

HIV-1 整合酶与 L708, 906 抑制剂结合模式及运动性的研究 *

胡建平^{1)**} 刘 崑²⁾ 唐典勇¹⁾ 张元勤¹⁾ 常 珊³⁾

(¹⁾ 乐山师范学院化学与生命科学学院分子设计中心, 乐山 614004; (²⁾ 成都大学药食同源植物资源开发重点实验室, 成都 610106;

(³⁾ 华南农业大学信息学院, 广州 510642)

摘要 前期工作已用分子对接方法获得了 HIV-1 整合酶与 L708,906 抑制剂分子的复合物模型(IN_L708, 906), 现从距离、能量和氢键三个方面详细地分析了 IN_L708, 906 模型中的关键残基. 结果表明, 复合物模型与蛋白质晶体库中整合酶(IN)与 5CITEP 的结合模式相近. 用主成分分析和动力学交叉相关图方法分别研究了 IN_L708, 906 复合物模型和 IN 单体的运动模式及相关性差异. 计算结果显示, L708, 906 抑制剂的结合使得 IN 功能 loop 区残基柔性下降、分子规律性运动的丧失及集团运动相关性的无序增加, 这些可能是酶活性下降的主要因素. 模拟结果将有利于基于芳香二酮酸类的抗 HIV 药物设计.

关键词 整合酶, L708,906, 运动模式, 运动相关性, 药物设计

学科分类号 O641, Q7

DOI: 10.3724/SP.J.1206.2010.00438

整合酶(integrase, IN)是 HIV-1 增殖必不可少的关键酶, 已成为抗 HIV-1 药物研发中的一个重要靶点^[1-3], 而芳香二酮酸类(aryl diketoacids, ADK)化合物能明显抑制 IN 的整合功能^[4-5]. 前期工作已经研究了 12 个 ADK 类抑制剂与 IN 的结合自由能, 并从距离角度初步分析了关键残基, 给出一个有意义的预测模型^[6]. 其中小分子抑制剂 L708, 906 是一个代表性的二酮酸类 IN 抑制先导化合物, L708, 906 的 IC_{50} 是 0.1 $\mu\text{mol/L}$, 对应的结合自由能值为 -36.29 kJ/mol^[7]. 但 L708, 906 与 IN 的详细结合模式及抑制机理尚未见报道, 结合模式尤其关键残基的提取是药物设计的基础.

点突变研究^[8-9]结果表明, 高度保守的 DDE 基序残基(Asp64, Asp116 和 Glu152)与二价金属离子(Mn^{2+} 或 Mg^{2+})结合, 共同构成 IN 的活性中心. Espeseth 小组^[10]对病毒末端序列进行了多次突变实验, 并结合光偶联实验, 确定了 IN 与病毒 DNA 的 LTR 关键残基和碱基, 并且认为 IN 的功能 loop 区(残基 137~149)参与结合病毒 DNA, 并将病毒

DNA 整合到宿主 DNA 中. 目前, 晶体库里有唯一的 IN 与二酮酸类抑制剂 5CITEP 的复合物晶体结构^[11]. 本文从能量、距离及氢键三个角度详细分析了 IN_L708, 906 复合物模型, 并与 IN_5CITEP 复合物结构进行对比. 最后分别用主成分分析(principal component analysis, PCA)方法和动力学交叉相关图(dynamical cross-correlation map, DCCM)方法获得了体系发生功能性慢运动的主要区域及运动的相关性, 这些研究将有利于揭示 L708, 906 抑制 IN 的可能机理. 图 1 给出了 L708, 906 和 5CITEP 分子结构, 前者还提供了分子中有可能参与形成氢键的原子标号.

* 四川省教育厅基金(08ZB054)和四川省中医药管理局科技专项基金(201003)资助项目.

** 通讯联系人.

Tel: 18981384068, E-mail: hujianping@emails.bjut.edu.cn

收稿日期: 2010-08-24, 接受日期: 2011-01-17

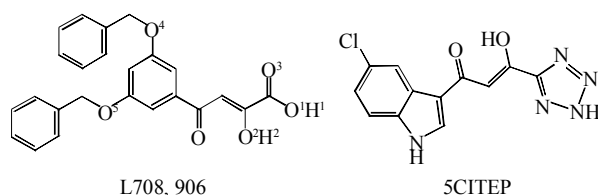


Fig. 1 The structures of L708, 906 and 5CITEP inhibitors

1 计算方法

1.1 MM-GBSA 能量分解

能量分解采用 MM-GBSA (generalized Born/surface area) 方法^[12-13], 基本思路是把每个残基的能量贡献近似分为分子力学 (molecular mechanics, MM) 方法计算的真空下分子内能, 广义波恩 (generalized Born, GB) 模型^[14-15]计算的极性溶剂化能(用 ΔGB 表示), 和 LCPO 模型^[16-17]计算的非极性溶剂化能(用 ΔSA 表示), 并且把能量分解到残基的主链原子和侧链原子上. 另外, 分子内能是由极性部分(用 ΔELE 表示)和非极性部分(用 ΔVDW 表示)组成, 而 LCPO 模型核心是非极性溶剂化能与分子溶剂可接近表面积正相关. 通过 MM-GBSA 能量分解可以考察蛋白质中主要残基对于结合底物分子所提供的贡献.

1.2 PCA 分析

PCA 方法可以从分子动力学 (molecular dynamics, MD) 模拟轨迹中提取出全局性的功能慢运动^[18-21]. 该方法把体系运动模式分为两部分, 即大振幅功能慢运动和几乎没有生物学功能的高频局部振动. PCA 已经成功地应用到很多生物体系的功能性运动研究中^[18-22]. 本文基于 IN 和 IN_L708, 906 体系 MD 模拟平衡后的 6 ns 轨迹(即 4~10 ns), 共 6 000 个构象中蛋白 C α 原子位置信息, 首先剔除 C α 原子的全局转动和平动, 并构建协方差矩阵. 将协方差矩阵对角化获得 154 个本征值及对应的本征向量. 选取本征值最大前 3 位的运动模式, 分析其位置涨落, 即可直观地看出功能性运动的分布和幅度.

1.3 DCCM 分析

生物大分子结构域的协同性运动对于生物大分子与其配体的结合以及发挥其他的特异性生物学功能起着重要作用. 一般来说, 生物大分子体系中能发挥特定功能结构域的原子运动协同性较高^[23-25].

目前, 国际上一般用 MD 模拟采样, 并用统计方法分析轨迹的动力学特征, 并最终获得体系运动的相关性^[26]. 运动的相关程度用交叉相关系数 C_{ij} 大小表示. C_{ij} 计算公式为:

$$C_{ij} = \frac{\langle \Delta \vec{r}_i \cdot \Delta \vec{r}_j \rangle}{\left(\langle (\Delta \vec{r}_i)^2 \rangle \langle (\Delta \vec{r}_j)^2 \rangle \right)^{1/2}}$$

其中 i 和 j 是第 i 个和第 j 个原子, 而 $\Delta \vec{r}_i$ 和 $\Delta \vec{r}_j$ 表示第 i 和第 j 个元素偏离平均位置的向量, 尖括号表示系综平均. C_{ij} 以矩阵的形式给出, 如果 $C_{ij} = 1$, 表示 i 和 j 完全正相关, 运动方向相同; 如果 $C_{ij} = -1$, 表示 i 和 j 完全负相关, 运动方向相反; 如果 $C_{ij} = 0$, 表示 i 和 j 完全不相关, 运动方向垂直. 本工作 C_{ij} 计算是基于 IN 和 IN_L708, 906 体系的 C α 原子位置.

2 结果与讨论

2.1 结合模式

图 2a 给出了 IN 中与 L708, 906 质心距离在 0.4 nm 之内的接触残基^[27]. 为了与 IN_5CITEP 复合物晶体结构对比, 图 2b 从相同的视角给出 IN 中与 5CITEP 相应接触残基. 从图 2 可知, 抑制剂 L702, 906 的接触残基有 14 个, 即 Asp64、Cys65、Thr66、His67、Leu68、Glu92、Asp116、Asn120、Phe121、Gln148、Ile151、Glu152、Asn155、Lys159, 而同为二酮酸类抑制剂 5CITEP 的接触残基有 7 个, 即 Asp64、Cys65、Thr66、His67、Asp116、Ile151、Asn155. 分析这些接触残基发现, L708, 906 和 5CITEP 结合到受体 IN 中的相同位置, L708, 906 的接触残基包括了 5CITEP 的 7 个接触残基, L708, 906 分子体积比 5CITEP 大, 接触残基比 5CITEP 多, 与 IN 结合的相对更牢固. L708, 906 与 IN 结合更紧

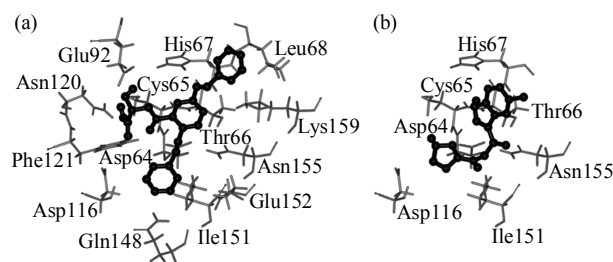


Fig. 2 The binding modes of IN with the inhibitors L708, 906 and 5CITEP

(a) IN_L708, 906 docked mode. (b) IN_5CITEP complex from protein data bank.

密($IC_{50} = 0.1 \mu\text{mol/L}$)可能是其抑制性大于 5CITEP ($IC_{50} = 2.3 \mu\text{mol/L}$)的主要原因^[7]. 接触残基分析可推测, L708, 906 的抑制机理与 5CITEP 相同, 均是结合到 IN 中包含 DDE 基序残基的活性口袋区, 竞争性占有了病毒 DNA 的结合位置.

接触残基分析仅从几何的角度定性给出了结合

部位的残基信息, 为了定量分析 IN 中各个残基对于结合抑制剂 L708, 906 的贡献差别, 我们用 MM-GBSA 方法计算了二者识别的关键残基贡献排序. 表 1 列出了前 20 位有利于 IN 与抑制剂 L708, 906 复合物稳定的关键残基.

Table 1 The energy decomposition of the top twenty key residues concerned with binding inhibitor L708, 906 (kJ/mol)

No.	Residues	ΔVDW	ΔELE	ΔSA	ΔGB	$\Delta Etot$	References
1	Arg166	-5.54	-187.37	-166.88	-19.92	-379.71	The integrase tetramer model constructed by NIH ^[28]
2	Arg107	-5.1	-159.46	-135.62	-34.17	-334.35	The complex model of IN dimer with EBR28 peptide ^[29]
3	Arg199	-5.51	-126.87	-104.98	-66.49	-303.85	The complex model of IN with viral DNA ^[30]
4	Arg187	-3.22	-131.45	-103.06	-63.81	-301.54	The 1EX4 crystal structure refined by Chen <i>et al.</i> ^[31]
5	Asp64	-0.75	-107.87	-90.85	21.27	-178.2	Belonging to the DDE motif residues confirmed by the mutation experiment ^[32]
6	Asn184	-6.79	-90.03	-71.78	-0.91	-169.51	The complex model of IN dimer with EBR28 peptide ^[29]
7	Asn155	-6.16	-84.35	-67.15	-6.04	-163.7	The complex model of IN with LCA molecule ^[27]
8	Asn117	-2.08	-84.16	-61.37	-13.73	-161.34	The complex model of IN with LCA molecule ^[27]
9	Glu157	-4.21	-92.23	-76.79	14.54	-158.69	The drug filtering model by Wielens <i>et al.</i> ^[32]
10	Gln62	-8.46	-80.97	-63.27	-0.53	-153.23	The drug filtering model by Wielens <i>et al.</i> ^[32]
11	Gln137	-5.36	-80.95	-58.05	-7.3	-151.66	Belonging to the DDE motif residues confirmed by photo-cross-linking experiment ^[33]
12	Asp116	-1.49	-78.04	-63.67	-1.48	-144.68	Belonging to the DDE motif residues confirmed by the mutation experiment ^[9]
13	Gln148	-8.3	-75.82	-55.56	-3.5	-143.18	Belonging to the DDE motif residues confirmed by photo-cross-linking experiment ^[33]
14	Gln177	-3.59	-74.73	-50.97	-10.42	-139.71	The complex model of IN dimer with EBR28 peptide ^[29]
15	Glu152	-5.6	-72.89	-58.08	-2.27	-138.84	Belonging to the DDE motif residues confirmed by the mutation experiment ^[9]
16	Asn144	-3.53	-71.26	-49.74	-13.39	-137.92	Belonging to the DDE motif residues confirmed by photo-cross-linking experiment ^[33]
17	Gln146	-1.94	-69.03	-44.77	-16.94	-132.68	Belonging to the DDE motif residues confirmed by photo-cross-linking experiment ^[34]
18	Gln168	0.87	-70.45	-47.21	-14.84	-131.63	The complex model of IN dimer with EBR28 peptide ^[29]
19	Glu138	-7.06	-55.09	-39.52	-15.66	-117.33	Belonging to the DDE motif residues confirmed by Photo-cross-linking experiment ^[33]
20	Asp167	-1.15	-28.17	-11.05	-59	-99.37	The drug filtering model by Wielens <i>et al.</i> ^[32]

对比表 1 和图 2 可知, 有助于结合的关键残基大部分分布在接触残基附近, 有些就是接触残基, 比如 Asp64、Asp116、Gln148、Glu152 和 Asn155. 由表 1 可知, 20 个关键残基中主要可以分为 4 部分: a. 最有利于 IN 与 L708, 906 结合的是 4 个碱性残基 Arg(即 Arg166, Arg107, Arg199 和 Arg187); b. 3 个 DDE 基序残基(Asp64, Asp116 和 Glu152); c. 5 个功能 loop 区残基(Gln137, Gln148, Asn144, Gln146 和 Glu138); d. 4 个其他残基(Asn184, Gln177, Gln168 和 Asp167), Hu 等创建的 IN 二聚

体与小肽 EBR28 复合物模型和 Wielens 等的药物筛选模型均报道这 4 个残基是 IN 抑制剂设计的重要靶点. 值得一提的是, 表 1 中残基 Asn155, Asn117, Glu157 和 Gln62 均属于 DDE 基序残基的相邻残基, 故没有单独列出. 综上, 所有的 20 个关键残基在之前的文献中均被报道过, 这进一步证明了 IN 与 L708, 906 结合模式的精确性及用于抑制剂设计参考的可行性.

氢键是维系蛋白质分子稳定性及其与底物识别的重要作用力. 表 2 列出了两个体系(IN 和

IN_L708, 906)占有率超过 60%的氢键及与抑制剂 L708, 906 形成的所有氢键, 这里氢键采用几何判据^[29]: 供体 - 氢 - 受体角大于 135°, 供体 - 受体间距离小于 0.35 nm. 从表 2 可知, 体系 IN 中稳定的氢键有 6 个, 而 IN_L708, 906 有 8 个稳定氢键, 其中有 3 个功能 loop 区残基(Glu138, Gln146 和 Gln137)参与形成. 抑制剂 L708, 906 与 IN 的结合增强了体系的稳定性并使功能 loop 区柔性下降.

另外, 抑制剂 L708, 906 与 IN 存在的氢键仅有 3 个, 涉及的残基有功能 loop 区残基 Gln148 和 DDE 基序残基 Asp116. 有意思的是 L708, 906 仅是二酮酸的羧酸羟基一侧(O^1-H^1)提供氢, 而空间距离很近的另一侧的烯醇式羟基(O^2-H^2)没有参与形成氢键, 可见在长时间的 MD 过程中, 分子也没有明显的位移或翻转.

Table 2 The hydrogen bonds formed in the IN and IN_L708, 906 systems

Systems	Acceptor	Donor	Distance/nm	Angles/(°)	Frequencies/%
IN	Asn120-O	Thr97-OG1-HG1	0.274 ± 0.011	162.39 ± 9.65	83.82
	His78-ND1	Ser81-OG-HG	0.282 ± 0.009	163.58 ± 8.95	77.90
	Gly193-O	Arg187-NH1-HH11	0.282 ± 0.009	158.83 ± 9.40	75.12
	Ser57-O	His78-NE2-HE2	0.282 ± 0.009	159.01 ± 10.63	71.64
	Pro58-O	Trp61-NE1-HE1	0.284 ± 0.009	150.24 ± 10.82	63.87
	Asn155-O	Thr66-OG1-HG1	0.276 ± 0.011	157.76 ± 10.28	62.17
IN_L708, 906	Asn120-O	Thr97-OG1-HG1	0.270 ± 0.011	162.77 ± 9.01	84.14
	Glu138-O	Thr115-OG1-HG1	0.276 ± 0.012	160.59 ± 9.96	83.26
	Pro109-O	Trp61-NE1-HE1	0.284 ± 0.009	158.65 ± 11.19	81.22
	Gln146-O	Gln62-NE2-HE21	0.285 ± 0.009	162.70 ± 9.29	78.14
	His78-ND1	Ser81-OG-HG	0.285 ± 0.009	165.51 ± 9.26	72.59
	Thr115-OG1	Gln137-NE2-HE22	0.288 ± 0.007	162.00 ± 9.11	68.45
	Ile84-O	Asn184-ND2-HD21	0.287 ± 0.008	164.35 ± 8.48	66.37
	Ile203-O	Thr206-OG1-HG1	0.276 ± 0.011	162.21 ± 10.65	61.97
L708, 906	Gln148-O	L708, 906-O1-H1	0.273 ± 0.011	161.54 ± 9.54	60.24
	Asp116-OD1	L708, 906-O1-H1	0.270 ± 0.012	165.69 ± 9.08	52.82
	Asp116-OD2	L708, 906-O1-H1	0.276 ± 0.011	162.21 ± 10.65	35.65

综合上述接触残基, 关键残基和氢键的分析, 可明显看到 L708, 906 与 IN 的结合模式和 IN_5CITEP 晶体复合物结构相似, 均是结合到 DDE 基序残基附近, DDE 基序残基和功能 loop 区残基对于 IN 与 L708, 906 的识别起着重要作用. 从距离、能量和氢键三个角度的分析均显示, Asp116 和 Gln148 是利于 IN 与 L708, 906 识别的关键残基.

2.2 慢运动模式对比

图 3 给出了体系 IN 和 IN_L708, 906 的整体及功能 loop 区(残基 137~149) C α 原子的方均根偏差(root mean square deviation, RMSD)随时间的变化. 从图 3 可知, 两个体系在 4 ns 后基本达到平衡, 其

中 IN_L708, 906 复合物模型整体 C α 原子的 RMSD 值(0.170 ± 0.017) nm 小于单体 IN(0.215 ± 0.024) nm, 可见复合物的稳定性要好于单体. 在体系 IN 中, 功能 loop 区 C α 原子的 RMSD 值(0.248 ± 0.032) nm 要高于分子整体(0.215 ± 0.024) nm, 即功能 loop 区柔性要大于分子整体的平均水平. 这与文献报道^[10]的功能 loop 区保持充分柔性是 IN 维持催化活性所必需的一致. 在体系 IN_L708, 906 复合物模型中, 功能 loop 区 C α 原子的 RMSD 值(0.161 ± 0.021) nm 与分子整体(0.170 ± 0.017) nm 接近. 可推测, L708, 906 的结合导致 IN 功能 loop 区柔性下降, 进而抑制 IN 活性, 这点与文献^[10]是吻合的.

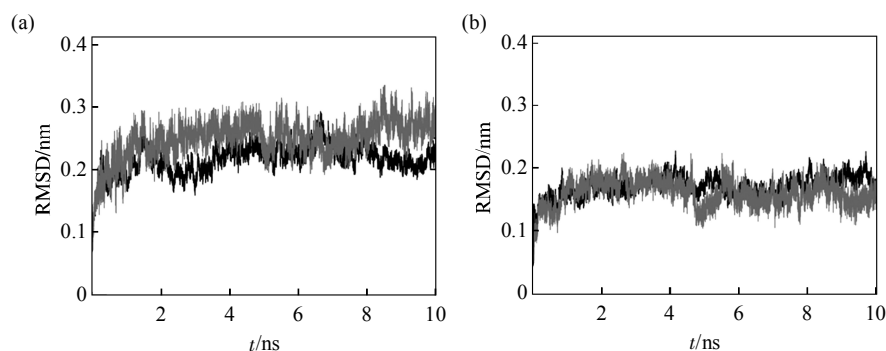


Fig. 3 Comparative molecular dynamics analysis between the IN system and the IN_L708, 906 system

(a) RMSD of the C α atoms of the whole IN system (IN-all) and its functional loop region (IN-loop) versus simulation times. —: IN-all; — —: IN-loop.
(b) RMSD of the C α atoms of the whole IN_L708, 906 system (COM-all) and its functional loop region (COM-loop) versus simulation times. —: COM-all; — —: COM-loop.

为了研究结合抑制剂 L708, 906 之后 IN 所有残基的柔性变化, 对比分析了体系 IN 和 IN_L708, 906 的主成分动力学, 图 4 给出了两个体系最大前 20 个本征值分布. 图 4 中, 两个体系的本征值大小及分布趋势都基本接近. 随后计算最大前几个本征值所占的比重, 方法是将本征值除以所有本征值之和, 权重体现了这个本征值对应的运动模式在整个运动中所占的比例. 计算结果表明, 体系 IN 前 3、4、5 个运动模式占总运动模式的比重分别是 55.23%、61.66%和 66.35%, 而体系 IN_L708, 906 对应的比重分别为 50.06%、56.05%和 60.85%. 由此可见, 体系最大的前 3~5 个本征值对应功能性慢运动可以基本代表体系整个运动.

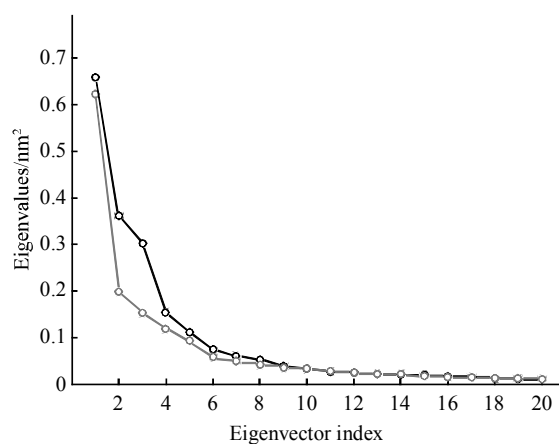


Fig. 4 The distribution of the top 20 eigenvalues from the IN and IN_L708, 906 systems

○—○: IN; □—□: IN_L708, 906.

把两个体系的 MD 模拟轨迹分别投影到各自对应的前 5 个本征向量上, 图 5 给出了这些轨迹投影的运动范围. 体系 IN 前 5 个主要慢运动模式的涨落范围分别是 (0.047 ± 8.07) , (0.043 ± 6.03) , (-0.203 ± 5.417) , (0.047 ± 3.902) 和 (0.001 ± 3.315) nm, 体系 IN_L708, 906 对应的涨落范围是 (-0.141 ± 7.851) , (0.504 ± 4.314) , (0.202 ± 3.847) , (0.067 ± 3.400) 和 (0.046 ± 3.022) nm. 可见, 相对于 IN 来说, IN_L708, 906 运动涨落幅度要小很多. 从图 5 还可知, 从第 3 个本征向量往后, 涨落分布的范围越来越窄, 趋近于高斯分布, 这是体系内部局部快运动的运动特征. 而前 3 个运动模式基本上是功能性慢运动.

图 6 给出了两个体系 C α 原子在前 3 个功能性慢运动中的位置涨落. 图 6 中, 体系 IN 表现出有层次的柔性分布, 其中功能 loop 区(残基 137~149)和 $\alpha 5$ 区与 $\alpha 6$ 区之间的 loop 区(残基 185~195)表现出较大的柔性, 而 DDE 基序残基(Asp64, Asp116 和 Glu152)的运动性较小. 而 C α 原子涨落的层次性分布在 IN_L708, 906 复合物体系中完全丧失, 2 个 loop 区和 DDE 基序残基均表现出中等程度的柔性.

综上, 通过 RMSD 和 PCA 分析可知, IN_L708, 906 复合物比 IN 单体稳定, L708, 906 抑制 IN 活性有可能由于结合导致功能 loop 区柔性的下降. PCA 分析结果还表明, L708, 906 的结合导致 IN 规律性运动的丧失及杂乱运动的增加, 可能也是导致该酶活性降低的主要因素.

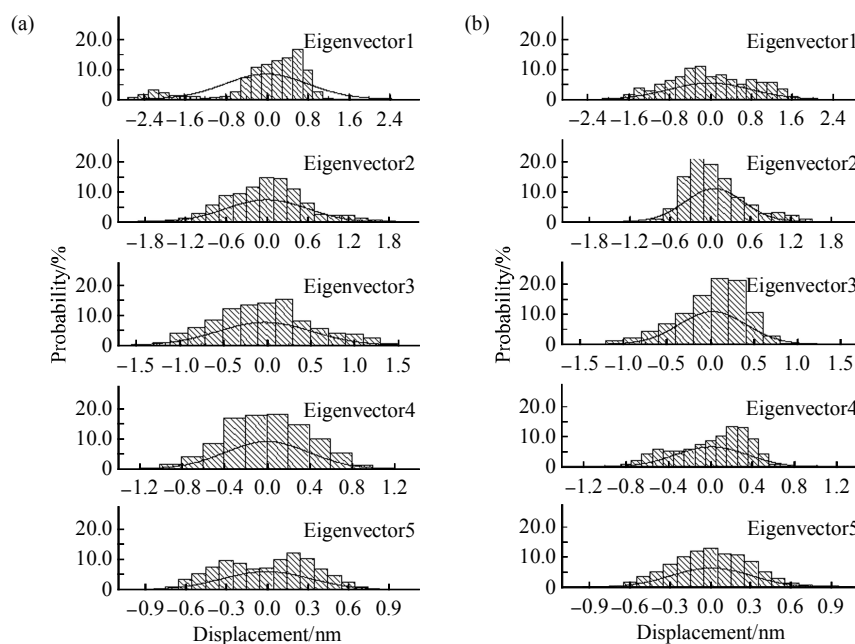


Fig. 5 Displacements for the 5 highest eigenvectors of the following two systems

(a) The IN system. (b) The IN_L708, 906 system.

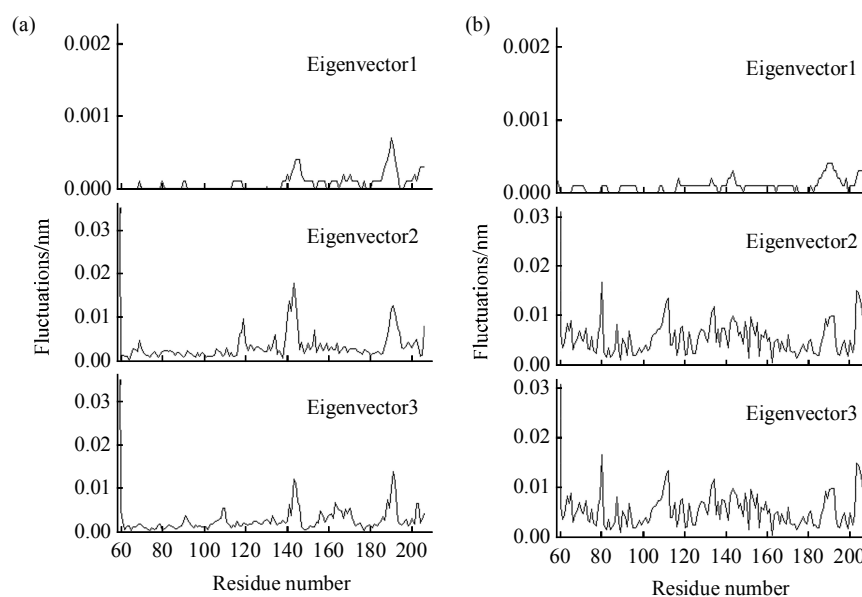


Fig. 6 Atomic fluctuations for the 3 highest eigenvectors of the following two systems

(a) The IN system. (b) The IN_L708, 906 system.

2.3 DCCM 对比

上述分析给出了体系 IN 在结合 L708,906 前后的柔性变化,但是对于各个区域的协同性运动并没有提到.区域协同性运动是生物大分子体系某区域维持生物活性所必需的.图 7 给出了两个体系(IN 和 IN_L708, 906)的 DCCM,图 7 中左上角表示正

相关,而负相关收集在右下角内,相关系数绝对值大于 0.75 的强相关用黑色表示,而相关系数在 0.4 和 0.7 之间的中等相关用灰色表示,剩余部分表示相关系数绝对值小于 0.4 的弱相关.从图 7 可知,体系 IN 的结构功能性分区比较清晰,6 个 α 螺旋区(α_1 区:残基 94~107, α_2 区:残基 119~122,

$\alpha 3$ 区: 残基 124~132, $\alpha 4$ 区残基: 149~165, $\alpha 5$ 区残基: 172~184, $\alpha 6$ 区残基: 196~208)的运动正相关性较高, 但结合 L708, 906 之后, 清晰的结构功能性分区完全弱化, 体系 IN_L708, 906 中出现了更广泛的正相关集团运动. 图 7 中, IN 中两

个 loop 区(残基 137~149, 残基 185~195)与其他区域几乎没有运动相关性, 但在体系 IN_L708, 906 中的对应区域与其他区域尚有一定的运动相关性. 体系 IN 两个 loop 区中存在的一些负相关运动在体系 IN_L708, 906 中完全消失了.

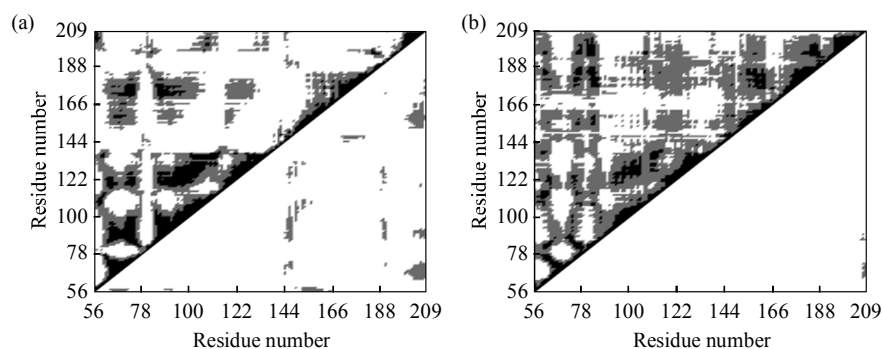


Fig. 7 Calculated DCCM from the MD simulations trajectories of the following two systems

(a) The IN system. (b) The IN_L708, 906 system.

值得一提的是, 该结论与之前整合酶 G140A/G149A 及 T66I/S153Y 突变体抗药性机制研究结果^[34]类似. 综上, 通过 DCCM 分析发现, IN 体系清晰的结构功能分区和功能 loop 区的独立运动是该酶保持较高活性所必需的.

3 结 论

用距离、能量和氢键方法详细分析了 IN 与 L708, 906 的结合模式. 结果表明, DDE 基序和功能 loop 区是 IN 结合 L708, 906 的重要区域, 而 Asp116 和 Gln148 是利于二者识别的关键残基. 对比分析 L708, 906 结合前后的 IN 运动模式差别发现, 复合物模型比单体的运动幅度低且更稳定. L708, 906 的结合导致 IN 规律性运动的丧失、功能 loop 区柔性下降以及体系运动相关性的无序增加均可能是 IN 活性降低的原因.

参 考 文 献

- [1] Flexner C. HIV drug development: the next 25 years. *Nat Rev Drug Discov*, 2007, **6**(12): 959-966
- [2] Michelini Z, Galluzzo C M, Negri D R M, *et al.* Evaluation of HIV-1 integrase inhibitors on human primary macrophages using a luciferase-based single-cycle phenotypic assay. *J Virol Methods*, 2010, **168**(1-2): 272-276
- [3] Ke G T, Li P, Hu J P, *et al.* Docking study of HIV-1 integrase tetramer with different length segments of viral end DNA. *Acta Biophys Sin*, 2010, **26**(10): 902-906
- [4] Van Lunzen J. Treatment of heavily antiretroviral-experienced HIV-infected patients. *AIDS Rev*, 2007, **9**(4): 246-253
- [5] Li H C, Wang C, Sanchez T, *et al.* Amide-containing diketoacids as HIV-1 integrase inhibitors: synthesis, structure-activity relationship analysis and biological activity. *Bioorg Med Chem*, 2009, **17**(7): 2913-2919
- [6] Hu J P, Zhang X Y, Tang D Y, *et al.* Study on the interactions between HIV-1 integrase and Aryl Diketoacid inhibitors with molecular simulation methods. *Acta Chim Sin*, 2009, **67**(19): 2177-2183
- [7] Pommier Y, Johnson A A, Marchand C. Integrase inhibitors to treat HIV/AIDS. *Nat Rev Drug Discov*, 2005, **4**(3): 236-248
- [8] Leavitt A D, Shiue L, Varmus H E. Site-directed mutagenesis of HIV-1 integrase demonstrates differential effects on integrase functions *in vitro*. *J Biol Chem*, 1993, **268**(3): 2113-2119
- [9] Cannon P M, Wilson W, Byles E, *et al.* Human immunodeficiency virus type 1 integrase: effect on viral replication of mutations at highly conserved residues. *J Virol*, 1994, **68**(8): 4768-4775
- [10] Espeseth A S, Felock P, Wolfe A, *et al.* HIV-1 integrase inhibitors that compete with the target DNA substrate define a unique strand transfer conformation for integrase. *Proc Natl Acad Sci USA*, 2000, **97**(21): 11244-11249
- [11] Goldgur Y, Craigie R, Cohen G H, *et al.* Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: a platform for antiviral drug design. *Proc Natl Acad Sci USA*, 1999, **96**(23): 13040-13043
- [12] Kollman P A, Massova I, Reyes C, *et al.* Calculating structures and free energies of complex molecules: combining molecular

- mechanics and continuum models. *Acc Chem Res*, 2000, **33**(12): 889–897
- [13] Wang W, Donini O, Reyes C, *et al.* 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
- [14] Simonson T. Macromolecular electrostatics: continuum models and their growing pains. *Curr Opin Struct Biol*, 2001, **11**(2): 243–252
- [15] Bashford D, Case D A. Generalized born models of macromolecular solvation effects. *Ann Rev Phys Chem*, 2000, **51**: 129–152
- [16] Still W C, Tempczyk A, Hawley R C, *et al.* Semianalytical treatment of solvation for molecular mechanics and dynamics. *J Am Chem Soc*, 1990, **112**(16): 6127–6129
- [17] Weiser J, Shenkin P S, Still W C. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *J Comput Chem*, 1999, **20**(2): 217–230
- [18] Amadei A, Linssen A B M, Berendsen H J C. Essential Dynamics of Proteins. *Proteins*, 1993, **17**(4): 412–425
- [19] Amadei A, Linssen A B M, de Groot B L, *et al.* An efficient method for sampling the essential subspace of proteins. *J Biomol Struct Dynam*, 1996, **13**(4): 615–626
- [20] Abseher R, Nilges M. Efficient sampling in collective coordinate space. *Proteins*, 2000, **39**(1): 82–88
- [21] Capozzi F, Luchinat C, Micheletti C, *et al.* Essential dynamics of helices provide a functional classification of EF-hand proteins. *J Proteome Res*, 2007, **6**(11): 4245–4255
- [22] Blinov N, Berjanskii M, Wishart D S, *et al.* Structural domains and main-chain flexibility in prion proteins. *Biochemistry*, 2009, **48**(7): 1488–1497
- [23] Kormos B L, Baranger A M, Beveridge D L. Do collective atomic fluctuations account for cooperative effects? molecular dynamics studies of the U1A-RNA complex. *J Am Chem Soc*, 2006, **128**(28): 8992–8993
- [24] Kormos B L, Baranger A M, Beveridge D L. A study of collective atomic fluctuations and cooperativity in the U1A-RNA complex based on molecular dynamics simulations. *J Struct Biol*, 2007, **157**(3): 500–513
- [25] Hu J P, Wang C X. Molecular dynamics simulation of HIV-1 integrase dimer complexed with viral DNA. *Chin J Chem*, 2010, **28**: 33–40
- [26] Harte W E Jr, Swaminathan S, Mansuri M. M. *et al.* Domain communication in the dynamical structure of human immunodeficiency virus 1 protease. *Proc Natl Acad Sci USA*, 1990, **87**(22): 8864–8868
- [27] Hu J P, Chang S, Chen W Z, *et al.* Study on the drug resistance and the binding mode of HIV-1 integrase with LCA inhibitor. *Sci China Ser B Chem*, 2007, **50**(5): 665–674
- [28] Karki R G, Tang Y, Burke T R J, *et al.* Model of full-length HIV-1 integrase complexed with viral DNA as template for anti-HIV drug design. *J Comput Aided Mol Des*, 2004, **18**(12): 739–760
- [29] Hu J P, Gong X Q, Su J G, *et al.* Study on the molecular mechanism of inhibiting HIV-1 integrase by EBR28 peptide *via* molecular modeling approach. *Biophys Chem*, 2008, **132**(2–3): 69–80
- [30] Hu J P, Ke G T, Chang S, *et al.* Studies on the binding modes of HIV-1 integrase with viral DNA *via* molecular docking method. *Chem J Chinese U*, 2008, **29**(7): 1432–1437
- [31] Chen J C, Krucinski J, Miercke L J, *et al.* Crystal structure of the HIV-1 integrase catalytic core and c-terminal domains: a model for viral DNA binding. *Proc Natl Acad Sci USA*, 2000, **97**(15): 8233–8238
- [32] Wielens J, Crosby I T, Chalmers D K. A three-dimensional model of the human immunodeficiency virus type 1 integration complex. *J Comput Aided Mol Des*, 2005, **19**(5): 301–317
- [33] 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**(35): 10655–10665
- [34] 胡建平, 唐典勇, 范晶, 等. HIV 整合酶 G140A/G149A 及 T66I/S153Y 突变后的构象变化. *化学学报*, 2010, **68**(15): 1499–1506
- Hu J P, Tang D Y, Fan J, *et al.* *Acta Chim Sin*, 2010, **68**(15): 1499–1506

Study on The Binding Mode and Mobility of HIV-1 Integrase With L708, 906 Inhibitor*

HU Jian-Ping^{1)**}, LIU Wei²⁾, TANG Dian-Yong¹⁾, ZHANG Yuan-Qin¹⁾, CHANG Shan³⁾

¹⁾ Molecular Design Center, Life and Chemistry College, Leshan Normal University, Leshan 614004, China;

²⁾ Key Laboratory of Medicinal and Edible Plants Resources Development, Chengdu University, Chengdu 610106, China;

³⁾ College of Informatics, South China Agricultural University, Guangzhou 510642, China)

Abstract The complex (IN_L708, 906) model of HIV-1 integrase with L708, 906 inhibitor was obtained *via* molecular docking method in previous work. The key residues of IN_L708, 906 complex model were detailedly analyzed from the three perspectives (i.e. distance, energy and hydrogen bond). The results show that the complex model is similar to the IN_5CITEP complex structure from proteins data bank. The difference of the motion modes and correlativity between IN_L708, 906 model and IN monomer was investigated with principal component analysis and dynamical cross-correlation map methods. The computational results indicate that the association with L708, 906 leads to the flexibility decrease of functional loop region, the lose of regular motion, as well as the disordered increase of correlative motion, which may be the main reasons for the activity attenuation of IN. All the simulation results may help the anti-HIV drug design based on the structures of Aryl diketoacids.

Key words integrase, L708, 906, motion mode, motion correlativity, drug design

DOI: 10.3724/SP.J.1206.2010.00438

*This work was supported by grants from The Scientific Research Fund of Sichuan Provincial Education Department (08ZB054) and The Scientific Special Fund of Sichuan Traditional Chinese Medicine Administration(201003).

**Corresponding author.

Tel: 86-18981384068, E-mail: hujianping@emails.bjut.edu.cn

Received: August 24, 2010 Accepted: January 17, 2011