

# 端粒 G-四链体结合酞菁分子的分子模拟研究

楼晓勤,黄依宁,许奕,王梦露,李建峰,王志国

(杭州师范大学 医学部,浙江 杭州 311121)

**摘要** 为揭示负电性稳定剂与人端粒 G-四链体的作用机制,采用分子对接和分子动力学方法对新型负电性四磺酸酞菁分子(anionic phthalocyanine 3,4',4'',4'''-tetrasulfonic acid)和人端粒平行构象 G-四链体进行了系统研究。构象分析发现,酞菁分子能够单独或同时稳定结合于 G-四链体的顶层和底层 G-四分体平面并有效提高 G-四链体的稳定性;结合自由能分析表明,酞菁分子堆叠于顶层 G-四分体的结合模式在能量上更为有利,同时第二个酞菁分子的引入能够进一步提升结合强度。这些发现对新型端粒 G-四链体稳定剂的开发具有重要的指导意义。

**关键词** G-四链体;端粒;分子对接;分子动力学;结合自由能

**中图分类号** Q641.3

**文献标志码** A



开放科学(资源服务)标识码(OSID)

## Molecular Simulation Study of the Binding Interaction between Telomere G-quadruplex and Phthalocyanine Molecule

LOU Xiaoqin, HUANG Yining, XU Yi, WANG Menglu,  
LI Jianfeng, WANG Zhiguo

(School of Medicine, Hangzhou Normal University, Hangzhou 311121, China)

**Abstract** In order to reveal the mechanism of interaction between the negatively charged stabilizer and the human telomere G-quadruplex, the current work applies molecular docking and molecular dynamics methods to the novel anionic phthalocyanine molecule (anionic phthalocyanine 3,4',4'',4'''-tetrasulfonic acid) and human telomere parallel G-quadruplex. Conformational analysis finds that the phthalocyanine molecule bind stably to the top and bottom G-tetrads of the G-quadruplex alone or at the same time, which effectively improve the stability of the G-quadruplex; binding free energy analysis shows that the top G-tetrad stacking binding mode of the phthalocyanine molecule is energetically more favorable, and the introduction of the second phthalocyanine molecule would further enhance the binding strength. These findings have important guiding significance for the development of novel telomere G-quadruplex stabilizers.

**Key words** G-quadruplex; telomere; molecular docking; molecular dynamics; binding free energy

收稿日期:2020-11-17

基金项目:国家自然科学基金项目(81971311);浙江省自然科学基金项目(LY18H250002)资助

通讯作者:李建峰,男,汉族,博士,副教授,硕士生导师,研究方向:脑衰老和 PD, E-mail: lijianfeng@hznu.edu.cn;王志国,男,汉族,博士,副研究员,硕士生导师,研究方向:计算生物学, Email: zhgwang@aliyun.com。

## 0 引言

G-四链体(G4)作为一种核酸高级结构,通常形成于端粒末端 DNA<sup>[1]</sup>、癌基因启动子<sup>[2]</sup>、非翻译 mRNA<sup>[3]</sup>等 G-碱基富集区。人端粒 DNA 末端有一段由[T<sub>2</sub>AG<sub>3</sub>]重复序列构成的长约 100—200 碱基的单链突出,该突出在生理条件下易折叠形成高稳定性的 G4 结构,从而抑制端粒酶延长端粒<sup>[4-7]</sup>。研究发现,约 85% 的癌症依赖端粒酶活性,而稳定 G4 能够抑制肿瘤的发生和发展<sup>[8,9]</sup>。因此,G-四链体稳定剂的开发成为癌症治疗领域关注的热点。

端粒 G4 在链取向、糖苷构象和末端延申序列等方面的差异导致了其构象的多样性,典型的 G4 结构呈现平行构象<sup>[10]</sup>、反平行构象<sup>[11]</sup>和杂合构象<sup>[12,13]</sup>。针对 G4 的结构特征,人们已经开发出一系列靶向性小分子稳定剂,包括黄连素<sup>[14]</sup>、咪唑<sup>[15]</sup>、吡啶<sup>[16]</sup>和喹啉<sup>[17]</sup>类化合物。这些化合物在结构上具有平面共轭性和正电性,使得它们能够通过  $\pi$ - $\pi$  和静电相互作用与 G4 相结合。然而静电作用的特性决定了这些稳定剂不能有效区分 G4 和 DNA 双螺旋结构,使其在选择性上表现不佳,导致目前尚无靶向 G4 的癌症治疗药物得以应用。因此,选择性和特异性是目前 G4 稳定剂研究领域关注的焦点。

最近研究发现,负电性的四磺酸酞菁分子(anionic phthalocyanine 3,4',4'',4'''-tetrasulfonic acid, APC)表现出良好的水溶性和端粒酶抑制活性,并且即使在过量双螺旋结构 DNA 存在的条件下 APC 仍表现出对端粒 G4 的高选择性<sup>[18]</sup>,而 APC 与最主要的端粒 G4 结构—平行构象 G4<sup>[19]</sup>之间作用机制尚属未知。本文应用分子模拟方法确定了 APC 与平行构象 G4 的结合模式,通过主成分、非键相互作用和氢键分析解析了 G4 和 APC 复合物的构象变化特征,并进一步通过结合自由能计算和自由能分解确定了在二者结合中起关键作用的能量类型和关键碱基,从而在原子和分子水平上揭示了 APC 与平行构象端粒 G4 之间的作用机制。

## 1 方法

### 1.1 APC 与端粒平行构象 G4 结构

APC 结构(图 1a)由 GaussView 软件构建并应用密度泛函理论在 B3LYP/6-31G(d)水平下进行构型优化,其原子电荷在 HF/6-31G(d)水平下通过限制性静电势(RESP)计算得到<sup>[20]</sup>。分子动力学模拟中,APC 分子的其它力场参数由 AmberTools 软件基于 AMBER GAFF 力场构建<sup>[21]</sup>。端粒平行构象 G4 结构由 PDB 数据库下载(PDB:1KF1)<sup>[10]</sup>并删除末端冗余核苷酸获得(图 1b),以使 G4 序列(5'-GGGTTAGGGT-TAGGGTTAGGG-3')<sup>[18]</sup>与实验报道一致。

### 1.2 分子对接

以达到动力学平衡状态的端粒平行构象 G4 结构为受体(图 1c,1d),以优化后的 APC 结构为配体,应用 AutoDock 4.2.6 软件<sup>[22]</sup>进行分子对接计算。对接计算应用 Lamarck 遗传算法,APC 分子中所有可旋转键均设为柔性,一个包含  $100 \times 100 \times 100$  个格点、格点间距为  $0.375 \text{ \AA}$  的正方体盒子用于定义配体结合空间。对 500 次遗传计算的运行结果进行分组和比较,选取构象最多的两个组中结合能最负的构象作为 APC 的结合构象<sup>[20]</sup>。

### 1.3 分子动力学

应用 AMBER 12 软件<sup>[23]</sup>进行分子动力学模拟。将 G4 和 G4-APC 复合物分别置于尖端截断的八面体 TIP3P 水盒中心,设置其与水盒边界的距离为  $10.0 \text{ \AA}$ ,在水环境中添加钾离子以使体系保持电中性。对 G4 应用 FF99SB 力场及 parmbsc1 和  $\chi_{OL3+OL15}$  矫正<sup>[24]</sup>,对环境和 G4 中的 K<sup>+</sup> 分别应用 Amber 标准参数(半径  $2.658 \text{ \AA}$ , 井深  $0.00328 \text{ kJ} \cdot \text{mol}^{-1}$ )和校正参数(半径  $1.705 \text{ \AA}$ , 井深  $0.1936829 \text{ kJ} \cdot \text{mol}^{-1}$ )<sup>[25]</sup>。分子动力学模拟过程中,周期性边界条件、静电与范德华相互作用截断值、能量最小化、升温过程、温度与压力控制等参数设置均与文献报导一致<sup>[26]</sup>。

### 1.4 主成分分析

主成分分析(PCA)通过过滤模拟体系大量的局域性波动而揭示分子动力学中的本质运动<sup>[27]</sup>。应用交互式本质动力学方法(IED)<sup>[28]</sup>和 AmberTools 的 CPPTRAJ 模块对所有 G4-APC 复合物中 G4 结构的骨架

原子进行分析,以豪猪图方式显示前两个本征向量所代表的运动方式。

### 1.5 非共价键相互作用

应用 NCIPLLOT 软件<sup>[29]</sup> 计算 APC 与 G4 之间的非共价键相互作用,步长设为 0.10。应用 VMD 软件<sup>[30]</sup> 构建相互作用等值面,isovalue 设为 0.3 au。

### 1.6 结合自由能

分子力学/广义波恩表面积方法(MM/GBSA)<sup>[31]</sup> 能够准确预测荷电分子的溶剂化自由能<sup>[32]</sup>,故应用此方法根据以下方程计算 APC 分子与端粒 G4 之间的结合自由能

$$\Delta G_{\text{bind}} = G_{\text{complex}} - (G_{\text{G4}} + G_{\text{APC}}), \quad (1)$$

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S \approx \Delta E_{\text{MM}} + \Delta G_{\text{solv}} - T\Delta S, \quad (2)$$

$$\Delta E_{\text{MM}} = \Delta E_{\text{int}} + \Delta E_{\text{ele}} + \Delta E_{\text{vdW}}, \quad (3)$$

$$\Delta G_{\text{solv}} = \Delta G_{\text{GB}} + \Delta G_{\text{SA}}, \quad (4)$$

其中  $E_{\text{MM}}$  是气相状态下的相互作用能,包括内部应变能  $E_{\text{int}}$  (键,角和二面角能量)、静电相互作用能  $E_{\text{ele}}$  和范德华(van der Waals, vdW)作用能  $E_{\text{vdW}}$ 。 $G_{\text{solv}}$  为溶剂化自由能,包括极性项  $G_{\text{GB}}$  和非极性项  $G_{\text{SA}}$ 。应用单轨迹近似计算结合自由能时, $\Delta E_{\text{int}}$  对结合自由能的贡献可忽略。应用广义波恩模型计算  $\Delta G_{\text{GB}}$  时,体系内部和外部介电常数分别设为 4 和 80。应用 LCPO 算法<sup>[33]</sup> 计算  $\Delta G_{\text{SA}}$ :  $\Delta G_{\text{SA}} = \gamma\Delta\text{SASA} + \beta$ , 其中  $\gamma$  和  $\beta$  分别设为  $0.0072 \text{ kcal} \cdot \text{\AA}^{-2}$  和 0。应用 AmberTools 中的 NModes 模块通过正则模式分析(NMA)计算熵变( $T\Delta S$ )对结合自由能的贡献。从各 G4-APC 复合物最后 40 ns 分子动力学轨迹中平均提取 500 个结构用于计算  $\Delta E_{\text{ele}}$ ,  $\Delta E_{\text{vdW}}$ ,  $\Delta G_{\text{GB}}$  和  $\Delta G_{\text{SA}}$ 。由于 NMA 计算量较大,故从最后 40 ns 分子动力学轨迹中平均提取 100 个结构用于计算熵变。

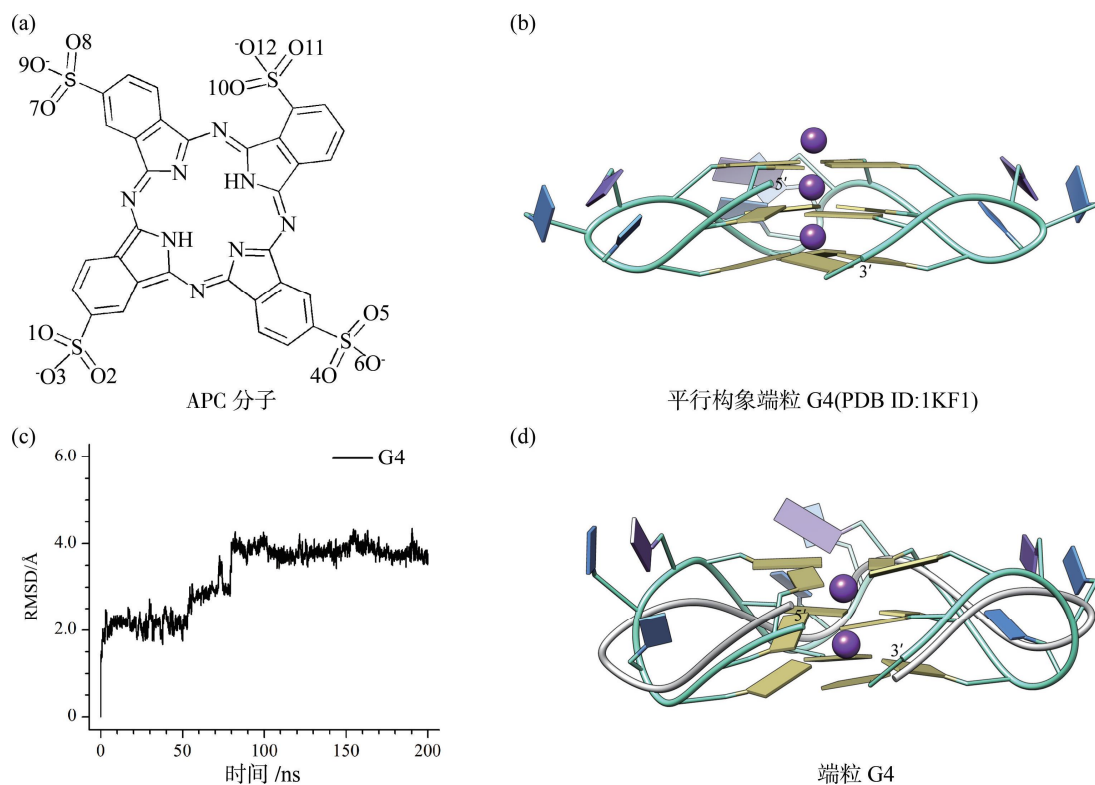


图 1 APC 分子及端粒平行构象 G4 (a) APC 分子结构; (b) 端粒平行构象 G4 晶体结构, 胸腺嘧啶、腺嘌呤和鸟嘌呤分别以蓝色、紫色和黄色扁平长方体表示,  $\text{K}^+$  以紫色圆球表示; (c) 端粒 G4 的 RMSD 值; (d) 分子动力学平衡状态 G4 与晶体结构的构象比对, 晶体结构以灰色飘带显示

## 2 结果与讨论

### 2.1 端粒 G4 与 APC 分子的结合模式

均方根偏差(RMSD)分析表明,平行构象端粒 G4 在 200 ns 时已达到平衡状态(图 1c)。三维构象比对发现,平衡状态 G4 相比晶体结构的构象差异主要位于 loop 区域(包括 dT4dT5dA6, dT10dT11dA12 和 dT16dT17dA18),参与形成 G-四分体的鸟嘌呤则吻合较好(图 1d)。值得注意的是,晶体结构中位于顶层

G-四分体外侧的  $K^+$  离子在分子动力学模拟过程中自发逃逸到溶剂环境中(图 1b, 1d), 使顶层 G-四分体暴露从而成为潜在的配体结合位点, 此发现与文献报导一致。

以平衡状态 G4 作为受体结构的分子对接计算发现, APC 分子主要以平行方式堆积于端粒 G4 顶层和底层 G-四分体外侧(图 2)。结合于顶层 G-四分体时, APC 共轭骨架与 dG1, dG7, dG13 和 dG19 形成  $\pi$ - $\pi$  相互作用, 磺酸官能团既与 dG1, dA12 和 dA18 形成阴离子- $\pi$  相互作用又与 dG19 形成氢键相互作用  $O8(APC^c) \cdots H22-N2(dG19)$  (图 2a)。APC 分子结合于底层 G-四分体时, 主要通过共轭骨架与 dG3, dG9, dG15 和 dG21 形成  $\pi$ - $\pi$  相互作用, 同时通过磺酸官能团与 dG3, dG9 和 dG21 形成氢键相互作用, 即  $O7(APC^b) \cdots H22-N2(dG3)$ ,  $O10(APC^b) \cdots H22-N2(dG9)$  和  $O2(APC^b) \cdots H22-N2(dG21)$  (图 2b)。当两分子 APC 同时结合于端粒 G4 的顶层和底层 G-四分体时,  $APC^{2t}$  和  $APC^{2b}$  与端粒 G4 的作用方式分别与  $APC^c$  和  $APC^b$  基本相同(图 2c), 表明已经结合的 APC 分子( $APC^{2t}$ )对第二个 APC 分子( $APC^{2b}$ )与端粒 G4 的结合影响甚微。

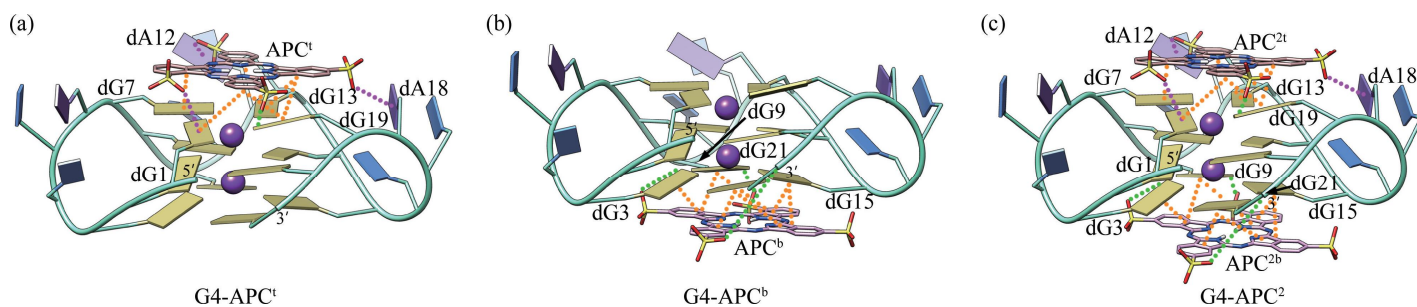


图 2 分子对接(a)APC 分子结合于顶层 G-四分体(标记为  $APC^c$ );(b)APC 分子结合于底层 G-四分体(标记为  $APC^b$ );(c)APC 分子同时结合于顶层和底层 G-四分体(分别标记为  $APC^{2t}$  和  $APC^{2b}$ ), 橙色虚线代表  $\pi$ - $\pi$  相互作用, 紫色虚线代表阴离子- $\pi$  相互作用, 绿色虚线代表氢键相互作用

## 2.2 APC 分子与端粒 G4 的结合特征

$APC^c$  分子的 RMSD 值在整个分子动力学过程中始终波动很小, 显示其与端粒 G4 结合的稳定性(图 3(a))。虽然端粒 G4 的 RMSD 值在 130-140 ns 时有较大波动, 但在动力学模拟后期趋于稳定, 表明 G4 结构达到平衡状态。均方根波动(RMSF)可以表征蛋白质或核酸等生物大分子中残基的柔性, 分析发现  $APC^c$ -G4 体系中 G4 具有与晶体结构相似的 RMSF 变化趋势, 即 loop 区域碱基的 RMSF 数值较高而参与构成 G-四分体的鸟嘌呤呈现较低 RMSF 值(图 3(a)), 表明 G4 结构的 loop 区域具有较大柔性而 G-四分体则稳定性较高。结构比对显示, 动力学平衡后  $APC^c$  仍平行堆积于顶层 G-四分体, 此时 G4 与晶体结构在构象上的差异主要位于 loop 区域(图 3(b)), 与 RMSF 结果一致。主成分分析发现, 前五个本征向量对  $APC^c$ -G4 体系中 G4 所有本质运动(由三十个本征向量近似表征)的贡献达 59.13%(图 3(c)), 因此分子动力学过程中 G4 主要的本质运动方式可由前几个本质向量进行表征。豪猪图分析显示, 第一本征向量主要表征了 5'-端碱基向上的运动, 第二本征向量主要表征了第一和第二 loop 结构的右向和左向运动(图 3(c))。这两个本征向量所表征的运动趋势与结构比对的结果相一致(图 3(b), (d))。非共价键相互作用分析表明, 平衡状态下  $APC^c$  主要与端粒 G4 顶层 G-四分体(包括 dG1, dG7, dG13 和 dG19)形成范德华相互作用(图 3(e))。分子对接中  $APC^c$  通过磺酸官能团所形成的阴离子- $\pi$  相互作用和分子间氢键相互作用由于稳定性差而消失。需要注意的是,  $APC^c$  与 dG1 形成了一个新的分子间氢键  $O8(APC^c) \cdots HO5'-O5'(dG1)$  (图 3e), 有利于  $APC^c$  与 G4 的稳定结合。

$APC^b$  分子和端粒 G4 在分子动力学过程中的 RMSD 波动均很小, 显示二者结构的稳定性(图 4(a))。 $APC^b$ -G4 体系中 G4 具有与晶体结构相似的 RMSF 变化趋势(图 4(a)), 表明分子动力学模拟的可靠性。与 RMSD 和 RMSF 结果一致, 结构比对发现平衡状态下  $APC^b$  仍平行堆积于底层 G-四分体且 G4 与晶体结构的构象差异主要位于 loop 区域(图 4(b))。 $APC^b$ -G4 体系中, 前五个本征向量对 G4 本质运动的贡献为 53.47%(图 4(c))。豪猪图显示, 第一本征向量主要表征了第二和第三 loop 的右向和左后向运动, 第二本征向量主要表征了第二和第三 loop 的内向和外向运动(图 4(c))。非共价键相互作用分析表明, 平衡状态下  $APC^b$  主要通过范德华相互作用结合于端粒 G4 底层 G-四分体(包括 dG3, dG9, dG15 和 dG21)外侧(图 4(e))。 $APC^b$  通过磺酸官能团与底层 G-四分体碱基所形成的分子间氢键(图 2(b))在分子动力学模拟过程中



很快消失,显示出底层堆积结合模式下氢键相互作用的不稳定性。

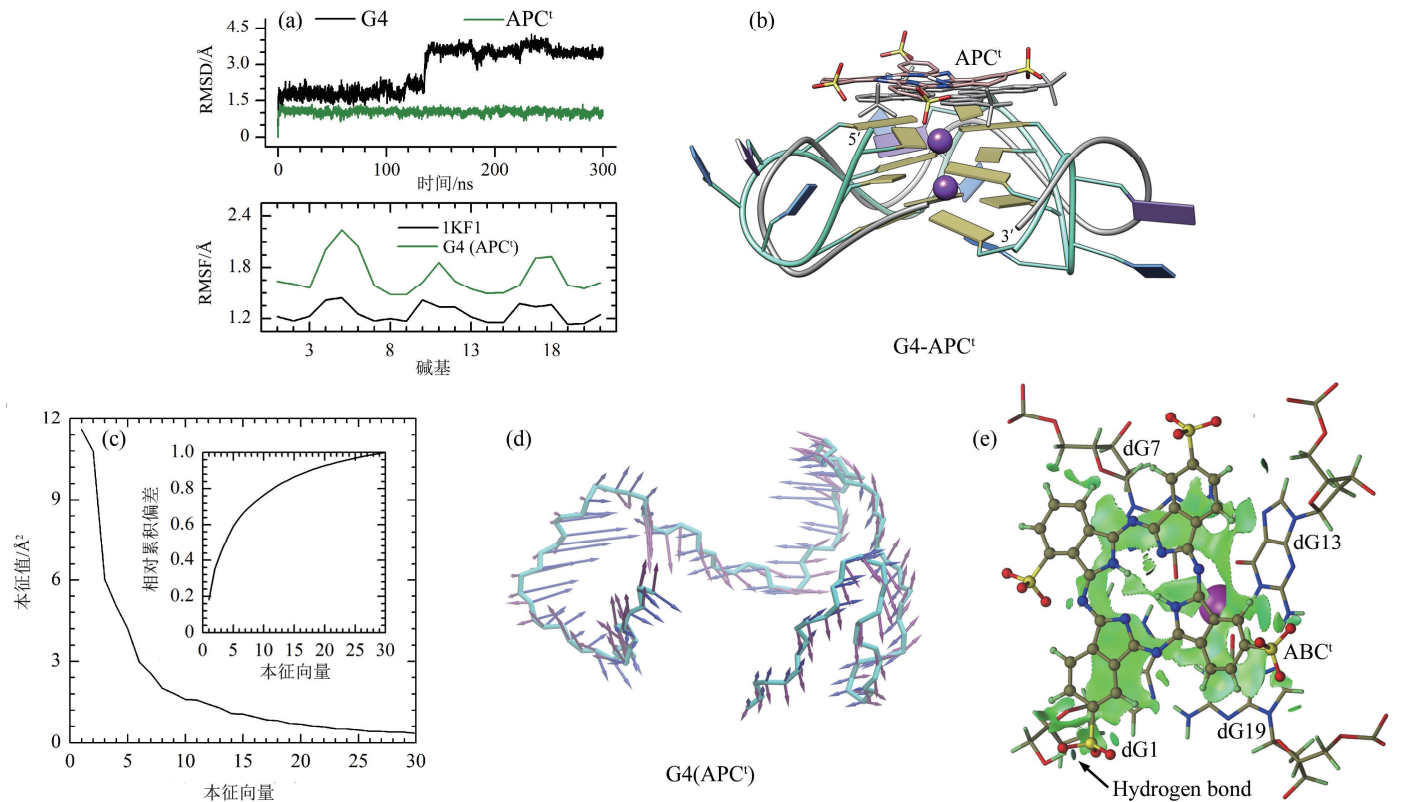


图 3 APC<sup>c</sup>与端粒 G4 的结合特征(a) RMSD 和 RMSF 分析;(b) 结构比对,G4 晶体结构以灰色飘带显示,APC<sup>c</sup>的对接构象以灰色棍状模型显示;(c) 主成分分析;(d) 豪猪图,暗红色和靛蓝色箭头分别表示第一和第二本征向量,箭头方向指示运动方向,箭头长度表示运动幅度;(e) 非共价键相互作用,绿色等值面表示范德华相互作用,蓝绿色等值面表示氢键相互作用

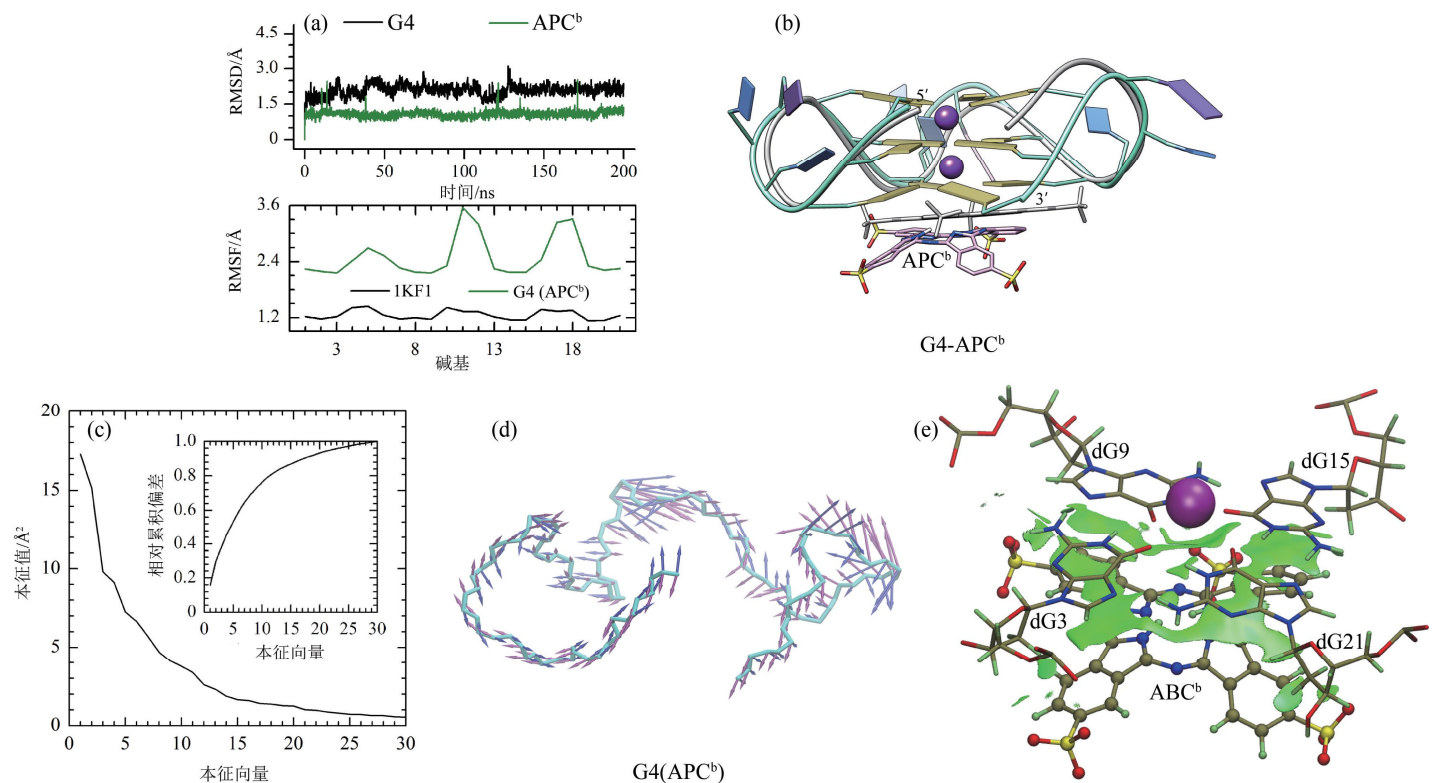


图 4 APC<sup>b</sup>与端粒 G4 的结合特征(a) RMSD 和 RMSF 分析;(b) 结构比对,G4 晶体结构以灰色飘带显示,APC<sup>b</sup>的对接构象以灰色棍状模型显示;(c) 主成分分析;(d) 豪猪图,暗红色和靛蓝色箭头分别表示第一和第二本征向量,箭头方向指示运动方向,箭头长度表示运动幅度;(e) 非共价键相互作用,绿色等值面表示范德华相互作用,蓝绿色等值面表示氢键相互作用

双分子 APC 和端粒 G4 的 RMSD 波动均较小,表明它们所形成的三明治结构的稳定性(图 5(a))。APC<sup>2</sup>-G4 体系中 G4 同样呈现出 loop 碱基高 RMSF 值和 G-四分体碱基低 RMSF 值的特性。与晶体结构的 RMSF 变化趋势相比,dA6 和 dA12 的 RMSF 值明显偏高,表明双分子 APC 的结合使得 dA6 和 dA12 在结

合复合物中的柔性增加(图 5(a))。结构比对发现,APC<sup>2t</sup>和 APC<sup>2b</sup>的平衡构象与对接构象重叠较好,而端粒 G4 与晶体结构的构象差异主要位于 5-末端和 loop 区域(图 5(b))。APC<sup>2</sup>-G4 体系中,前五个本征向量对 G4 本质运动的贡献达 61.59%(图 5(c))。豪猪图显示,第一本征向量主要表征了第一 loop 结构的左向运动和 5-末端的下向运动,第二本征向量主要表征了第一 loop 的上向运动和碱基 dG7 及 dG8 的下向运动(图 5(d))。非共价键相互作用分析表明,平衡状态下 APC<sup>2t</sup>和 APC<sup>2b</sup>分别与端粒 G4 顶层和底层 G-四分体形成广泛的范德华相互作用(图 5(e),(f)),以维持它们与端粒 G4 的稳定结合。此外,APC<sup>2t</sup>还同时与 loop 碱基 dA12 形成范德华相互作用以及与 dG1 形成分子间氢键相互作用 O7(APC<sup>2t</sup>)...HO5'-O5'(dG1),有利于进一步提升 APC<sup>2t</sup>与端粒 G4 结合强度(图 5(e))。

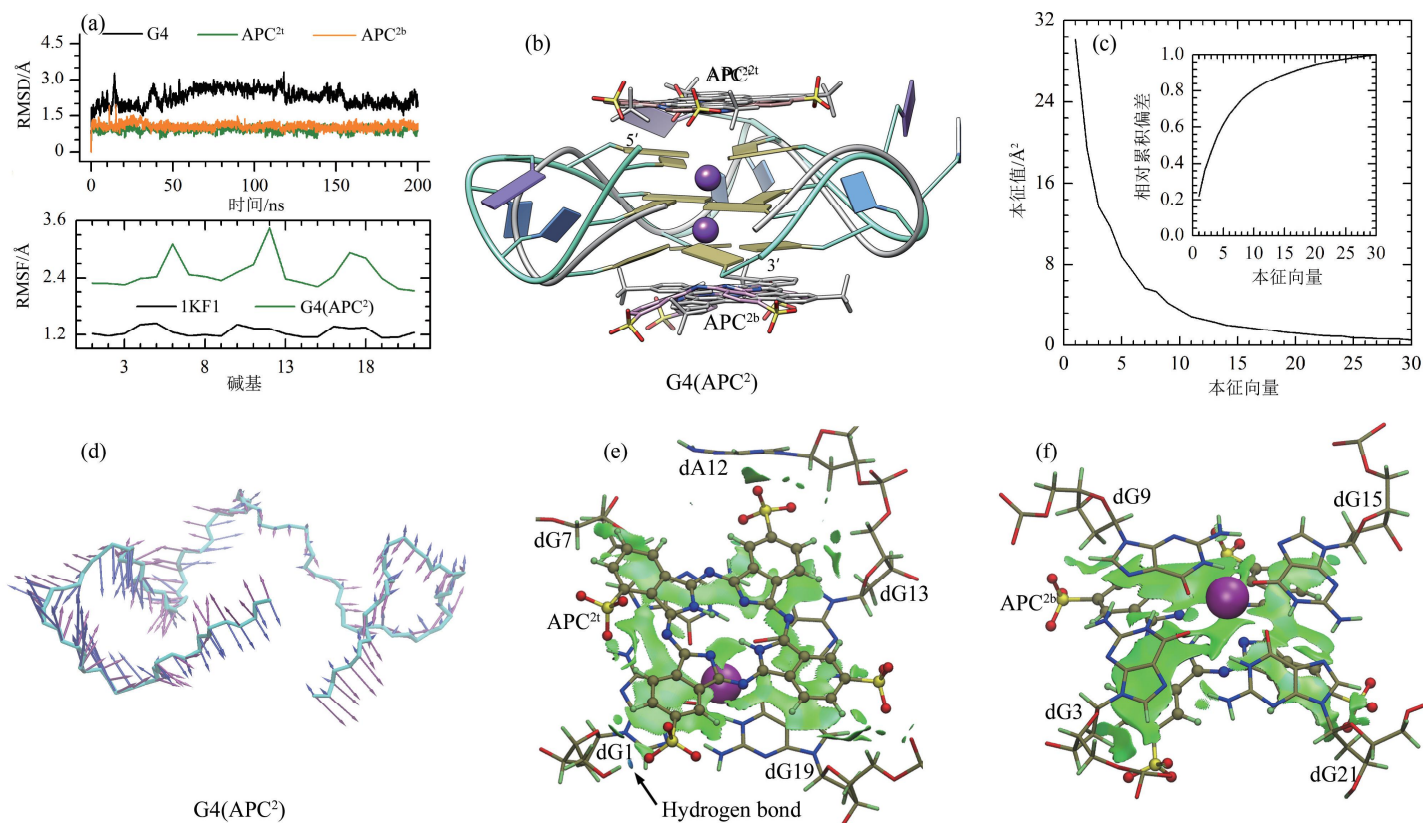


图 5 双分子 APC 与端粒 G4 的结合特征 (a) RMSD 和 RMSF 分析;(b) 结构比对,G4 晶体结构以灰色飘带显示,APC 分子的对接构象以灰色棍状模型显示;(c) 主成分分析;(d) 豪猪图,暗红色和靛蓝色箭头分别表示第一和第二本征向量,箭头方向指示运动方向,箭头长度表示运动幅度;(e) 非共价键相互作用,绿色等值面表示范德华相互作用

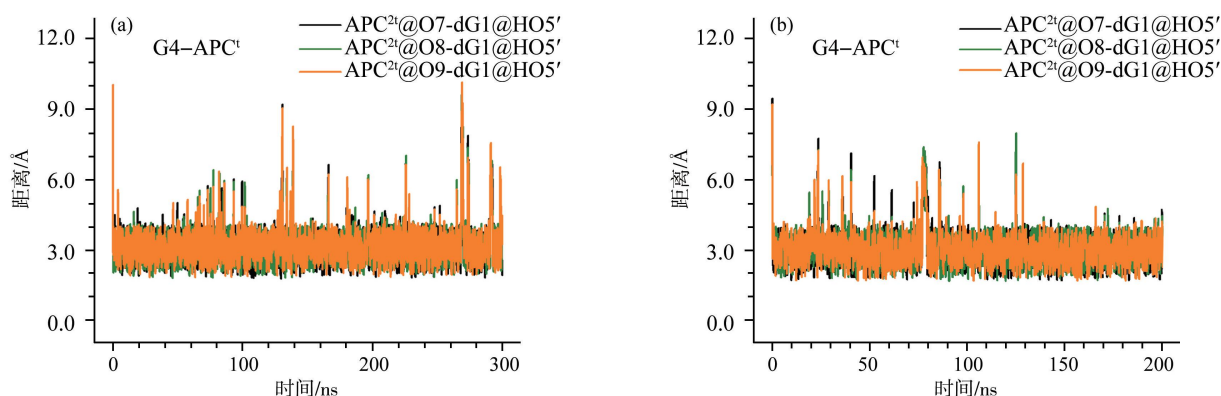
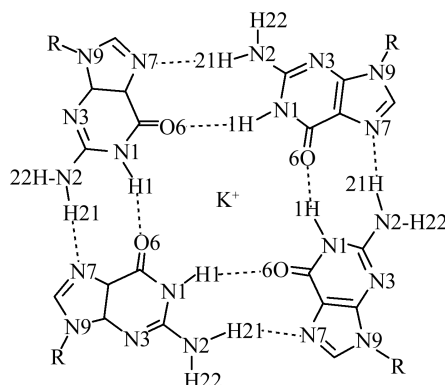


图 6 分子间氢键的距离 (a) APC<sup>1</sup>与端粒 G4 碱基 dG1 间的氢键距离;(b) APC<sup>2t</sup>与端粒 G4 碱基 dG1 间的氢键距离

分子间氢键是配体与生物大分子之间重要的相互作用方式,是二者稳定结合的重要贡献因素<sup>[34]</sup>。由以上分析可知,只有堆积于顶层 G-四分体外侧时 APC 分子才能与端粒 G4 结构形成分子间氢键。为了考察分子间氢键的稳定性,图 6 给出了分子动力学过程中氢键的距离变化。由图 3(e),5(e)和 6 可知,APC<sup>1</sup>和 APC<sup>2t</sup>均可通过相同的磺酸基 O 原子与端粒 G4 中 dG1 碱基的羟基 H 原子形成氢键。由于连接磺酸基与 APC 骨架的 C-S 单键可自由旋转,使得磺酸基中三个 O 原子(O7,O8 和 O9)与 dG1 羟基 H 原子的平均距离均约为 3.06 Å。同时,该磺酸基中每个 O 原子参与的分子间氢键的持续时间均约占分子动力学模拟总时

间的32.67%,这意味着在超过 98%的动力学模拟时间里 APC 分子和 dG1 碱基间均存在氢键相互作用,表明分子间氢键相互作用的稳定性。

表 1 平行构象端粒 G-四链体的 Hoogsteen 氢键特征



模型	Hoogsteen 氢键 <sup>1</sup>	最小值		最大值		平均值	
		时长/%	键长/Å	时长/%	键长/Å	时长/%	键长/Å
G4	O6...H1-N1 <sup>t</sup>	91.73	2.94	96.42	2.97	94.47	2.96
	N7...H21-N2 <sup>t</sup>	89.28	3.06	98.31	2.98	94.67	3.02
	O6...H1-N1 <sup>c</sup>	92.64	3.17	95.54	3.07	94.27	3.13
	N7...H21-N2 <sup>c</sup>	99.29	2.96	99.89	2.93	99.49	2.96
	O6...H1-N1 <sup>b</sup>	98.99	3.00	99.28	3.01	99.15	2.99
	N7...H21-N2 <sup>b</sup>	99.52	2.98	99.61	2.97	99.56	2.97
G4 (APC <sup>a</sup> )	O6...H1-N1 <sup>t</sup>	99.85	2.93	99.94	2.91	99.89	2.92
	N7...H21-N2 <sup>t</sup>	99.44	3.01	99.80	2.98	99.66	2.99
	O6...H1-N1 <sup>c</sup>	93.39	3.15	99.69	3.00	95.34	3.11
	N7...H21-N2 <sup>c</sup>	99.43	2.97	99.59	2.94	99.49	2.97
	O6...H1-N1 <sup>b</sup>	99.12	3.01	99.30	2.98	99.21	3.00
	N7...H21-N2 <sup>b</sup>	98.96	2.97	99.50	2.97	99.31	2.98
G4 (APC <sup>b</sup> )	O6...H1-N1 <sup>t</sup>	98.23	2.95	99.38	2.94	98.66	2.96
	N7...H21-N2 <sup>t</sup>	96.30	3.04	99.20	2.98	98.35	3.01
	O6...H1-N1 <sup>c</sup>	97.40	3.09	98.00	3.09	97.78	3.10
	N7...H21-N2 <sup>c</sup>	99.70	2.96	99.86	2.94	99.79	2.95
	O6...H1-N1 <sup>b</sup>	99.53	2.97	99.89	2.94	99.74	2.96
	N7...H21-N2 <sup>b</sup>	99.70	3.00	99.90	2.98	99.81	2.98
G4 (APC <sup>c</sup> )	O6...H1-N1 <sup>t</sup>	99.88	2.93	99.95	2.91	99.92	2.93
	N7...H21-N2 <sup>t</sup>	98.65	2.99	99.83	2.98	99.76	2.99
	O6...H1-N1 <sup>c</sup>	95.23	3.15	97.74	3.08	99.45	3.12
	N7...H21-N2 <sup>c</sup>	99.46	3.00	99.89	2.94	99.68	2.97
	O6...H1-N1 <sup>b</sup>	99.80	2.96	99.91	2.93	99.87	2.95
	N7...H21-N2 <sup>b</sup>	99.77	2.99	99.90	2.97	99.84	2.98

注:<sup>1</sup> 上标字母 t, c 和 b 分别表示氢键位于顶层(top), 中间层(central)和底层(bottom)G-四分体平面中。

### 2.3 APC 分子对端粒 G4 结构的稳定作用

Hoogsteen 分子内氢键 O6...H1-N1 和 N7...H21-N2 是维持 G-四分体平面结构稳定的关键作用力,其状态可用以表征 G4 结构的稳定性<sup>[26]</sup>。表 1 列出了结合 APC 分子前后端粒 G4 各 G-四分体中的 Hoogsteen 氢键参数,可以发现所有氢键在整个分子动力学过程中的平均时长均高于 94.27%,表明端粒 G4 结构



的稳定性。APC<sup>c</sup>分子的结合将顶层 G-四分体 O6...H1-N1 和 N7...H21-N2 的平均时长分别从未结合时的 94.47%和 94.67%提升到 99.89%和 99.66%,APC<sup>b</sup>分子则将底层 G-四分体 O6...H1-N1 和 N7...H21-N2 的平均时长分别由 99.15%和 99.56%提升到 99.74%和 99.81%,表明 APC 分子提升了其所结合区域的稳定性。双分子 APC 的结合则将端粒 G4 所有 G-四分体 O6...H1-N1 和 N7...H21-N2 的平均时长提升至 99.45%和 99.68%以上,同时将 Hoogsteen 氢键时长的最小值由未结合状态的 89.28%提升至 95.23%,从而从整体上提高了端粒 G4 结构的稳定性。

结合自由能可用于评估配体分子与蛋白质或核酸等生物大分子之间的亲和能力<sup>[31]</sup>。表 2 列出了应用 MM/GBSA 方法计算得到的 APC 分子与端粒 G4 之间的结合自由能,可以发现负电性 APC 分子与端粒 G4 负电性磷酸骨架之间的静电排斥作用使得静电相互作用不利于二者的结合。同时,APC 分子与端粒 G4 的结合是一个体系自由度减少的过程,因此熵变是对结合自由能的不利贡献。极性和非极性溶剂化作用以及范德华相互作用是 APC 分子与端粒 G4 稳定结合的主要贡献者,其中极性和非极性溶剂化能恰能大致抵消静电作用能对结合的不利贡献,而熵变对结合自由能的贡献又大致相同,这就使得范德华作用能的大小实质上决定了 APC 分子与端粒 G4 亲和力的强弱(表 2)。

比较不同结合模式下结合自由能的大小可知,当 APC 分子以 1:1 和 2:1 比例与端粒 G4 结合时均是堆积于顶层 G-四分体外侧的结合方式(APC<sup>c</sup>和 APC<sup>2c</sup>)的能量上更为有利(表 2),这和非共价键相互作用和分子间氢键的分析结果一致。同时,在以 2:1 比例结合时 APC<sup>2c</sup>和 APC<sup>2b</sup>均较 1:1 比例结合时的 APC<sup>c</sup>和 APC<sup>b</sup>在能量上更为有利,表明第二个 APC 分子的结合不仅能更全面的提升端粒 G4 的稳定性(表 1),同时也提升了两个 APC 分子各自与端粒 G4 的亲和力(表 2)。

表 2 平行构象端粒 G-四链体与 APC 分子的结合自由能

配体	能量组分 <sup>1</sup> /kcal·mol <sup>-1</sup>					
	$\Delta E_{\text{ele}}$	$\Delta E_{\text{vdW}}$	$\Delta G_{\text{GB,sol}}$	$\Delta G_{\text{np,sol}}$	$-T\Delta S$	$\Delta G_{\text{bind}}$
APC <sup>c</sup>	362.64 ± 6.18	-62.54 ± 2.76	-349.38 ± 5.54	-5.66 ± 0.25	27.34 ± 8.05	-27.59
APC <sup>b</sup>	380.85 ± 5.51	-61.47 ± 2.38	-366.97 ± 5.02	-5.29 ± 0.19	26.17 ± 6.57	-26.71
APC <sup>2c</sup>	439.95 ± 5.04	-67.52 ± 3.26	-425.49 ± 4.67	-6.22 ± 0.31	26.76 ± 7.58	-32.52
APC <sup>2b</sup>	448.03 ± 6.32	-61.74 ± 2.04	-433.43 ± 5.85	-5.28 ± 0.18	25.09 ± 7.36	-27.33

注:<sup>1</sup> 结合自由能  $\Delta G_{\text{bind}}$ 所包含的能量组分:静电作用能  $\Delta E_{\text{ele}}$ ,范德华作用能  $\Delta E_{\text{vdW}}$ ,溶剂化作用能极性部分  $\Delta G_{\text{GB,sol}}$ ,溶剂化作用能非极性部分  $\Delta G_{\text{np,sol}}$ 和熵的贡献  $-T\Delta S$

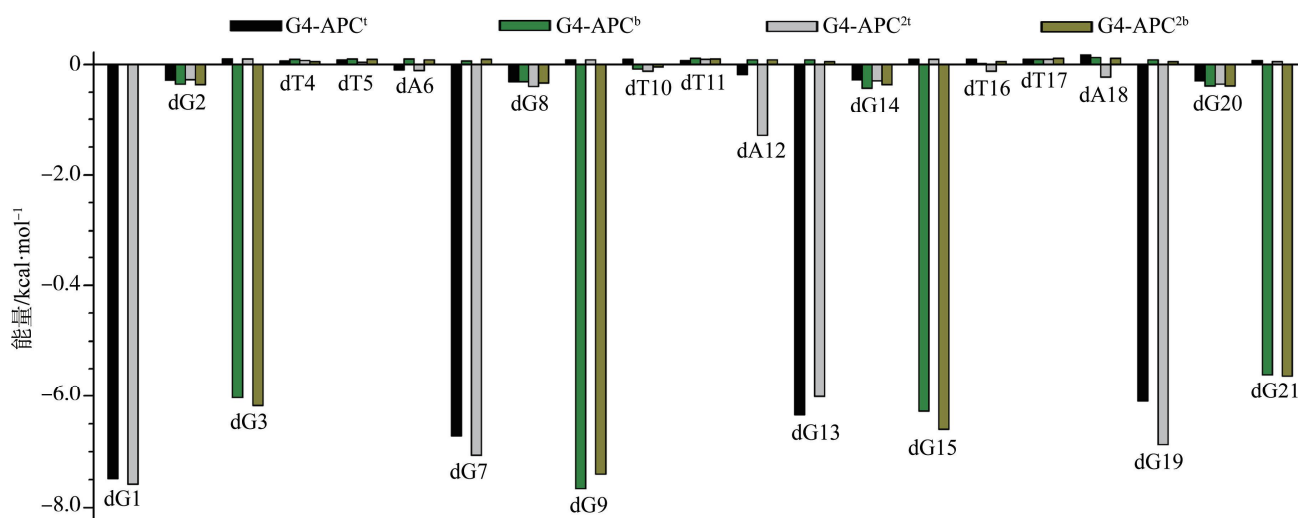


图 7 基于碱基的结合能分解

进一步的结合能分解分析发现,构成顶层 G-四分体的碱基 dG1,dG7,dG13 和 dG19 对 APC<sup>c</sup>和 APC<sup>2c</sup>与端粒 G4 的结合贡献最大,而构成底层 G-四分体的碱基 dG3,dG9,dG15 和 dG21 对 APC<sup>b</sup>和 APC<sup>2b</sup>与端粒 G4 的结合贡献最大(图 7)。需要注意的是,碱基 dA12 对 APC<sup>2c</sup>的结合也有重要贡献(-1.29 kcal·mol<sup>-1</sup>),但是其对 APC<sup>c</sup>的结合则贡献微弱(-0.18 kcal·mol<sup>-1</sup>)(图 7),这也是与端粒 G4 结合时 APC<sup>2c</sup>较 APC<sup>c</sup>亲和力更强的重要原因之一。结合能分解与非共价键相互作用分析的结果完全一致,进一步增强了当前研究结



果的可靠性。

## 2 结论

本文应用分子对接和分子动力学方法对具备高选择性的负电性 APC 分子与人端粒平行构象 G4 的作用机制进行了系统研究。研究发现,APC 分子倾向于以平行堆积的方式结合于端粒 G4 顶层和底层 G-四分体平面的外侧。相比底层堆积的结合方式,APC 分子结合于顶层 G-四分体时除可以形成广泛的范德华相互作用外,还能够与 dG1 碱基形成稳定的分子间氢键相互作用,使得顶层堆积的结合方式在能量上更为有利。APC 分子以 2:1 比例结合时能够更全面的提升端粒 G4 结构的稳定性,并且每个 APC 分子与端粒 G4 的亲合力均较单独结合时更强。结合自由能计算表明静电相互作用不利于 APC 分子与端粒 G4 的结合,而范德华作用能作为二者结合的重要贡献因素,在本质上决定了 APC 分子在不同结合位点亲和力的强弱。这些发现在原子和分子水平上揭示了 APC 分子与端粒平行构象 G4 的作用机制,对新型特异性端粒 G-四链体稳定剂的开发具有重要意义。

## 参 考 文 献

- [1] XU Y,ISHIZUKA T,KURABAYASHI K,et al.Consecutive formation of G-quadruplexes in human telomericoverhang DNA;a protective capping structure for telomere ends [J].*Angewandte Chemie*,2009,121(42):7973-7976.
- [2] SIDDIQUI-JAIN A,GRAND C L,BEARSS D J,et al.Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription [J].*Proceedings of the National Academy of Sciences of the United States of America*,2002,99(18):11593-11598.
- [3] COLLIE G W,HAIDER S,NEIDLE S,et al.A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex [J].*Nucleic Acids Research*,2010,38(16):5569-5580.
- [4] BOCHMAN M L,PAESCHKE K,ZAKIAN V A.DNA secondary structures:stability and function of G-quadruplex structures [J].*Nature Reviews Genetics*,2012,13(11):770-780.
- [5] PUNNOOSE J A,MA Y,HOQUE M E,et al.Random formation of G-quadruplexes in the full-length human telomere overhangs leads to a kinetic folding pattern with targetable vacant G-tracts [J].*Biochemistry*,2018,57(51):6946-6955.
- [6] WANG Q,LIU J,CHEN Z,et al.G-quadruplex formation at the 3' end of telomere DNA inhibits its extension by telomerase,polymerase and unwinding by helicase [J].*Nucleic Acids Research*,2011,39(14):6229-6237.
- [7] PHAN A T.Human telomeric G-quadruplex;structures of DNA and RNA sequences [J].*The FEBS Journal*,2010,227(5):1107-1117.
- [8] WU Y,BROSH RM,G-quadruplex nucleic acids and human disease [J].*The FEBS Journal*,2010,277(17):3470-3488.
- [9] HAN H,HURLEY L H.G-quadruplex DNA;a potential target for anti-cancer drug design [J].*Trends in Pharmacological Sciences*,2000,21(4):136-142.
- [10] PARKINSON G N,LEE M P H,NEIDLE S.Crystal structure of parallel quadruplexes from human telomeric DNA [J].*Nature*,2002,417(6891):876-880.
- [11] PHAN A T,PATEL D J.Two-repeat human telomeric d(TAGGGTTAGGGT) sequence forms interconverting parallel and antiparallel G-quadruplexes in solution;distinct topologies,thermodynamic properties,and folding/unfolding kinetics [J].*Journal of American Chemistry Society*,2003,125(49):15021-15027.
- [12] PHAN A T,KURYAVYI V,LUU K N,et al.Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K<sup>+</sup> solution [J].*Nucleic Acids Research*,2007,35(19):6517-6525.
- [13] DAI J,CARVER M,PUNCHIHEWA C,et al.Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K<sup>+</sup> solution;insights into structure polymorphism of the human telomeric sequence [J].*Nucleic Acids Research*,2007,35(15):4927-4940.
- [14] MA Y,OU T,HOU J,et al.9-N-Substituted berberine derivatives;Stabilization of G-quadruplex DNA and down-regulation of oncogene c-myc [J].*Bioorganic & Medicinal Chemistry*,2008,16(16):7582-7591.
- [15] PANDA D,DEBNATH M,MANDAL S,et al.A nucleus-imaging probe that selectively stabilizes a minor conformation of c-MYC G-quadruplex and down-regulates c-MYC transcription in human cancer cells [J].*Scientific Reports*,2015,5:13183.
- [16] SEENISAMY J,REZLER E M,POWELL T J,et al.The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4 [J].*Journal of the American Chemical Society*,2004,126(28):8702-8709.
- [17] DAI J,CARVER M,HURLEY L H,et al.Solution structure of a 2:1 quindoline-c-MYC G-quadruplex;insights into G-quadruplex-interactive small molecule drug design [J].*Journal of the American Chemical Society*,2001,123(44):17673-17680.
- [18] YAKU H,MURASHIMA T,MIYOSHI D,et al.Anionic phthalocyanines targeting G-quadruplexes and inhibiting telomerase activity in

- the presence of excessive DNA duplexes [J]. *Chemical Communications*, 2010, 46(31): 5740-5742.
- [19] LIU H, ZHAO Q, ZHANG T, et al. Conformation selective antibody enables genome profiling and leads to discovery of parallel G-quadruplex in human telomeres [J]. *Cell Chemical Biology*, 2016, 23(10): 1261-1270.
- [20] WANG Z, LIU J P. Effects of the central potassium ions on the G-quadruplex and stabilizer binding [J]. *Journal of Molecular Graphics and Modelling*, 2017, 72: 168-177.
- [21] WANG J, WOLF R M, CALDWELL J W, et al. Development and testing of a general amber force field [J]. *Journal of Computational Chemistry*, 2004, 25(9): 1157-1174.
- [22] MORRIS G M, RUTH H, WILLIAM L, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility [J]. *Journal of Computational Chemistry*, 2010, 30(16): 2785-2791.
- [23] CASE D A, CHEATHAM T E III, DARDEN T, et al. The Amber biomolecular simulation programs [J]. *Journal of Computational Chemistry*, 2010, 26(16): 1668-1688.
- [24] MACHIREDDY B, KALRA G, JONNALAGADDA S C, et al. Probing the binding pathway of BRACO19 to a parallel-stranded human telomeric G-quadruplex using molecular dynamics binding simulation with AMBER DNA OL15 and ligand GAFF2 force fields [J]. *Journal of Chemical Information and Modeling*, 2017, 57(11): 2846-2864.
- [25] JOUNG I S, CHEATHAM T E III. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations [J]. *Journal of Physical Chemistry B*, 2008, 112(30): 9020-9041.
- [26] WANG Z, LI J, LIU J, et al. Molecular insight into the selective binding between human telomere G-quadruplex and a negatively charged stabilizer [J]. *Clinical and Experimental Pharmacology and Physiology*, 2020, 47(5): 892-902.
- [27] DAVID C C, JACOBS D J. Principal component analysis: a method for determining the essential dynamics of proteins [J]. *Methods in Molecular Biology*, 2014, 1084: 193-226.
- [28] MONGAN J. Interactive essential dynamics [J]. *Journal of Computer-Aided Molecular Design*, 2004, 18(6): 433-436.
- [29] CONTRERAS-GARCÍA J, JOHNSON E R, KEINAN S, et al. NCIPLLOT: a program for plotting noncovalent interaction regions [J]. *Journal of Chemical Theory and Computation*, 2011, 7(3): 625-632.
- [30] HUMPHREY W, DALKE A, SCHULTEN K. VMD: visual molecular dynamics [J]. *Journal of Molecular Graphics*, 1996, 14(1): 33-38.
- [31] KOLLMAN P A, MASSOVA I, REYES C, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models [J]. *Accounts of Chemical Research*, 2000, 33(12): 889-897.
- [32] MACHIREDDY B, SULLIVAN H J, WU C. Binding of BRACO19 to a telomeric G-quadruplex DNA probed by all-atom molecular dynamics simulations with explicit solvent [J]. *Molecules*, 2019, 24(6): 1010.
- [33] WEISER J, SHENKIN P S, STILL W C. Approximate solvent-accessible surface areas from tetrahedrally directed neighbor densities [J]. *Biopolymers*, 2015, 50(4): 373-380.