text{ cm}^{-1}$ ）对应于磷酸基团的反对称 $\text{PO}_2^-$ 伸缩振动<sup>[13]</sup>。上述频率梯度覆盖了从核酸螺旋集体运动、核糖环呼吸、磷酸骨架振动到局域化学键伸缩的多个分子自由度层级，使得系统考察不同弱相互作用成份对RNA组装体的差异化贡献成为可能。

## 2.2 表征参量的定义与物理含义

为定量刻画太赫兹电场对RNA四级结构的调控效应，本文定义了四个表征参

量，分别从链间相互作用和链-溶剂界面两个正交维度描述组装体的结构响应。组装体的四级结构及各参量的可视化对应关系见图S1。

(1) 链间总氢键数 $n_H$ 。 $n_H$ 定义为不同RNA链之间所形成的氢键总数，反映链间直接结构接触的程度。 $n_H$ 越大，四级结构的链间耦合越紧密，组装体的拓扑刚性越强。计算上，使用MDAnalysis v2.0分析各副本最后2500帧（平衡阶段）轨迹<sup>[27]</sup>，氢键判据为供体-受体距离 $d \leq 3.5 \text{ \AA}$ 且供体-氢-受体角度 $\theta \geq 150^\circ$ 。供体和受体限定于Watson-Crick (G-C) 配对和G-四链体 (G-G) 配对涉及的原子（典型WC配对见图S1d），每条氢键严格分类为WC型、G-四链体Hoogsteen型或其他。对每帧求和后取帧平均，再跨20个副本求均值。

(2) 链间紧致因子 $CCR$  (compactness factor,  $CCR$ )<sup>[28]</sup>。 $CCR$ 定义为不同链间磷 (P) 原子空间靠近程度的加权度量，反映组装体四级结构的几何紧致度。 $CCR$ 越大，骨架越紧凑，链间磷 (P) 原子的平均间距越小。计算上，提取G和C残基的磷原子坐标（见图S1b），对每帧按下式计算不同链间的紧致度之和：

$$CCR = \sum_{chain(i) \neq chain(j)} \frac{1}{1 + d_{ij}} \quad (1)$$

其中 $d_{ij}$ 为P原子 $i$ 和 $j$ 在周期性边界条件下的最小镜像距离，截断半径设为17.56 nm。副本水平的均值由最后2500帧取平均得到。

(3) 溶剂可及表面积 ( $SASA$ )<sup>[29]</sup>。 $SASA$ 定义为RNA全原子范德华表面可被溶剂探针触及的面积，反映组装体暴露于体相溶剂的程度。 $SASA$ 越小，链间堆积越紧密，四级结构整体越紧致。计算上，使用GROMACS sasa工具（图S1c分子表面示意），以半径0.14 nm的球形探针（对应于水分子有效半径）沿RNA全原子范德华表面滚动，逐帧计算各副本最后2500帧，以150核苷酸归一化得到每核苷酸 $SASA$ （单位： $\text{nm}^2/\text{nt}$ ）。

(4) RNA-溶剂氢键数 $n_W$ 。 $n_W$ 定义为RNA与水分子之间形成的氢键总数，反

映组装体界面的水合程度。 $n_w$ 越大，界面水分子越多，水合层越厚。计算上，使用GROMACS gmx hbond工具提取每帧RNA-水分子间的氢键总数，截断判据与 $n_H$ 相同（ $d \leq 3.5 \text{ \AA}$ ， $\theta \geq 150^\circ$ ），以150核苷酸归一化得到每核苷酸氢键数，副本水平的均值取自最后2500帧。

上述四个表征参量分属两个正交维度：（1） $n_H$ 和 $CCR$ 量化链间耦合的强度与几何紧致度，二者合起来表征四级结构内部接触网络的完整性和紧凑性；（2） $SASA$ 和 $n_w$ 量化组装体暴露于溶剂的程度与界面水合状态，二者合起来刻画链-溶剂界面的几何和化学特征。这一分组为后续通过线性置换模型分析链间接触维度与溶剂界面维度之间的统计耦合关系奠定了基础。

### 2.3 线性置换模型与统计显著性检验

为定量揭示链间氢键 $n_H$ 和链间紧致度 $CCR$ 在统计上如何耦合到溶剂可及表面积 $SASA$ 和界面水分子数 $n_w$ 的变化，在每个频率下对副本数据拟合如下三个线性置换模型：

$$\text{模型1: } SASA = k_H \cdot n_H + k_C \cdot CCR + const \quad (2)$$

$$\text{模型2: } n_w = r_H \cdot n_H + r_C \cdot CCR + const \quad (3)$$

$$\text{模型3: } n_w = \rho_{HB} \cdot SASA + const \quad (4)$$

在模型1中， $k_H$ 和 $k_C$ 分别度量单位 $n_H$ 和单位 $CCR$ 变化所引起的 $SASA$ 变化，反映链间相互作用置换溶剂可及表面积的效率。负值表示链间接触增强导致表面掩埋，正值表示链间接触增强反而暴露更多表面。在模型2中， $r_H$ 和 $r_C$ 分别度量单位 $n_H$ 和单位 $CCR$ 变化所释放（负值）或招募（正值）的界面水分子数，反映链间接触对界面水分子的置换能力。在模型3中， $\rho_{HB}$ 度量单位 $SASA$ 变化所能容纳的水合氢键数目，即界面水合密度。 $\rho_{HB}$ 可被视为RNA表面亲水性的一个可操作度量，若 $\rho_{HB}$ 在不同频率下保持恒定，则表明太赫兹电场并未改变RNA表面单位面积对

水分子的化学亲和力。

所有模型系数均通过最小二乘拟合求解，以决定系数 $R^2$ 衡量拟合优度。然而，最小二乘仅给出基于现有20个副本的点估计，无法回答这些系数在不同抽样下的波动幅度以及处理组与对照组的差异是否统计可靠。为此，对每个频率的副本数据进行1000次有放回Bootstrap重抽样，获得各系数的经验分布：以Bootstrap分布的2.5%和97.5%分位数分别作为95%置信区间的下、上限。对于各处理频率（3.0-37.3 THz，共7个）与对照组（0 THz）系数的比较，计算两组Bootstrap分布差异的双侧 $p$ 值，即构建处理组系数与对照组系数的差异分布（1000个差异值），取差异 $\leq 0$ 和 $\geq 0$ 的比例中较小者的2倍。与经典 $t$ 检验假设系数服从正态分布不同，Bootstrap直接从数据中通过重抽样构建系数的经验分布，无需任何分布假设，尤其适用于本文 $n=20$ 的小样本情形。用以下方式描述算法：

$$\text{Bootstrap估计量: } \hat{\theta}_b^* = \hat{\theta}(X_b^*), \quad b=1, \dots, B \quad (5)$$

其中：

$X_b^*$ 是从原始样本  $X = \{x_1, \dots, x_n\}$  中有放回抽取的大小为 $n$ 的子样本；

$\hat{\theta}(\cdot)$ 是任何统计量（这里是普通最小二乘法（OLS）回归系数）；

$B$ 是Bootstrap次数，这里是1000。

置信区间（百分位法）：

$$CI_{95\%} = \left[ \hat{\theta}_{(0.025)}^*, \hat{\theta}_{(0.975)}^* \right] \quad (6)$$

差异 $p$ 值：

$$p = 2 \times \min \left( \frac{1}{B} \sum_{b=1}^B 1(\Delta_b^* \leq 0), \frac{1}{B} \sum_{b=1}^B 1(\Delta_b^* \geq 0) \right) \quad (7)$$

其中  $\Delta_b^* = \hat{\theta}_{b,treat}^* - \hat{\theta}_{b,ctrl}^*$  是第 $b$ 次两个组的系数差异。

此外，对每个处理频率各表征参量的副本均值与对照组进行双样本Student  $t$

检验（假设等方差），给出各频率下各参量的均值、标准差（SD）、标准误（SEM）及t统计量和p值。所有统计检验包括Bootstrap的显著性水平定为 $\alpha=0.05$ ，图中星号标注规则为 $*p<0.05$ ， $**p<0.01$ ， $***p<0.001$ 。

### 3 结果与讨论

#### 3.1 太赫兹频率对四级结构表征参量的直接影响

首先考察太赫兹电场对 RNA 链间相互作用的直接效应。在所考察的 3-37.3 THz 范围内，各处理组的链间氢键数  $n_H$ （处理组范围 8.48-11.57，对照组  $10.31\pm5.25$ ）和链间紧致因子  $CCR$ （处理组范围 6.68-7.12，对照组  $6.65\pm0.99$ ）与对照组均无显著差异（t 检验， $p$  均 $>0.05$ ，完整数据见补充材料表 S1）。所有处理组的  $n_H$  和  $CCR$  均值均落在对照组的一个标准差范围之内，且  $n_H$  的符号方向在不同频率间并无系统性偏移。该结果表明，0.5 V/nm 场强下的太赫兹电场不足以破坏 G-四链体和 Watson-Crick 配对所提供的主链堆积与链间氢键网络，RNA 组装体的四级拓扑结构在整个 3-37.3 THz 频段内保持完整。因此，下文将讨论的  $SASA$  和  $n_W$  的任何显著变化，均发生在链间接触不被破坏的背景之下。

#### 3.2 溶剂可及表面积与水合数的频率选择性响应

与链间氢键  $n_H$  和  $CCR$  的惰性不同，溶剂可及表面积  $SASA$  和 RNA-溶剂氢键数  $n_W$  表现出显著且具有频率选择性的响应（图 1，表 1）。11.9-20.5 THz 区间采样间隔较宽，该区间内是否存在  $SASA$  或耦合系数显著变化的中间频率尚无法基于当前数据排除。

表 1 各太赫兹频率下  $SASA$  和  $n_W$  的均值变化

Table 1. Mean values of  $SASA$  and  $n_W$  at each terahertz frequency.

| 频率 (THz) | $SASA$ (nm <sup>2</sup> /nt) | $\Delta SASA$ (%) | $p$ ( $SASA$ ) | $n_W$ | $p$ ( $n_W$ ) |
|----------|------------------------------|-------------------|----------------|-------|---------------|
| 对照       | 1.763                        |                   |                | 10.06 |               |
| 3.0      | 1.696                        | -3.8              | <0.001***      | 9.86  | <0.001***     |

|      |       |      |         |       |       |
|------|-------|------|---------|-------|-------|
| 8.0  | 1.724 | -2.2 | 0.012*  | 10.00 | 0.226 |
| 11.9 | 1.728 | -2.0 | 0.028*  | 9.97  | 0.076 |
| 32.2 | 1.722 | -2.3 | 0.007** | 10.00 | 0.193 |
| 37.3 | 1.721 | -2.4 | 0.012*  | 9.98  | 0.087 |

注：仅列出 *SASA* 显著下降的频率 ( $p < 0.05$ )，20.5 THz 和 25.5 THz ( $p > 0.05$ ) 省略。完整处理组频率数据见补充材料表 S1。

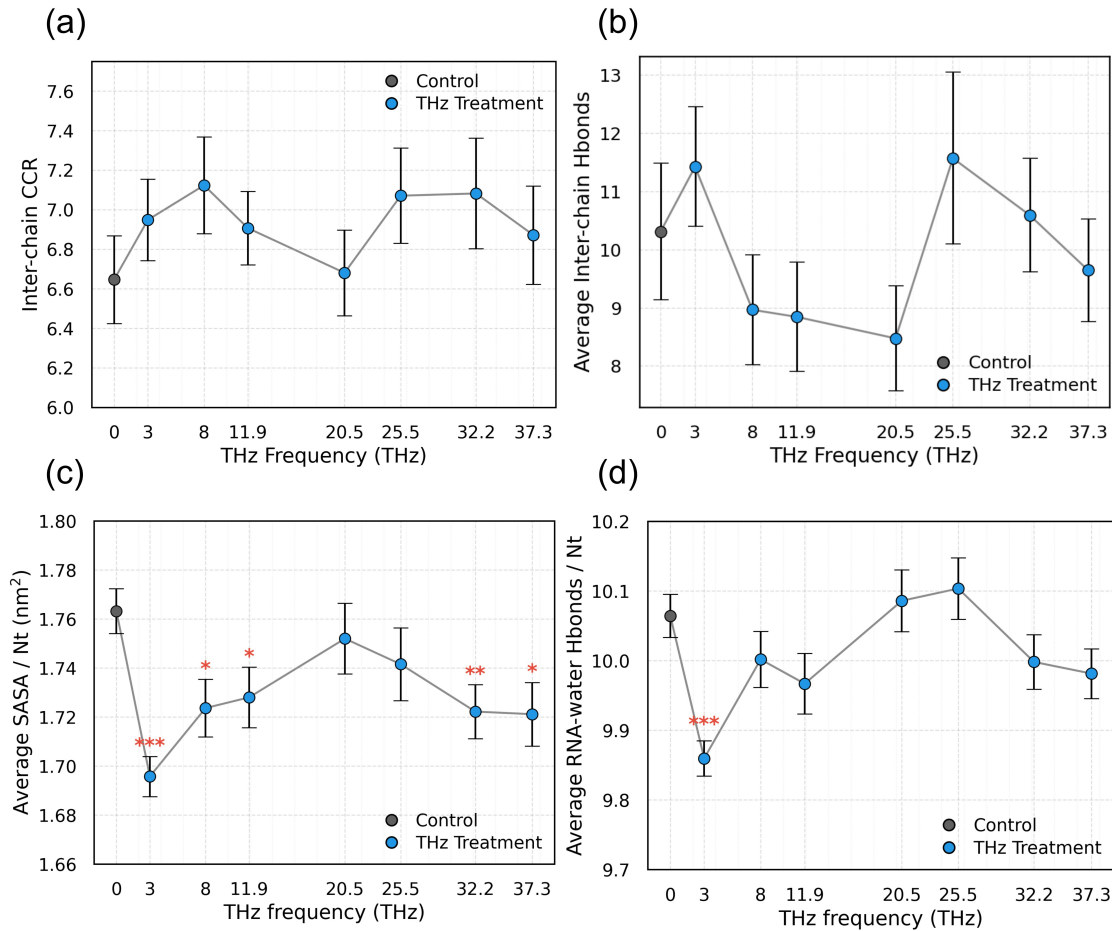


图 1 各表征参量随 THz 频率的变化趋势。(a) 平均链间紧致因子 (*CCR*)。 (b) 链间平均每核苷酸 (Nucleotide, Nt) RNA-RNA 氢键数目。(c) 平均每核苷酸溶剂可及表面积 (*SASA*)。 (d) 平均每核苷酸 RNA-水氢键数目。

Fig. 1. Variation trends of order parameters with THz frequency: (a) Average inter-chain compactness factor (*CCR*); (b) Average number of inter-chain RNA-RNA hydrogen bonds (hbonds) per nucleotide (Nt); (c) A solvent-accessible surface area (*SASA*) per Nt; (d) Average number of RNA-water hydrogen bonds per Nt.

在 8 个频率中，*SASA* 在以下 5 个频率显著低于对照组 ( $p < 0.05$ )：3.0 THz、8.0 THz、11.9 THz、32.2 THz 和 37.3 THz，降幅范围为 2.0%~3.8% (如表 1)。其

中 3.0 THz 的降幅最大 ( $\Delta SASA = -0.067 \text{ nm}^2/\text{nt}$ , 降幅约 3.8%,  $p = 3 \times 10^{-6}$ ), 表明低频段远红外激发对 RNA 四级结构产生了最强紧致化效应。值得注意的是, 20.5 THz 和 25.5 THz 处的 *SASA* 与对照组无显著差异 ( $p = 0.516$  和  $p = 0.222$ ), 构成一个松弛窗口, RNA 的表面暴露程度恢复至对照水平。

RNA-溶剂氢键数  $n_w$  的响应比 *SASA* 更为保守: 仅在 3.0 THz 处观察到  $n_w$  的显著下降 (9.860 vs 对照 10.064,  $p = 9 \times 10^{-6}$ )。在 *SASA* 显著下降的其他 4 个频率 (8.0 THz、11.9 THz、32.2 THz、37.3 THz),  $n_w$  与对照无显著差异 ( $p = 0.076\text{--}0.226$ , 完整数据见表 S1)。这一 *SASA* 减小而  $n_w$  不降的分离现象表明, 太赫兹电场诱导的紧致化并非仅通过直接排除界面水分子实现, 可能还伴随着界面水合状态的补偿性调整。

### 3.3 链间接触对表面掩埋与水分子释放的置换效率

为揭示紧致化的分子机制, 我们以各频率的副本数据拟合线性置换模型, 考察  $n_H$  和 *CCR* 对 *SASA* 和  $n_w$  的统计驱动效率 (表 2, 公式 (2) 和 (3))。

在对照组 (0 THz) 中, 模型 1 ( $SASA \sim n_H + CCR$ ) 的决定系数  $R^2 = 0.531$ 。*CCR* 与 *SASA* 呈预期的负相关 ( $k_C = -1.95 \times 10^{-2}$ ), 即链间紧致导致部分表面掩埋; 而  $n_H$  与 *SASA* 呈弱正相关 ( $k_H = +3.83 \times 10^{-3}$ ), 表明在控制 *CCR* 不变的前提下, 链间氢键的形成伴随 *SASA* 的微弱增加。

表 2 关键频率下线性置换模型的回归系数

Table 2. Regression coefficients of the linear displacement model at key frequencies.

| 频率 (THz) | $k_H (\times 10^{-3})$ | $k_C (\times 10^{-2})$ | $R^2_1$ | $r_H (\times 10^{-2})$ | $r_C (\times 10^{-1})$ | $R^2_2$ |
|----------|------------------------|------------------------|---------|------------------------|------------------------|---------|
| 对照       | +3.8                   | -1.95                  | 0.531   | +0.92                  | +0.001                 | 0.121   |
| 11.9     | -4.0                   | -4.50                  | 0.676   | -2.31                  | -1.10                  | 0.616   |
| 20.5     | +3.9                   | -4.83                  | 0.493   | +0.97                  | -1.70                  | 0.648   |
| 32.2     | -0.2                   | -3.05                  | 0.598   | -0.85                  | -0.62                  | 0.236   |
| 37.3     | -3.3                   | -2.48                  | 0.340   | -1.49                  | -0.61                  | 0.403   |

注：粗体表示 Bootstrap 差异检验  $p < 0.05$ （与对照组相比）。完整处理组频率数据见补充材料表 S2。

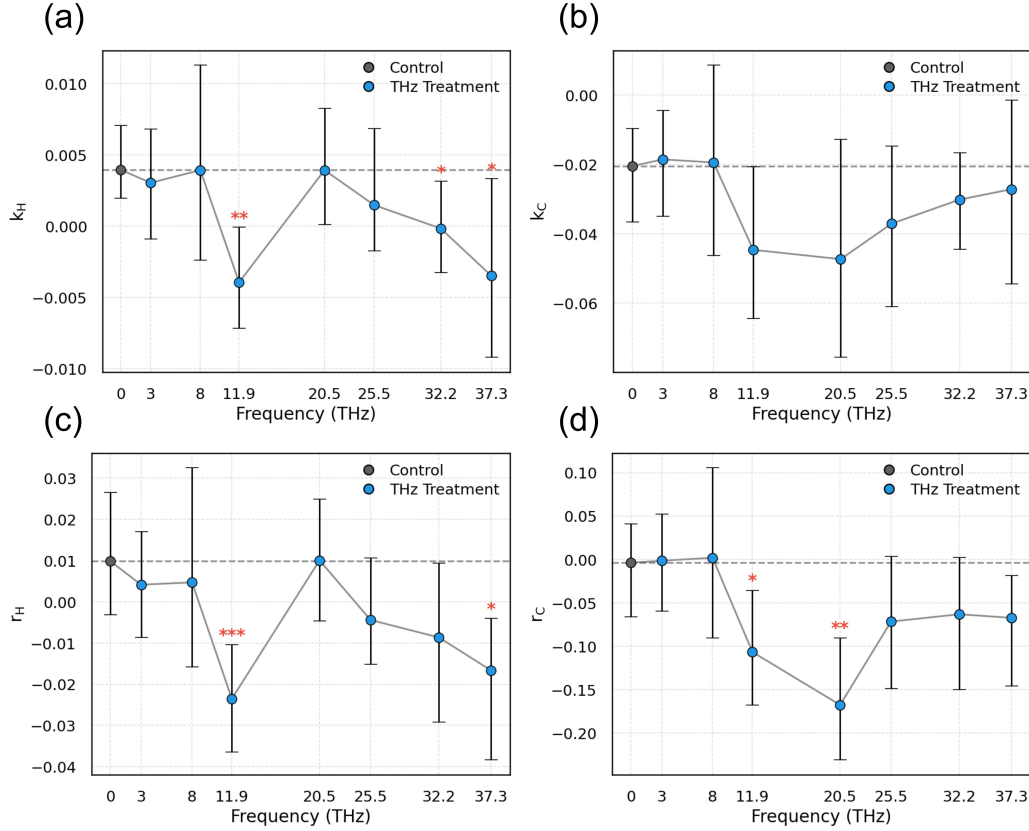


图 2 模型 1 和 2 中线性置换系数的 Bootstrap 估计值随太赫兹频率的变化, 对应表 S3。(a)  $k_H$ , 链间氢键对 SASA 的置换系数; (b)  $k_C$ , 链间紧致度对 SASA 的置换系数, 负值和正值分别表示增强表面掩埋或暴露溶剂表面。(c)  $r_H$ , 链间氢键对界面水分子数的置换系数; (d)  $r_C$ , 链间紧致度对界面水分子数的置换系数, 负值和正值分别表示释放或招募的界面水分子数。误差棒为 Bootstrap 95%置信区间, 灰色虚线标示对照组 (0 THz) 的 Bootstrap 均值, 红色星号表示与对照组的 Bootstrap 差异显著性 (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ )。

Fig. 2. Bootstrap estimates of linear displacement coefficients in Models 1 and 2 as functions of THz frequency, corresponding to Table S3: (a)  $k_H$ , displacement coefficient of inter-chain hydrogen bonds on SASA; (b)  $k_C$ , displacement coefficient of inter-chain compactness on SASA; (c)  $r_H$ , displacement coefficient of inter-chain hydrogen bonds on the number of interfacial water molecules; (d)  $r_C$ , displacement coefficient of inter-chain compactness on the number of interfacial water molecules. Error bars represent 95% bootstrap confidence intervals. The gray dashed line indicates the bootstrap mean of the control group (0 THz). Red asterisks denote significant differences from the control group (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).

太赫兹电场在特定频率显著重塑了这一置换关系。如表 2 和图 2 所示, 11.9 THz 是全频段中效应最集中的频率, 表现为回归系数的集体翻转和  $R^2$  的剧烈跃升。具体而言: 模型 1 的  $R^2$  从 0.531 升至 0.676 (+27%), 模型 2 ( $n_W \sim n_H + CCR$ )

的  $R^2$  从 0.121 升至 0.616 (+409%，达对照组的约五倍)。三个核心系数的符号均发生翻转。 $k_H$  从  $+3.83 \times 10^{-3}$  翻转至  $-3.98 \times 10^{-3}$  (Bootstrap  $p=0.006$ )， $r_H$  从  $+9.21 \times 10^{-3}$  翻转至  $-2.31 \times 10^{-2}$  (Bootstrap  $p<0.001$ )， $r_C$  从  $+9.2 \times 10^{-5}$  翻转至  $-1.10 \times 10^{-1}$  (Bootstrap  $p=0.034$ )。符号翻转的物理含义是：在 11.9 THz 电场作用下，链间氢键从 SASA 增大因素转变为 SASA 减小因素，链间接触的 SASA 掩埋效率和水分子释放效率同时被激活为全频段最强。我们将 11.9 THz 称为协同开关频率。

如表 2 和图 2 所示，在 20.5-32.2 THz 的中高频段，协同性发生解体，即不同频率选择性激活了不同的置换通道。20.5 THz 处，仅  $r_C$  显著强于对照组 ( $-0.170$ ，Bootstrap  $p=0.002$ )，而  $k_H$  和  $r_H$  已恢复至对照水平（符号均为正， $p>0.9$ ），模型 2 的  $R^2$  维持高位 (0.648)，表明该频率选择性增强了 CCR 对水分子的排除效率。37.3 THz 处激活通道发生转移： $r_H$  显著转负 ( $-0.015$ ，Bootstrap  $p=0.010$ )， $k_H$  处于边际显著 ( $-0.003$ ，Bootstrap  $p=0.048$ )，表明  $n_H$  重新获得了对水分子的排除驱动力。

### 3.4 界面水合密度的恒常性与面积驱动的紧致化机制

为进一步检验太赫兹电场是否改变了 RNA 表面单位面积对水分子的化学亲和力和力，我们将全部 8 个频率 $\times$ 20 副本共 160 个独立样本点的 ( $SASA$ ,  $n_W$ ) 合并，进行跨频率的全局线性回归。拟合结果显示， $n_W$  与  $SASA$  之间存在显著的线性关系 ( $R^2=0.664$ ，斜率  $\rho_{HB}=2.66$ ；图 3)，即约 66% 的  $n_W$  跨副本变异可由  $SASA$  单独解释。这一全局  $R^2$  表明  $SASA$  是  $n_W$  变化的主导因素，但并非唯一因素，仍有约 34% 的方差源自频率依赖的  $\rho_{HB}$  波动及界面水分子取向、第一溶剂化层重排等未建模的自由度。

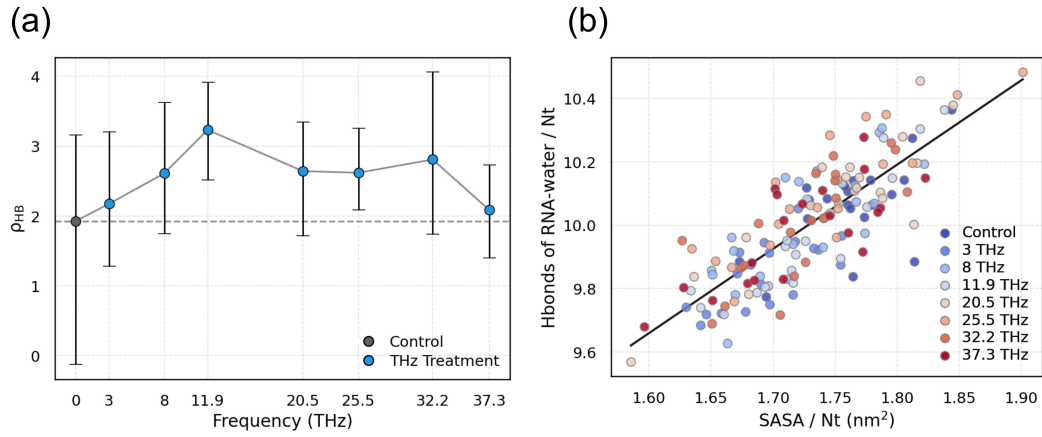


图 3 界面水合密度  $\rho_{HB}$  的频率依赖性与全频率  $n_W$ - $SASA$  全局线性拟合。(a)  $\rho_{HB}$  的 Bootstrap 估计值随太赫兹频率的变化（对应表 S3），误差棒为 Bootstrap 95%置信区间，灰色虚线标示对照组；(b) 全部 8 个条件共 160 个副本的  $n_W$ - $SASA$  散点图及全局最小二乘拟合。黑色实线的斜率  $\rho_{HB}=2.66$  代表每  $1 \text{ nm}^2$   $SASA$  约承载 2.7 个 RNA-水氢键,  $R^2=0.66$  表明  $n_W$  的变化约 66% 可由  $SASA$  解释。

Fig. 3. Frequency dependence of interfacial hydration density  $\rho_{HB}$  and global linear fitting of  $n_W$  versus  $SASA$  across all frequencies: (a) bootstrap estimates of  $\rho_{HB}$  as a function of THz frequency (corresponding to Table S3), with error bars representing 95% bootstrap confidence intervals and the gray dashed line indicating the control group; (b) scatter plot of  $n_W$  versus  $SASA$  for all 160 replicates across 8 conditions, along with global least-squares fitting. The black solid line has a slope  $\rho_{HB}=2.66$ , indicating that approximately 2.7 RNA–water hydrogen bonds are accommodated per  $\text{nm}^2$  of  $SASA$ .  $R^2=0.66$  suggests that about 66% of the variation in  $n_W$  can be explained by  $SASA$ .

表 3 各频率下的界面水合密度  $\rho_{HB}$

Table 3. Interfacial hydration density  $\rho_{HB}$  at each frequency.

| 频率(THz) | $\rho_{HB}(\text{OLS})$ | $R^2$ | Bootstrap | Bootstrap 95% CI | diff <sub>p</sub> |
|---------|-------------------------|-------|-----------|------------------|-------------------|
| 对照      | 2.03                    | 0.358 | 1.927     | [-0.13, 3.16]    | —                 |
| 3.0     | 2.16                    | 0.486 | 2.173     | [1.28, 3.20]     | 0.850             |
| 8.0     | 2.61                    | 0.578 | 2.61      | [1.75, 3.63]     | 0.494             |
| 11.9    | 3.23                    | 0.840 | 3.23      | [2.52, 3.91]     | 0.116             |
| 20.5    | 2.68                    | 0.763 | 2.64      | [1.71, 3.35]     | 0.414             |
| 25.5    | 2.58                    | 0.760 | 2.62      | [2.09, 3.26]     | 0.410             |
| 32.2    | 2.72                    | 0.583 | 2.81      | [1.74, 4.06]     | 0.384             |
| 37.3    | 2.10                    | 0.586 | 2.09      | [1.40, 2.73]     | 0.934             |

注：完整处理组频率数据见补充材料表 S3。

然而，对各频率单独拟合的  $\rho_{HB}$  进行 Bootstrap 显著性检验表明，所有处理组的  $\rho_{HB}$  与对照组（OLS 点估计 2.03，Bootstrap 均值 1.93）均无显著差异（Bootstrap 差异  $p$  值均  $>0.05$ ；表 3 和 S3）。尽管  $\rho_{HB}$  的点估计在不同频率间存在波动（2.10-3.23），但 Bootstrap 检验表明这种波动在  $n=20$  的小样本下无法与随机抽样变异区分开来。因此，在统计意义上，太赫兹电场并未系统性地改变 RNA 表面单位面积对水分子的化学亲和力。

基于上述结果，紧致化过程中  $n_w$  的响应可通过面积项（ $\Delta SASA \times \rho_{HB}$ ， $\Delta$  表示处理组均值减去对照组均值）与密度项（ $SASA \times \Delta \rho_{HB}$ ）的分解来理解。如表 S4 所示，以对照组  $\rho_{HB}=2.03$  为基准，计算“若  $\rho_{HB}$  恒定、仅由  $SASA$  减小应导致的理论  $n_w$  降幅”，并与实际  $n_w$  变化比较（详见表 S4）。在  $SASA$  降幅最大的 3.0 THz（-3.8%），实际  $n_w$  下降（-0.20）甚至超过了纯面积驱动的预测值（-0.14），表明该低频电场对界面水分子有额外的排除效应。而在  $SASA$  降幅仅为 2.0-2.4% 的 8.0 THz、32.2 THz 和 37.3 THz，实际  $n_w$  降幅小于纯面积驱动的预测值（残差：+0.002 至 +0.018，即实际  $n_w$  比预测值高），指示  $\rho_{HB}$  的微弱上调（虽未达统计显著）在一定程度上补偿了面积损失对  $n_w$  的下拉效应，使得  $n_w$  在统计上维持在对照水平。

综上，太赫兹电场对 RNA 四级结构的紧致化以  $SASA$  的面积压缩为主驱动，界面水合密度  $\rho_{HB}$  在统计意义上守恒。电场并不改变 RNA 表面的水合化学亲和力，而是通过调节  $SASA$  的面积间接控制界面水分子数量。 $SASA$  降幅大时（3.0 THz）， $n_w$  显著下降； $SASA$  降幅小时（2.0-2.4%）， $\rho_{HB}$  的微弱上调将  $n_w$  维持在对照水平附近，最终界面水分子数量取决于  $SASA$  减幅与  $\rho_{HB}$  上调方向之间的竞争。

### 3.5 太赫兹频率依赖的协同性调制与物理机制

综合前面分析结果，太赫兹电场对 RNA 四级结构的响应表现出显著的频率选择性。我们从三个维度，表征参量均值的显著性变化、线性置换系数的符号与

Bootstrap 显著性、 $R^2$ 的增强模式，将 0-37.3 THz 划分为四个功能频段。

(i) 基线区（无外场对照）。在无外场条件下， $CCR$  与  $SASA$  呈预期的负相关 ( $k_C=-1.95\times10^{-2}$ )，反映链间紧致导致的部分表面掩埋； $n_H$  与  $SASA$  呈弱正相关 ( $k_H=+3.83\times10^{-3}$ )，表明链间氢键在自然状态下并不有效掩埋溶剂可及表面。模型 1 的  $R^2=0.531$ ，即  $n_H$  和  $CCR$  联合解释了约 53% 的  $SASA$  变异；模型 2 的  $R^2$  仅为 0.121，说明链间接触对界面水分子数  $n_W$  的解释力极弱。

(ii) 低频紧致区 (3.0 THz 和 8.0 THz, 约  $100\text{ cm}^{-1}$  和  $267\text{ cm}^{-1}$ )。这是  $SASA$  降幅最大的频段。3.0 THz 处  $SASA$  和  $n_W$  同时显著下降 ( $\Delta SASA=-3.8\%$ ,  $\Delta n_W=-0.20$ ;  $p$  均小于 0.001)，为全频段中唯一同时出现面积收缩和水分子显著排除的频率。8.0 THz 处仅  $SASA$  显著下降 ( $-2.2\%$ ,  $p=0.012$ )， $n_W$  与对照无差异。两个频率的共有特征是：回归系数 ( $k_H$ 、 $k_C$ 、 $r_H$ 、 $r_C$ ) 与对照组均无 Bootstrap 显著差异 ( $p>0.05$ )，且模型 1 的  $R^2$  反而下降 (3.0 THz: 0.28; 8.0 THz: 0.12)。  $R^2$  下降系数不变的模式表明，该频段的紧致化并非通过增强链间接触对  $SASA/n_W$  的置换效率实现，或许更接近一种直接的力学压缩，即链间相互作用与溶剂界面的耦合方式未被系统性地改变。该频段落在核酸大分子中弱非键相互作用和螺旋整体扭转/弯曲的集体振动区内<sup>[24, 25]</sup>。我们推测，电场与此类螺旋集体振动模的耦合增强了沿螺旋轴的压缩涨落，以外部力学驱动的方式实现紧致化；水分子的氢键弯曲/伞状集体运动也可能辅助驱动疏水塌缩<sup>[14]</sup>，但其对 RNA 四级结构的相对贡献尚待定量区分。

(iii) 全协同区 (11.9 THz, 约  $397\text{ cm}^{-1}$ )。此频率是全文中最显著的特征频率，表现为回归系数的集体翻转和  $R^2$  的剧烈跃升： $k_H$ 、 $r_H$  和  $r_C$  的符号均由正翻转为负 (Bootstrap  $p$  分别为 0.006、 $<0.001$  和 0.034；精确数值见表 S2 和 S3)；模型 1 的  $R^2$  从 0.531 升至 0.676，模型 2 的  $R^2$  从 0.121 升至 0.616，后者较对照组提升了逾四倍。我们将这种四个系数同时激活的现象称为协同开关：在 11.9 THz

电场下，链间氢键从对抗紧致化转变为驱动紧致化，链间接触的 SASA 掩埋效率和水分子释放效率同时达到全频段最强， $n_H$  和 CCR 突然获得了对  $n_W$  的统计解释力。在频率归属上，11.9 THz 精准落入核糖环振动的特征频段 ( $300-500\text{ cm}^{-1}$ )<sup>[26]</sup>。先前拉伸动力学模拟表明 9-15 THz 外场可能通过核糖环共振显著增强 RNA 发夹的展开拉力<sup>[30]</sup>。我们据此推测，11.9 THz 电场通过核糖环的共振激发降低了链间接触形成过程中的去溶剂化自由能垒，使得在自然状态下不易维持的紧密链间接触在电场作用下触发了置换效率的协同跃升。

(iv) 选择性增强区 (20.5-37.3 THz, 约  $684-1244\text{ cm}^{-1}$ )。该频段区别于 11.9 THz 的关键特征是协同性的解体，不同回归系数在不同频率被选择性激活，而非集体翻转。20.5 THz 处仅  $r_C$  显著强于对照组 (Bootstrap  $p=0.002$ )， $k_H$  和  $r_H$  已恢复至对照水平，模型 2 的  $R^2$  维持高位 (0.648)；32.2 THz 处  $k_H$  出现边际显著的符号翻转 (Bootstrap  $p=0.046$ )，而  $r_C$  消退至不显著，标志着  $n_H$  对 SASA 的掩埋效率开始激活；37.3 THz 处激活通道发生转移， $r_H$  显著转负 ( $p=0.010$ ) 而  $r_C$  不再显著。在频率归属上，该频段对应 RNA 磷酸-糖骨架振动的指纹区， $1083\text{ cm}^{-1}$  (约 32.5 THz) 为对称  $\text{PO}_2^-$  伸缩振动， $1239\text{ cm}^{-1}$  (约 37.2 THz) 为反对称  $\text{PO}_2^-$  伸缩振动， $750-1500\text{ cm}^{-1}$  整体为碱基序列无关的磷酸骨架振动带<sup>[13]</sup>。我们推测，20.5 THz 处的效应源于电场与磷酸基团振动模式的耦合，选择性增强了 CCR 驱动的去水化通道；37.3 THz 处的效应则源于反对称  $\text{PO}_2^-$  伸缩振动与链间氢键几何的耦合，选择性激活了  $n_H$  驱动的去水化通道。25.5 THz 为两个通道均不显著的最低点，可能是两种振动耦合模式交叉衰减最深的区域。需要指出，上述机制解读是基于核糖环振动和磷酸伸缩振动的文献指认与本文回归分析结果的关联分析，其确切性在当前实验和模拟条件下有待独立验证，在此仅作为与数据一致的一种可能性供讨论。

## 4 结 论

本文通过全原子分子动力学模拟结合线性置换模型与 Bootstrap 统计分析，考察了 0-37.3 THz 正弦电场（场强 0.5 V/nm）对五链 5×G4C2 RNA 组装体四级结构的物理响应。结果表明，在所考察的全频段范围内，链间氢键数和链间紧致程度均未发生显著变化，组装体的四级拓扑结构保持完整；而 SASA 和界面水分子数  $n_w$  则表现出显著的频率选择性，即 SASA 在多个频率显著下降（最大降幅 3.8%，3.0 THz）， $n_w$  仅在降幅最大时显著减少，揭示紧致化并非简单的界面水分子排除过程。线性置换模型进一步表明，11.9 THz 处三个核心系数的符号同时由正翻转为负，模型决定系数剧烈跃升（模型 2 的  $R^2$  从 0.121 升至 0.616），触发了链间接触的 SASA 掩埋效率和水分子释放效率的全协同开关；20.5 THz 和 37.3 THz 则分别选择性增强了 CCR 驱动和  $n_H$  驱动的去水化通道。界面水合密度  $\rho_{HB}$  在所有频率下与对照组无统计差异，全频率  $n_w$ -SASA 全局线性拟合  $R^2=0.66$ ，表明太赫兹电场主要通过压缩 SASA 的面积实现紧致化，而非改变 RNA 表面的水合化学亲和力。上述结果揭示了太赫兹频率与 RNA 组装体链间-界面耦合之间的选择性相互作用，为利用太赫兹波调节 RNA 组装体结构提供了频率选择依据，也为太赫兹波与生物大分子相互作用的定量解析提供了方法学参考。

声明：在本工作的准备过程中，作者使用了 Reasonix 工具来数据分析和图表制作等。使用此工具后，作者对所有内容进行了严格的审查和编辑，并对出版物的内容承担全部责任。

附录 图 S1，表 S1, S2, S3, S4。

## 参考文献

- [1] DeJesus-Hernandez M, Mackenzie I R, Boeve B F, Boxer A L, Baker M, Rutherford N J, Nicholson A M, Finch N A, Flynn H, Adamson J 2011 *Neuron* **72** 245

- [2] Renton A E, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs J R, Schymick J C, Laaksovirta H, Van Swieten J C, Myllykangas L 2011 *Neuron* **72** 257
- [3] Majounie E, Renton A E, Mok K, Doppler E G, Waite A, Rollinson S, Chiò A, Restagno G, Nicolaou N, Simon-Sanchez J 2012 *Lancet Neurol.* **11** 323
- [4] van Blitterswijk M, DeJesus-Hernandez M, Niemantsverdriet E, Murray M E, Heckman M G, Diehl N N, Brown P H, Baker M C, Finch N A, Bauer P O 2013 *Lancet Neurol.* **12** 978
- [5] Reddy K, Zamiri B, Stanley S Y, Macgregor R B, Pearson C E 2013 *J. Biol. Chem.* **288** 9860
- [6] Jain A, Vale R D 2017 *Nature* **546** 243
- [7] Haeusler A R, Donnelly C J, Periz G, Simko E A, Shaw P G, Kim M S, Maragakis N J, Troncoso J C, Pandey A, Sattler R 2014 *Nature* **507** 195
- [8] Zhang Q, Valério M, Wang K, Yang L, Wang S, Souza P C, Gong Y 2026 *Molecules* **31** 1903
- [9] Zhang Q, Valério M, Grünwald L, Borges-Araújo L, Grünwald F, Wang S, Marrink S J, Gong Y, Souza P C 2026 *J. Phys. Chem. Lett.* **17** 6445
- [10] Ma S, Ding P, Zhou Z, Jin H, Li X, Li Y 2024 *Brain Sci.* **14** 279
- [11] Zhu Z, Wang T, Yang H, Zhu J, Francisco J S, Wang C 2025 *Research* **8** 1028
- [12] Sun L, Zhao L, Peng R Y 2021 *Mil. Med. Res.* **8** 28
- [13] Sarkar M, Dornberger U, Rozners E, Fritzsche H, Strömberg R, Gräslund A 1997 *Biochemistry* **36** 15463
- [14] Heyden M, Sun J, Funkner S, Mathias G, Forbert H, Havenith M, Marx D 2010 *Proc. Natl. Acad. Sci. USA* **107** 12068

- [15] Xu Y, Havenith M 2015 *J. Chem. Phys.* **143** 170901
- [16] Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark A E, Berendsen H J 2005 *J. Comput. Chem.* **26** 1701
- [17] Zgarbová M, Otyepka M, Sponer J, Mladek A, Banas P, Cheatham III T E, Jurecka P 2011 *J. Chem. Theory Comput.* **7** 2886
- [18] Zgarbová M, Sponer J, Otyepka M, Cheatham III T E, Galindo-Murillo R, Jurecka P 2015 *J. Chem. Theory Comput.* **11** 5723
- [19] Bussi G, Donadio D, Parrinello M 2007 *J. Chem. Phys.* **126** 014101
- [20] Parrinello M, Rahman A 1980 *Phys. Rev. Lett.* **45** 1196
- [21] Parrinello M, Rahman A 1981 *J. Appl. Phys.* **52** 7182
- [22] Li Y, Chang C, Zhu Z, Sun L, Fan C 2021 *J. Am. Chem. Soc.* **143** 4311
- [23] Bo W, Guo L, Yang Y, Ma J, Wang K, Tang J, Wu Z, Zeng B, Gong Y 2020 *IEEE Access* **8** 10305
- [24] Globus T, Bykhovskaia M, Woolard D, Gelmont B 2003 *J. Phys. D: Appl. Phys.* **36** 1314
- [25] Woolard D, Globus T, Gelmont B, Bykhovskaia M, Samuels A, Cookmeyer D, Hesler J, Crowe T, Jensen J, Jensen J 2002 *Phys. Rev. E* **65** 051903
- [26] Globus T, Woolard D, Samuels A, Gelmont B, Hesler J, Crowe T, Bykhovskaia M 2002 *J. Appl. Phys.* **91** 6105
- [27] Michaud-Agrawal N, Denning E J, Woolf T B, Beckstein O 2011 *J. Comput. Chem.* **32** 2319
- [28] Martínez L, Souza P C, Garcia W, Batista F A, Portugal R V, Nascimento A S, Nakahira M, Lima L M, Polikarpov I, Skaf M S 2010 *J. Phys. Chem. B* **114** 1529

- [29] Eisenhaber F, Lijnzaad P, Argos P, Sander C, Scharf M 1995 *J. Comput. Chem.* **16**  
273
- [30] Zhang Q, Yang L, Wang K, Guo L, Ning H, Wang S, Gong Y 2023 *iScience* **26**  
107571

## Frequency-selective responses of an RNA assembly to terahertz external fields\*

ZHANG Qin<sup>1)</sup> WANG Kaicheng<sup>1)</sup> SUN Yuankun<sup>1)</sup> WANG Shaomeng<sup>1)</sup> GONG Yubin<sup>1)†</sup>

1) (National Key Laboratory on Vacuum Electronics, Terahertz Radiation and Application Key  
Laboratory of Sichuan Province, School of Electronic Science and Engineering, University of  
Electronic Science and Technology of China, Chengdu 611731, China)

### Abstract

The hexanucleotide repeat expansion GGGGCC (G4C2) in the *C9orf72* gene is the most common genetic cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). The transcribed RNA self-assembles into multi-chain quaternary structures through multivalent base pairing, and this assembly behavior constitutes the molecular basis for pathological liquid–liquid phase separation and the formation of pathological condensates. Here, using all-atom molecular dynamics simulations, we investigate how terahertz (THz) frequencies modulate the physical coupling between interchain contacts and the solvent interface in a five-chain 5xG4C2 RNA assembly. Under a field strength of 0.5 V/nm, seven distinct frequencies spanning 0–37.3 THz were examined, each with 20 independent 200-ns replicates. Four structural parameters covering two orthogonal dimensions, interchain contacts and the solvent interface, were analyzed using Student's t-test, three linear displacement models, and bootstrap resampling (1000 iterations). Across the entire frequency range, neither the interchain hydrogen-bond network nor backbone stacking was significantly affected. In contrast, solvent-accessible surface area (SASA) exhibited a marked frequency-selective decrease, whereas the number of interfacial water molecules dropped only at the frequency of maximal compaction (3.0 THz), indicating that compaction is not simply due to interfacial water expulsion. Linear permutation model analysis revealed a

cooperative switch at 11.9 THz, where the signs of three core coefficients flipped simultaneously and the coefficient of determination rose sharply, demonstrating that interchain contacts **become** effective drivers of surface burial and water release. At 20.5 THz, the compactness-driven water-release channel was selectively activated; this channel subsided at 32.2 THz, where **a** hydrogen-bond-driven channel began to emerge, and became fully activated at 37.3 THz. Interfacial hydration density showed no statistically significant change at any frequency, and global regression confirmed that changes in interfacial water are dominated by *SASA* compression rather than **altered surface chemical affinity**. Collectively, these findings establish that THz frequency provides a continuously tunable means of resolving interactions within the assembly: different frequencies selectively elicit responses along distinct physical dimensions, **mechanical compaction**, cooperative switching of surface burial and water-release efficiency, and channel-selective **activation**, while the interfacial hydration density remains conserved. This frequency-encoding framework offers a physical basis for understanding **terahertz-nucleic-acid** supramolecular interactions at the molecular level.

Keywords: GGGGCC RNA repeats, Terahertz waves, Molecular Dynamics, RNA assembly

---

\* Project supported by the National Natural Science Foundation of China (Grant Nos. 62527804, T2241002, 62427806, 62588201).

† Corresponding ybgong@uestc.edu.cn  
The first author.E-mail: 202111022406@std.uestc.edu.cn