



HIV-1 整合酶与抑制剂 LCA 的结合模式 及抗药性研究

胡建平^{①②} 常 珊^① 陈慰祖^① 王存新^{①*}

(^① 北京工业大学生命科学与生物工程学院, 北京 100022; ^② 乐山师范学院化学与生命科学系, 乐山 614004)

摘要 I 型人体免疫缺陷病毒(HIV-1)整合酶(integrase, IN)是病毒生命周期中一个重要的酶, 也是研究抗 HIV 新药的一个重要靶点. 运用多构象分子对接和分子动力学(molecular dynamics, MD)模拟, 研究了野生型整合酶核心区及 G140S 点突变的突变态整合酶核心区与抑制剂 L-菊苣酸(L-chicoric acid, LCA)的结合模式, 并基于该结合模式探讨了 G140S 突变态整合酶对抑制剂 LCA 的抗药性. 结果表明: LCA 结合到 G140S 突变态整合酶核心区中的位置与结合到野生型整合酶核心区的位置不同, 结合位置的差异导致 LCA 抑制作用的部分丧失; IN 功能 Loop 区的柔性以及 Mg^{2+} 离子与三个关键残基 D64, D116 和 E152 之间的相互作用有助于 IN 发挥生物学功能; G140S 突变态整合酶核心区中的 E152 与 LCA 的排斥作用、K159 与 LCA 结合能力的变弱以及 Y143 指向 IN 的口袋区是产生抗药性的重要原因. 这些模拟结果与实验结果吻合, 可为基于 IN 的抗 HIV 药物分子设计提供一些有用信息.

关键词 分子对接 MD 模拟 HIV-1 IN 药物分子设计

分子动力学(molecular dynamics, MD)模拟是研究生物大分子结构、动力学和热力学性质的一个重要工具. 该方法不但能够给出生物大分子在原子水平上的运动细节, 而且可提供蛋白质分子的位置涨落和构象变化的详细信息, 在微观层次上理解和解释实验, 特别是在用实验手段无法研究或者实验研究花费昂贵的情况下更显示出其优势, MD模拟的结果对于理解生物大分子的结构-功能关系具有重要意义^[1,2]. 分子对接是按照几何互补、能量互补及化学环境互补的原则来评价配体与受体相互作用的好坏, 并找到两个分子之间最佳的结合模式. 目前, 分子对接已成为基于结构的药物分子设计中快速有效地发现先导化合物的首选方法^[3,4].

I 型人体免疫缺陷病毒(HIV-1)是获得性免疫缺

陷综合症(AIDS)的病源. HIV-1 整合酶(integrase, IN)介导逆转录病毒 DNA 整合到宿主细胞 DNA 中, 是 HIV-1 生命周期中必不可少的一个酶^[5], 也是设计抗 HIV 药物的重要靶点. HIV-1 IN 由 288 个氨基酸残基组成, 分子量为 32 kD. 从结构上看, HIV-1 IN 由 N 端结构域(残基 1~49)、核心区结构域(残基 50~212)和 C 端结构域(残基 213~288)组成. N 端结构域含有保守的 HHCC 结构, 并结合了一个 Zn^{2+} 离子, 其功能主要是促进 HIV-1 IN 的多聚化^[6,7]; 核心区结构域含有 6 个 α 螺旋和 5 个 β 折叠, 结构上拥有保守的 DDE 基序残基(D64, D116, E152), 并整合 1~2 个 Mg^{2+} 离子, 主要发挥核酸内切酶和聚核苷酸转移酶功能^[8~10]. C 端结构域保守性稍差, 能与 DNA 发生非特异性结合^[11,12]. 体外定点突变实验证明, 单独的核心区结构域就能

收稿日期: 2006-07-13; 接受日期: 2006-10-27

国家自然科学基金(批准号: 30670497, 30500429)和北京市自然科学基金(批准号: 5072002)资助项目

* 联系人, E-mail: cxwang@bjut.edu.cn

进行去整合反应^[13,14]. 在组织培养时发现LCA能抑制病毒DNA与宿主DNA的整合^[15], LCA分子结构见图 1. 体外实验发现: 如果HIV-1 IN 中 140 号残基发生点突变, 由Gly突变成Ser, 突变后的IN本身的催化活性下降, 而且突变态还会对LCA产生抗药性^[16,17].

目前唯一解析出的HIV-1 IN 与抑制剂 1-(5-氯-吡啶-3-甲酸)-3-四唑基-1,3-丙二酮烯醇(5CITEP, 见图 1)复合物的晶体结构(PDB代码: 1QS4)^[18], 提供了IN与抑制剂结合位点和结合模式的重要信息, 为基于IN的药物分子设计提供了参考. Sotriffer等^[19]用半柔性分子对接的方法获得了HIV-1 IN核心区与包含LCA在内的六种抑制剂小分子的结合模式, 刘春莉等^[20]用线性相互作用能方法(LIE)计算了包含LCA在内的一系列咖啡酰基类抑制剂与HIV-1 IN之间的结合自由能. 但这些研究工作并没有涉及到G140S突变态与LCA的结合模式, 以及G140S突变所引起的抗药机理等重要信息, 在分子对接方面也未考虑蛋白的柔性. 本工作基于Amber力场^[21~25], 用充分考虑配体、受体分子柔性的多构象分子对接方法^[26]探讨了野生型IN及G140S突变态IN与LCA的结合模式, 并用GBSA方法^[27,28]分析野生型IN及G140S突变态IN对于结合LCA具有主要作用的残基, 从动力学的角度解释了野生型IN及G140S突变态IN与LCA相互识别与抗药性机理, 得到了与实验一致的结果.

1 材料与方法

1.1 体系搭建

以HIV-1 IN和抑制剂5CITEP复合物晶体结构^[18]为参考(PDB代码: 1QS4)来搭建IN的核心区结构, 保留A链和Mg²⁺, 去除配体 5CITEP、结晶水分子、B链和C链. 依据非复合物晶体结构 1BIS^[29]中B链的同源Loop区补全A链柔性Loop区中 4 个尚未解析的残基 I141, P142, Y143 和N144. 用SYBYL 6.9 程序包中的Biopolymer模块将晶体结构中存在的两个突变残基

K158 和E131 替换为天然的残基F158 和W131, 从而获得野生型HIV-1 IN的核心区结构, 然后用最陡下降法和共轭梯度法分别进行 200 步的能量优化, 得到的体系用WT(wild type)表示. 再用SYBYL 6.9 程序包中的Biopolymer模块将WT中的 140 号残基Gly突变成Ser, 经过优化获得的突变态用G140S表示.

1.2 多构象分子对接

目前, 大多数分子对接方法都不考虑受体的柔性, 配体小分子的柔性也只是部分考虑. 事实上, 在配体与受体相互作用过程中, 受体的构象特别是活性部位的构象会发生一定的变化, 若能考虑受体的柔性, 将会得到与真实的结合方式更为吻合的对接结果. 基于此, 我们采用了多构象分子对接方法, 首先对搭建好的 WT 和 G140S 的两个蛋白体系做 2020 ps 的 MD 模拟(模拟步骤见 1.3 节), 然后每隔 200 ps 取一个构象分别与 LCA 进行对接, 这样 WT 和 G140S 均进行了独立的 10 次对接模拟. LCA 的分子结构用分子图形软件 Chemdraw 构建, 并为 LCA 赋上 Gasteiger-Hückel 电荷, 然后用 SYBYL 软件在 Tripos 力场下用最陡下降法能量优化 1000 步, 接着用共轭梯度法能量优化 1000 步, 收敛条件为能量梯度小于 4.182 kJ·mol⁻¹·nm⁻¹. 优化后的结构作为分子对接所需配体的初始结构.

采用AutoDock3.0 程序包^[26]对LCA与WT及G140S进行多构象分子对接. AutoDock3.0 程序能实现刚性受体与柔性配体的对接, 本文考虑了LCA所有单键的可旋转性, 共 17 个可旋转键. 对接运用 Lamarckian遗传算法与局部能量搜索相结合, 用半经验势函数作为能量打分函数, 对小分子构象和位置进行全局搜索, 从 100 个随机的个体开始全局优化, 最大的能量评价次数为 2.5×10⁶, 最多产生 2.7×10⁵ 个后代, 本地搜索迭代的最大数目为 3000. 受体大分子的格点盒子中心位于Mg²⁺离子上, 盒子大小设为

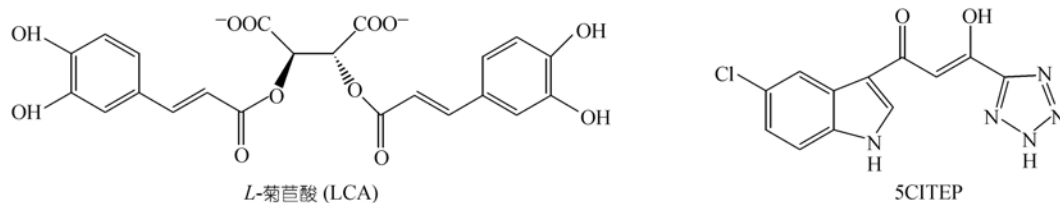


图 1 LCA 和 5CITEP 的分子结构

3×3×3 nm³, 格点间距为 0.0375 nm. 每次对接实验产

生 100 个对接结果. 从所有对接构象中选取对接能最

低的前 10 位构象进行分析, 结果发现结合位置差异不大, 将对接能最低的构象定为最终对接结果. 其中 WT 与 LCA 的对接结果用 WT_LCA 表示, G140S 与 LCA 的对接结果用 G140S_LCA 表示.

1.3 分子动力学模拟

分别对 WT, G140S, WT_LCA 和 G140S_LCA 四个体系进行 MD 模拟. 模拟采用 AMBER 程序 [25,30], 采用的力场是 AMBER 力场 [21~25], 该力场中蛋白的力场参数数据均基于实验值拟合 [21,22]. 模拟中, 用去头八面体盒子给溶质外围加上 0.8 nm 的水分子层, 溶剂采用 TIP3P 水模型 [31]. MD 模拟之前, 对体系分别进行两次能量优化, 首先约束溶质(约束力常数为 $2.09 \times 10^5 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$), 用最陡下降法优化 250 步, 共轭梯度法优化 750 步. 去约束后再进行 250 步最陡下降法优化, 2500 步共轭梯度法优化. 收敛条件为能量梯度小于 $4.182 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-1}$.

MD 模拟分为两步: 首先进行 20 ps 的约束溶质分子 MD 模拟(约束力常数为 $4.18 \times 10^3 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-2}$), 温度从 0 K 逐步升高到 300 K, 接着进行 2000

ps 的无约束恒温 300 K MD 模拟. 模拟中, 用 SYBYL 图象显示软件实时跟踪体系的构象, 采用 SHAKE 算法 [32] 约束键长, 非键相互作用的截断半径设为 1.2 nm, MD 模拟的积分步长为 2 fs.

1.4 能量分解

能量分解采用 GBSA (generalized born/surface area) 方法 [27,28], 基本思想是把受体中每个残基对结合底物的能量贡献近似分为用分子力学方法计算的真空下分子内能, 广义波恩(generalized born, GB) 模型 [33~35] 计算的极性溶剂化能, 和 LCPO 模型 [36,37] 计算的非极性溶剂化能, 并且把能量分解到残基的主链原子和侧链原子上. 通过能量分解可以考察蛋白中主要残基对于结合底物分子所提供的贡献.

2 结果与讨论

2.1 单体 MD 轨迹比较分析

由图 2 可见, 体系 WT 和 G140S 总的势能在 100 ps 后趋于平衡, 体系 WT 总势能平均值为 $-8.98 \times 10^5 \text{ kJ/mol}$, 体系 G140S 的总势能平均值为 -8.99×10^5

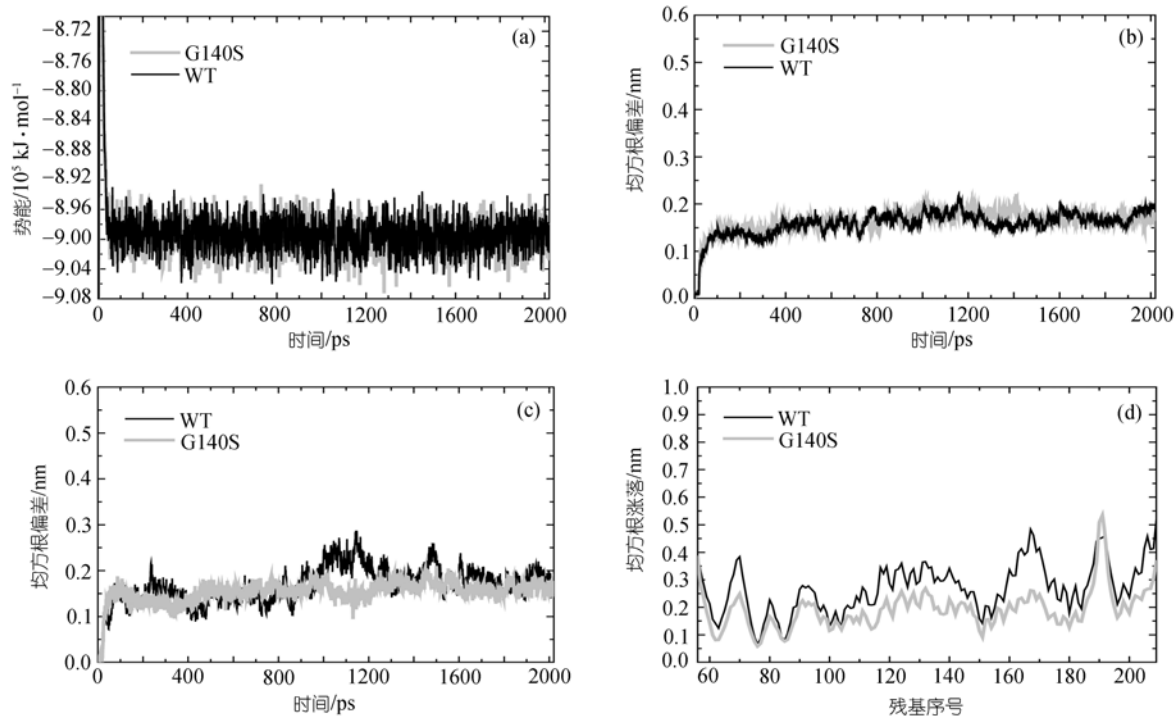


图2 WT 和 G140S 的动力学轨迹比较分析

(a) 体系势能随时间的变化; (b) 体系 C_{α} 原子的 RMSD 随时间的变化; (c) 体系功能 Loop 区 C_{α} 原子的均方根偏差随时间的变化; (d) 体系 C_{α} 原子的均方根涨落

kJ/mol, 涨落范围均小于 1%, 400 ps 后两个体系的 C_{α} 原子相对于初始构象的均方根偏差(RMSD)趋于稳定,

在 0.18 nm 附近涨落. 图 2(c)给出了实验上证明能与病毒DNA结合的功能Loop区(138~149)^[38,39]C α 原子的RMSD随时间的变化, 与图 2(b)比较可知, 突变后体系总的RMSD变化不大, 但是WT的功能Loop区(138~149)的RMSD比体系G140S大, 表明WT的功能Loop区表现出更大的运动性. 由图 2(d)可以看出, WT分子的整体以及功能Loop区的均方根涨落(RMSF)均大于体系G140S. 实验^[16,17]提到, 140 位点突变导致IN的活性降低, IN分子的整体以及功能Loop区的柔性随之下降, 这一点在模拟中得到再现.

表 1 列出了体系 WT 和 G140S 中氢键占有率超过 60%的所有氢键对, 可以看出, 体系 WT 中占有率

超过 60%的氢键有 4 个; G140S 中有 9 个, 其中功能 Loop 区参与形成 2 对氢键. 显然突变后的体系氢键总数增加了, 结构更稳定, 而功能 Loop 区参与形成的 2 个氢键直接降低了功能 Loop 区的运动性, 这与图 2(d)所得的结果一致.

2.2 结合模式预测

图 3 给出了多构象分子对接得到的WT与LCA可能的结合模式. 可以看出, 体系WT与LCA距离小于 0.4 nm的接触残基^[41]主要有: D64, C65, T66, H67, D116, Q148, E152, K156, 这与Sotriffer^[19]及刘等^[20]用半柔性分子对接得到的结果基本一致, 但本工作

表 1 WT 和 G140S 中形成的氢键^{a)}

体系	受体原子	供体原子	供体-受体间距离/nm	供体-氢-受体夹角/(°)	占有率 ^{b)} /%
WT	H78-ND1	S81-OG-HG	0.284 $\pm$ 0.01	163.62 $\pm$ 9.14	72.28
	P58-O	H78-NE2-HE2	0.287 $\pm$ 0.01	158.68 $\pm$ 10.66	71.16
	T93-O	T97-OG1-HG1	0.283 $\pm$ 0.01	160.56 $\pm$ 10.97	68.44
	G193-O	R187-NH1-HH11	0.286 $\pm$ 0.01	152.58 $\pm$ 11.16	63.81
G140S	P58-O	H78-NE2-HE2	0.280 $\pm$ 0.01	154.03 $\pm$ 12.66	88.91
	H78-ND1	S81-OG -HG	0.281 $\pm$ 0.01	164.32 $\pm$ 8.65	87.08
	G193-O	R187-NH1-HH11	0.283 $\pm$ 0.01	157.75 $\pm$ 9.85	81.31
	T93-O	T97-OG1-HG1	0.279 $\pm$ 0.01	161.01 $\pm$ 10.84	75.57
	I203-O	T206-OG1-HG1	0.273 $\pm$ 0.01	163.68 $\pm$ 9.53	75.40
	L63-O	Q148-NE2-HE21 ^{c)}	0.284 $\pm$ 0.01	160.14 $\pm$ 10.03	75.35
	I84-O	N184-ND2-HD21	0.286 $\pm$ 0.01	163.58 $\pm$ 8.64	74.43
	P109-O	W61-NE1-HE1	0.286 $\pm$ 0.01	157.63 $\pm$ 11.25	72.15
	Q146-O ^{c)}	Q62-NE2-HE21	0.286 $\pm$ 0.01	159.37 $\pm$ 10.51	60.32

a) WT和G140S中氢键占有率超过 60%的所有氢键对, 按氢键占有率排序, $\pm$ 号前为平均值, $\pm$ 号后为标准偏差; 氢键采用几何判据^[40]: 供体-氢-受体角大于 135°, 供体-受体间距离小于 0.35 nm; b) 氢键占有率指氢键在整个模拟轨迹中出现的几率; c) 功能Loop区的两个残基Q148, Q146参与形成氢键

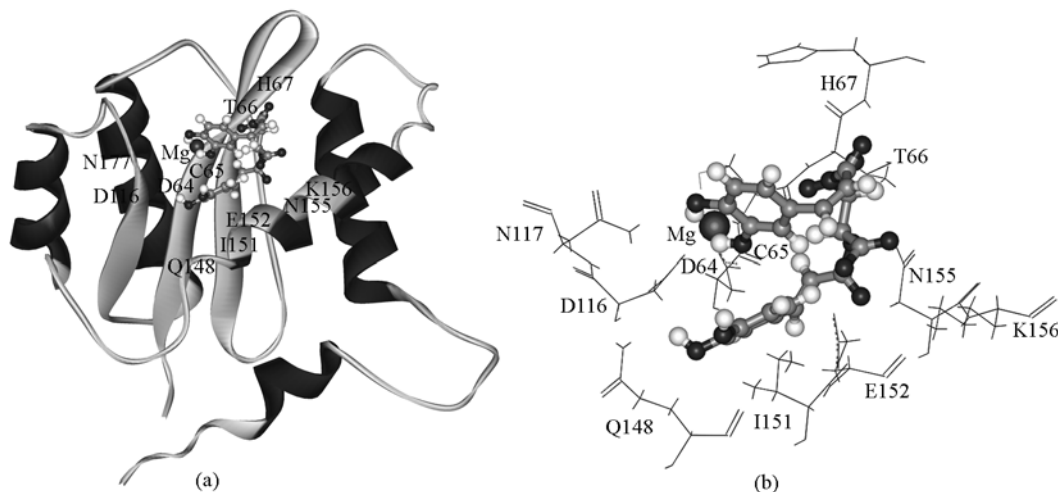


图 3 WT_LCA 的结合模式图

(a) 抑制剂 LCA 和 Mg²⁺离子用棍棒模型表示, 蛋白用带状模式表示, 接触残基分别用单字母表示; (b) 口袋区中接触残基用线条模型表示 WT_LCA 对接结果中的接触残基稍多一些, 这与考虑了 LCA 所有单键的可旋转性有关. LCA 的一个咖

啡酰基的两个酚羟基刚好把 Mg^{2+} 离子螯合于中间, 另外一个咖啡酰基主要指向残基 D64, D116 以及 Q148, 而酒石酸的一个羧酸基指向残基 C65, T66, 另一个羧酸基靠近残基 E152. 图 4 给出了复合物 WT_5CITEP 晶体结构中 WT 与 5CITEP 的结合模式. 对比图 3 和图 4, 发现抑制剂 LCA 结合到 WT 中的位置与 5CITEP 结合到 WT 中的位置类似. LCA 的一个咖啡酰基与 5CITEP 的 5-氯吡啶基取向近似, LCA 的酒石酸基团与 5CITEP 的烯醇式丙二酮取向相近.

迄今为止, 实验证明抑制剂 LCA 能与突变体 G140S 结合, 但具体结合位点还未见报道 [16,17]. 图 5 给出了通过多构象分子对接得到了 G140S 与 LCA 可

能的结合模式. 由图 5 可知, 体系 WT 突变成 G140S 后, LCA 的结合模式有较大变化, 更靠近功能 Loop 区. 具体来说, LCA 咖啡酰基的一个酚羟基与 Mg^{2+} 离子螯合, 指向残基 D64 和 E152 方向, 另外一个咖啡酰基靠近残基 Q148, 而酒石酸的一个羧酸基指向 Loop 区的残基 F139, P142, 另一个羧酸基则指向分子外侧. 体系 WT 与 LCA 的最低对接能为 -24.21 kJ/mol, G140S 与 LCA 的最低对接能为 -22.25 kJ/mol, 突变后导致 IN 结合 LCA 能力的部分下降, 这与生物学实验相吻合 [16,17]. 实验得到的 LCA 抑制体系 WT 的 IC_{50} 值为 82 nmol/L, LCA 抑制体系 G140S 的 IC_{50} 值为 372 nmol/L, 随着 140 位点突变的发生, 抑制作

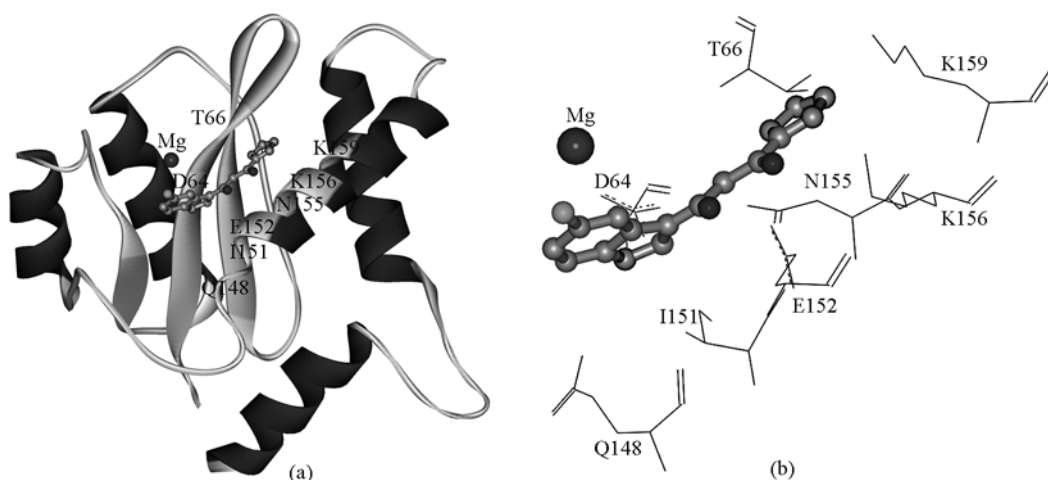


图4 WT_5CITEP 的结合模式图

抑制剂 5CITEP 用球棍模型表示, 其它的图形表示方法见图 3

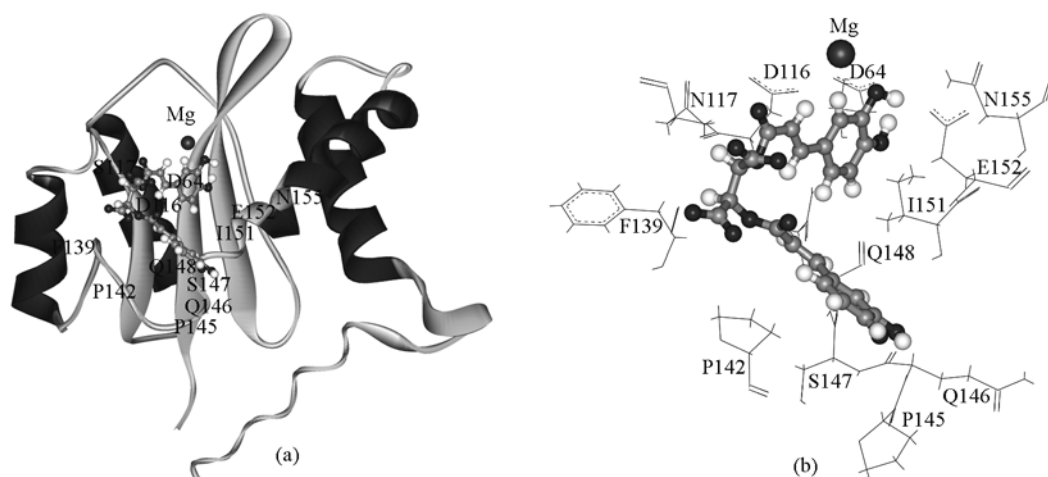


图5 G140S_LCA 的结合模式图

图形表示方法见图 3

用削弱了 4.54 倍. 可见结合位置的变化导致 LCA 抑制作用的部分丧失.

2.3 复合物 MD 轨迹比较分析

体系 WT_LCA 和 G140S_LCA 体系总势能在 100 ps 后趋于平衡, 600 ps 后两个体系的 C_{α} 原子相对于初始构象的 RMSD 达到平衡. 体系 G140S_LCA 的功能 Loop 区 C_{α} 原子的 RMSD 在 0.2 nm 附近涨落, 稍大于体系 WT_LCA 功能 Loop 区的 RMSD (0.17 nm), 即抗药性的产生伴随着 IN 功能 Loop 区的运动性上升, 这点与实验吻合 [16,17].

图 6(a) 给出了抑制剂 LCA 相对于初始构象的 RMSD 随时间的变化. 由图 6(a) 可知, LCA 在体系 G140S_LCA 中不是很稳定, 850 ps 之前维持在 0.25 nm 附近涨落, 850~1000 ps 之间 RMSD 有一个突然的跃迁, 到达 0.35 nm 附近. 通过将 LCA 在 G140S_LCA 体系的 20~850 ps 轨迹中的平均结构与在 1000~2020 ps 轨迹中的平均结构进行叠落发现, LCA 的两个羧酸基的指向没有发生太大变化, 变化较大的是两个咖啡酰基的二酚方向, 其中靠近 Mg^{2+} 离子的酚基发生轻微旋转, 变得更靠近 E152, 而另一个酚基发生大幅度旋转, 与 S140 更为靠近, 具体见图 6(b).

通过分析体系 WT_LCA 和 G140S_LCA 氢键对发现: 体系 WT_LCA 功能 Loop 区残基 N143 与 S147 之间形成氢键, 使得 Loop 成为一个整体, 由于空间位阻及内部牵制限制了功能 Loop 区的运动, WT_LCA 的功能 Loop 区的柔性小于 G140S_LCA, 这与功能 Loop 区 C_{α} 原子的 RMSD 分析得到的结论一

致.

表 2 给出了 IN 活性中心的 DDE 基序残基的质心与 Mg^{2+} 离子之间的距离及标准偏差. 从表 2 可看出, 四个体系模拟过程中 DDE 基序残基相互之间的距离保持稳定, 涨落幅度也较小, 可见 WT 是否突变成 G140S, 以及是否与抑制剂 LCA 结合对三个关键残基影响都不大. WT 突变成 G140S 后, IN 活性降低, 同时伴随着 Mg^{2+} 离子远离 IN 分子的 DDE 基序残基; 由于抗药性, 与 WT_LCA 相比, G140S_LCA 体系活性上升, 此时这个距离又有变小的趋势. 这表明 Mg^{2+} 离子与三个关键残基之间维持较为稳定的相互作用是整合酶发挥生物功能所必须的, 这一结果与实验吻合 [42,43].

实验证明病毒 DNA 在含有三个关键残基的 IN 核心区口袋中与 Y143, E152 结合形成 IN-DNA 复合物 [44,45]. 通过叠落 G140S_LCA 体系 1000~2020 ps 轨迹的平均结构与 WT_LCA 体系 650~2020 ps 轨迹的平均结构可以看出: 三个关键残基的位置基本保持稳定, G140S_LCA 体系中的 LCA 更接近功能 Loop 区. 而 G140S_LCA 体系中功能 Loop 区和 Y143 的侧链均指向口袋区催化中心, 有利于 IN-DNA 复合物的形成, 而 WT_LCA 体系中功能 Loop 区和 Y143 侧链指向分子外侧, 使得病毒 DNA 与功能 Loop 区无法靠近 (具体见图 7). 这与 IN 的功能 Loop 区的 Y143 参与结合病毒 DNA 这一生物功能是对应的.

图 8 给出了用 GBSA 方法计算得到的体系 WT

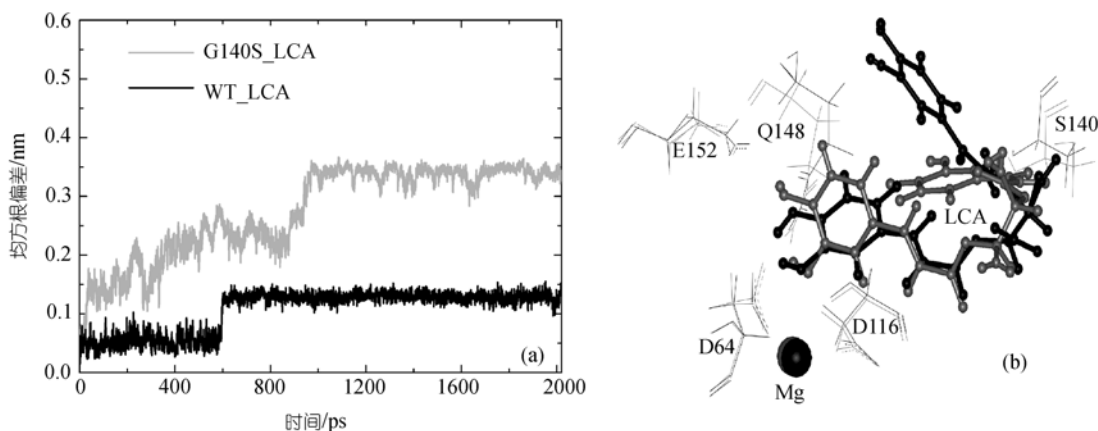


图 6 LCA 的运动性分析

(a) LCA 的 RMSD 随时间的变化; (b) G140S_LCA 体系 20~850 ps 与 1000~2020 ps 的平均结构的叠落, 20~850 ps 的平均结构用黑色表示, 1000~2020 ps 的平均结构用灰色表示, LCA 用球棍模型表示, LCA 附近 0.4 nm 以内的接触残基用单字母棍棒模型表示

表 2 四个体系中 DDE 基序残基质心与 Mg^{2+} 离子的平均距离及标准偏差 (nm)

体系	D64-D116	D64-E152	D116-E152	D64-Mg ²⁺	D116-Mg ²⁺	E152-Mg ²⁺
WT	0.54 ± 0.02	0.84 ± 0.06	1.09 ± 0.08	0.38 ± 0.01	0.44 ± 0.01	0.85 ± 0.07
G140S	0.55 ± 0.04	0.83 ± 0.06	1.08 ± 0.08	0.47 ± 0.01	0.50 ± 0.02	0.86 ± 0.14
WT_LCA	0.55 ± 0.03	0.75 ± 0.05	1.07 ± 0.06	0.47 ± 0.02	0.61 ± 0.05	0.91 ± 0.04
G140S_LCA	0.52 ± 0.02	0.80 ± 0.06	1.07 ± 0.05	0.47 ± 0.01	0.52 ± 0.01	0.86 ± 0.07

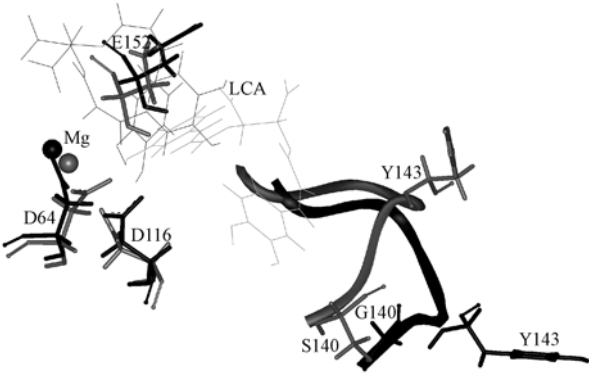


图 7 WT_LCA 体系的平均结构与 G140S_LCA 体系的平均结构叠落

WT_LCA 平均结构用黑色表示, G140S_LCA 平均结构用灰色表示, LCA 用球棍模型表示, DDE 基序残基、140 号残基、Y143 用单字母棍棒模型表示,功能 Loop 区用带状模型表示

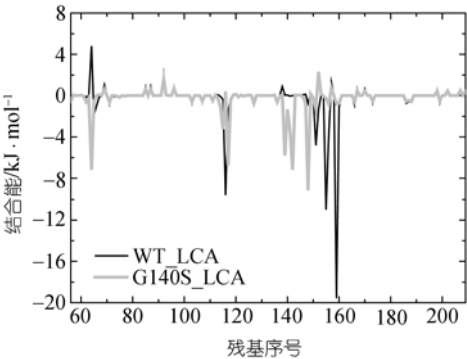


图 8 体系 WT 及 G140S 结合 LCA 时残基的能量贡献

及G140S结合LCA时各个残基的能量贡献. 从图 8 可以看出, WT_LCA体系中, 有利于LCA结合的残基有: C65, D116, N117, E152, N155, K159. 其中残基E152与LCA的结合能为-1.42 kJ/mol. K159 是与LCA结合贡献最大的残基, 结合能达到了-19.57 kJ/mol, 多构象分子对接结果中没有观察到这点(见图 3), 可见多构象分子对接的结果进行MD模拟修正是必要的. 光交联反应等实验提出K159 是参与结合病毒DNA的残基 [44,45], 可见LCA选择性紧密结合K159 是LCA抑制IN的一个重要原因. G140S_LCA体系中有利于结合LCA的残基有: D64, G115, N117, F139, S140, I141,

P142, Q148, 从能量分解来看, 功能Loop区残基对于结合LCA起到较大的作用, K159 对于结合LCA贡献不大, 仅-0.84 kJ/mol, E152 结合LCA的结合能达到 2.31 kJ/mol, LCA对E152 的排斥作用有利于E152 与病毒DNA结合. LCA与G140S的结合位点和LCA与WT的结合位点的差异、G140S_LCA体系中E152 与LCA的排斥作用以及LCA与K159 结合能力的变弱均是G140S对LCA产生抗药性的重要原因.

值得一提的是: HIV-1 IN 突变的140号残基本身对产生抗药性的影响具体从模拟上看有两点: 由图 3、图 5、图 7 可以看出, 随着突变的发生, LCA 与 IN 的结合位点更靠近功能 Loop 区(内含 140 号残基), 由图 8 可以看出, 突变体系 G140S_LCA 中 140 号残基更有利于结合 LCA; 其次, 突变给功能 Loop 区 140 号残基引入极性侧链, 并产生空间位阻, 使得 IN 的构象发生变化(比如功能 Loop 区 Y143 指向发生改变等), 并进一步改变了 LCA 与 IN 的结合模式, 这是产生抗药性最根本的原因.

3 结论

用多构象分子对接和 MD 模拟方法研究了体系 WT 及 G140S 与抑制剂 LCA 的结合模式, 以及体系 G140S 对 LCA 的抗药机制. 相对于 WT, G140S 整体稳定性上升, 体系有更多的氢键形成, 导致功能 Loop 区柔性下降, Mg²⁺离子与 DDE 基序残基距离变大, 这些因素导致了 IN 催化效应下降. LCA 在 WT 中的结合位点与晶体结构中的 5CITEP 位置类似, 而 LCA 与 G140S 的结合位点更靠近功能 Loop 区, 结合位置的变化导致 LCA 抑制作用的部分丧失. 通过对复合物 MD 模拟轨迹分析发现: 功能 Loop 区有较大柔性, Y143 侧链指向 IN 核心区口袋, Mg²⁺离子与 DDE 基序残基之间相互作用稳定, 这些都有利于 IN 发挥其生物功能. G140S_LCA 体系中残基 E152 与 LCA 的排斥作用以及 LCA 与 K159 结合能力的变弱是 G140S 对 LCA 产生抗药性的重要原因. 上述的模拟结果与实验相吻合, 并为进一步基于结构的 HIV-1 整合酶抑制剂设计提供了有用的信息.

参 考 文 献

- 1 Karplus M, Petsko G A. Molecular dynamics simulations in biology. *Nature*, 1990, 347: 631—639[DOI]
- 2 Smith L J, Dobson C M, van Gunsteren W F. Side chain conformational disorder in a molten globule: molecular dynamics simulations of human alpha-lactalbumin. *J Mol Biol*, 1999, 286: 1567—1580[DOI]
- 3 Blaney J M, Dixon J S. A good ligand is hard to find: automated docking methods. *Perspectives in Drug Discovery and Design*, 1993, 1: 301—319[DOI]
- 4 Lengauer T, Rarey M. Computational methods for biomolecular docking. *Curr Opin Struct Biol*, 1996, 6: 402—406[DOI]
- 5 Engelman A, Mizuuchi K, Craigie R. HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. *Cell*, 1991, 67: 1211—1221[DOI]
- 6 Zheng R, Jenkins T M, Craigie R. Zinc folds the N-terminal domain of HIV-1 integrase promotes multimerization, and enhances catalytic activity. *Proc Natl Acad Sci*, 1996, 93: 13659—13664[DOI]
- 7 Lee S P, Xiao J, Knutson J R, Lewis M S, Han M K. Zn^{2+} promotes the self-association of human immunodeficiency virus type-1 integrase *in vitro*. *Biochemistry*, 1997, 36: 173—180[DOI]
- 8 Engelman A, Craigie R. Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function *in vitro*. *J Virol*, 1992, 66: 6361—6369
- 9 Kulkosky J, Jones K S, Katz R A, Mack J P, Skalka A M. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. *Mol Cell Biol*, 1992, 12: 2331—2338
- 10 Polard P, Chandler M. Bacterial transposases and retroviral integrases. *Mol Microbiol*, 1995, 15: 13—23[DOI]
- 11 Engelman A, Hickman A B, Craigie R. The core and carboxyl-terminal domains of the integrase protein of human immunodeficiency virus type 1 each contribute to nonspecific DNA binding. *J Virol*, 1994, 68: 5911—5917
- 12 Vink C, Oude Groeneger A M, Plasterk R H. Identification of the catalytic and DNA-binding region of the human immunodeficiency virus type I integrase protein. *Nucleic Acids Res*, 1993, 21: 1419—1425[DOI]
- 13 Chow S A, Vincent K A, Ellison V, Brown P O. Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science*, 1992, 255: 723—726[DOI]
- 14 Bushman F D, Engelman A, Palmer I, Wingfield P, Craigie P. Domains of the integrase protein of human immunodeficiency virus type 1 responsible for polynucleotidyl transfer and zinc binding. *Proc Natl Acad Sci*, 1993, 90: 3428—3432[DOI]
- 15 Robinson W E Jr, Reinicke M G, Abdel-Malek S, Jia Q, Chow S A. Inhibitors of HIV-1 replication that inhibit HIV integrase. *Proc Natl Acad Sci*, 1996, 93: 6326—6331[DOI]
- 16 King P J, Lee D J, Reinke R A, Victoria J G, Beale K, Robinson W E Jr. Human immunodeficiency virus type 1 integrase containing a glycine to serine mutation at position 140 is attenuated for catalysis and resistant to integrase inhibitors. *Virology*, 2003, 306: 147—161[DOI]
- 17 King P J, Robinson W E Jr. Resistance to the anti-human immunodeficiency virus type 1 compound *L*-chicoric acid results from a single mutation at amino acid 140 of integrase. *J Virol*, 1998, 72: 8420—8424
- 18 Goldgur Y, Craigie R, Cohen G H, Fujiwara T, Yoshinaga T, Fujishita T, Sugimoto H, Endo T, Murai H, Davies D R. Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: a platform for antiviral drug design. *Proc Natl Acad Sci*, 1999, 96: 13040—13043[DOI]
- 19 Sotriffer C A, Ni H, McCammon A. Active site binding modes of HIV-1 integrase inhibitors. *J Med Chem*, 2000, 43: 4109—4117[DOI]
- 20 刘春莉, 李春华, 陈慰祖, 王存新. 用分子模拟方法研究 HIV-1 整合酶与咖啡酰基类抑制剂的相互作用. *物理化学学报*, 2005, 21(11): 1229—1234
- 21 Weiner S J, Kollman P A, Case D A, Singh U C, Ghio C, Alagona G, Profeta S Jr, Weiner P. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J Am Chem Soc*, 1984, 106: 765—784[DOI]
- 22 Weiner S J, Kollman P A, Nguyen D T, Case D A. An all atom force field for simulations of proteins and nucleic acids. *J Comput Chem*, 1986, 7: 230—252[DOI]
- 23 Homans S W. A molecular mechanical force field for the conformational analysis of oligosaccharides: comparison of theoretical and crystal structures of man alpha 1-3man beta 1-4glcna. *Biochemistry*, 1990, 29: 9110—9118[DOI]
- 24 Cornell W D, Cieplak P, Bayly C I, Gould I R, Merz K M Jr, Ferguson D M, Spellmeyer D C, Fox T, Caldwell J W, Kollman P A. A second generation force field for the simulation of proteins and nucleic acids. *J Am Chem Soc*, 1995, 117: 5179—5197[DOI]
- 25 Wang J M, Cieplak P, Kollman P A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules. *J Comput Chem*, 2000, 21: 1049—1074[DOI]
- 26 Morris G M, Goodsell D S, Halliday R S, Huey R, Hart W E, Belew R K, Olson A J. Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem*, 1998, 19: 1639—1662[DOI]
- 27 Kollman P A, Massova I, Reyes C, Kohn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case D A, Cheatham T E. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res*, 2000, 33: 889—897[DOI]
- 28 Wang W, Donini O, Reyes C, Kollman P A. Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. *Annu Rev Biophys Biomol Struct*, 2001, 30: 211—243[DOI]
- 29 Goldgur Y, Dyda F, Hickman A B, Jenkins T M, Craigie R, Davies D R. Three new structures of the core domain of HIV-1 integrase: an active site that binds magnesium. *Proc Natl Acad Sci*, 1998, 95:

- 9150—9154[DOI]
- 30 Wang J, Wolf R M, Caldwell J W, Kollman P A, Case D A. Development and testing of a general Amber force field. *J Comput Chem*, 2004, 25: 1157—1174[DOI]
- 31 Jorgensen W L, Chandrasekhar J, Madura J D, Impey R W, Klein M L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*, 1983, 79: 926—935[DOI]
- 32 Ryckaert J P, Ciccotti G, Berendsen H J C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J Comput Phys*, 1977, 23(3): 327—341[DOI]
- 33 Tsui V, Case D A. Theory and applications of the generalized born solvation model in macromolecular simulations. *Biopolymers: Nucleic Acid Sci*, 2001, 56: 275—291[DOI]
- 34 Simonson T. Macromolecular electrostatics: continuum models and their growing pains. *Curr Opin Struct Biol*, 2001, 11: 243—252[DOI]
- 35 Bashford D, Case D A. Generalized born models of macromolecular solvation effects. *Ann Rev Phys Chem*, 2000, 51: 129—152[DOI]
- 36 Still W C, Tempczyk A, Hawley R C, Hendrickson T. A general treatment of solvation for molecular mechanics. *J Am Chem Soc*, 1990, 112: 6127—6129[DOI]
- 37 Weiser J, Shenkin P S, Still W C. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *J Comput Chem*, 1999, 20: 217—230[DOI]
- 38 Heuer T S, Brown P O. Mapping features of HIV-1 integrase near selected sites on viral and target DNA molecules in an active enzyme-DNA complex by photo-cross-linking. *Biochemistry*, 1997, 36: 10655—10665[DOI]
- 39 Esposito D, Craigie R. Sequence specificity of viral end DNA binding by HIV-1 integrase reveals critical regions for protein-DNA interaction. *EMBO J*, 1998, 17: 5832—5843[DOI]
- 40 Wang C X, Shi Y Y, Zhou F, Wang L. Thermodynamic integration calculations of binding free energy difference for gly 169 mutation in subtilisin BPN. *Proteins*, 1993, 15: 5—9[DOI]
- 41 Kong R, Tan J J, Ma X H, Chen W Z, Wang C X. Prediction of the binding mode between BMS-378806 and HIV-1 gp120 by docking and molecular dynamics simulation. *Biochim Biophys Acta*, 2006, 1764(4): 766—772
- 42 Asante A E, Skalka A M. HIV-1 integrase: structural organization, conformational changes, and catalysis. *Adv Virus Res*, 1999, 52: 351—369
- 43 Wlodawer A. Crystal structures of catalytic core domains of retroviral integrases and role of divalent cations in enzymatic activity. *Adv Virus Res*, 1999, 52: 335—350
- 44 Jenkins T M, Esposito D, Engelman A, Craigie R. Critical contacts between HIV-1 integrase and viral DNA identified by structure based analysis and photo-cross-linking. *EMBO J*, 1997, 16: 6849—6859[DOI]
- 45 Wang J Y, Ling H, Yang W, Craigie R. Structure of a two domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. *EMBO J*, 2001, 20: 7333—7343[DOI]