

• 研究论文 •

## 应用 MM/PBSA 方法研究 CDK2 活性口袋内溶剂水分子对 CDK2-配体结合自由能的影响

杨丽君<sup>a,b</sup> 贾若<sup>a</sup> 杨胜勇<sup>\*,a</sup>

(<sup>a</sup> 四川大学华西医院 生物治疗国家重点实验室 成都 610041)

(<sup>b</sup> 西华师范大学化学化工学院 南充 637002)

**摘要** 应用 MM/PBSA 方法研究了 CDK2 活性口袋内溶剂水分子对 CDK2-配体结合自由能的影响。结果表明, 活性口袋内溶剂水分子对 CDK2-配体相互作用自由能有一定的贡献, 其贡献的大小随配体不同而有所差异, 导致这种差异的主要原因是活性位点内溶剂水分子与蛋白残基和配体之间形成了不同的氢键相互作用网络。

**关键词** MM/PBSA; CDK2; 结合自由能; 溶剂水分子; 分子动力学

## Influence of the Solvent Water Molecules at the Active Site of CDK2 on the Binding Free Energy of CDK2-ligand Complexes: an MM/PBSA Study

Yang, Lijun<sup>a,b</sup> Jia, Ruo<sup>a</sup> Yang, Shengyong<sup>\*,a</sup>

(<sup>a</sup> State Key Laboratory of Biotherapy, West China Hospital, West China Medical School, Sichuan University, Chengdu 610041)

(<sup>b</sup> College of Chemistry and Chemical Engineering, China West Normal University, Nanchong 637002)

**Abstract** An MM/PBSA method has been used to investigate the influence of the solvent water molecules at the active site of CDK2 on the binding free energy of CDK2-ligand complexes. The results show that the solvent water molecules at the active site of CDK2 have some contribution to the binding free energy of CDK2-ligand complex. However, the magnitude of contribution depends on the characteristic of the corresponding ligand. The difference of the contribution is mainly due to the different hydrogen bonding interaction network formed among the water molecules, the protein residues and ligand.

**Keywords** MM/PBSA; CDK2; binding free energy; solvent water molecule; molecular dynamics

近年来, 分子对接(Docking)技术已被广泛地应用于计算机辅助药物筛选<sup>[1~3]</sup>, 但与人们的预期相比, 其先导化合物的命中率仍然偏低。导致命中率偏低的原因是多方面的, 典型的包括: 在对接中缺乏对蛋白质柔性的考虑; 评价配体与受体相互作用的打分函数质量不高; 没有有效考虑靶蛋白活性口袋内溶剂水分子的影响等<sup>[4,5]</sup>。在本文中, 我们将主要探讨与靶蛋白活性口袋内

溶剂水分子对分子对接影响的相关问题。

靶蛋白活性口袋内溶剂水分子对基于分子对接的虚拟筛选的影响主要包括两个方面: 第一, 溶剂水分子是否对配体在活性口袋内的位置和取向有影响; 第二, 溶剂水分子是否对配体与受体的相互作用自由能有影响。近年来, 关于第一个问题已有许多研究报道, 例如, Yang等<sup>[6]</sup>发现, 在对接中溶剂水分子对配体的取向有很

\* E-mail: yangsy@scu.edu.cn; Tel.: 028-85164063

Received June 10, 2008; revised August 27, 2008; accepted October 22, 2008.

国家自然科学基金(No. 30772651)和四川省青年基金(No. 08ZQ026-030)资助项目。

大的影响; Pospisil, Verdonk 等<sup>[7,8]</sup>在研究不同的靶蛋白对接时也得出类似的结论. 而对于第二个问题, 即活性口袋内溶剂水分子对靶蛋白-配体相互作用自由能影响的研究则很少<sup>[9]</sup>. 鉴于此, 本文将应用 MM/PBSA 方法研究 CDK2 活性口袋内溶剂水分子对 CDK2 与配体相互作用自由能的影响. 我们选择 CDK2 作为研究对象, 主要基于以下两方面原因: (1) CDK2 是控制细胞增殖的关键激酶之一<sup>[10]</sup>, 其功能的失调被认为是导致许多肿瘤发生的重要原因<sup>[11~13]</sup>, 它已被列为最具前景的肿瘤治疗靶点之一而备受关注<sup>[14~16]</sup>; (2) 目前已有大量的高分辨率的 CDK2-配体复合物晶体结构数据报道<sup>[17]</sup>, 为我们选择合适的模型提供了方便.

## 1 模型与计算方法

### 1.1 CDK2 的结构与模型的选择

CDK2 的活性口袋(这里我们指 ATP 结合位点)位于 CDK2 的激酶域, 该激酶域通常由 300 个氨基酸残基组成, 并折叠成大小不同的两个瓣状结构(N 瓣和 C 瓣), 两瓣之间通过一段肽链连接(称为 Linker). N 瓣主要由 5 个  $\beta$  片层和一个  $\alpha$  螺旋(称为  $\alpha$ C 螺旋)组成, C 瓣主要由  $\alpha$  螺旋构成. N 瓣和 C 瓣之间的裂缝(Cleft)为 ATP 的结合位点, 也即 ATP 竞争性抑制剂的作用位点(图 1).

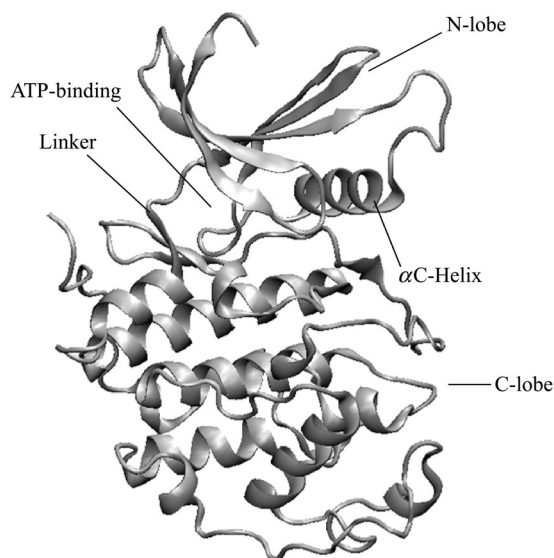


图 1 CDK2 的三维结构  
Figure 1 3D structure of CDK2

目前已有多个 CDK2-配体复合物的晶体结构报道, 我们选取了分辨率较高( $<0.25$  nm), 并且对应配体的  $IC_{50}$ (半数细胞抑制浓度, 可用于表征配体与靶蛋白的结合强弱)实验值已知的 4 个晶体结构, 它们的 PDB 代码分别是: 1ogu, 2bkz, 2c5t 和 2iw6, 其对应的小分子结构

见图 2. 选取这 4 个 CDK2-配体复合物的另一个原因是它们的对应配体的  $IC_{50}$  值较分散, 并且有一个较宽的分布.

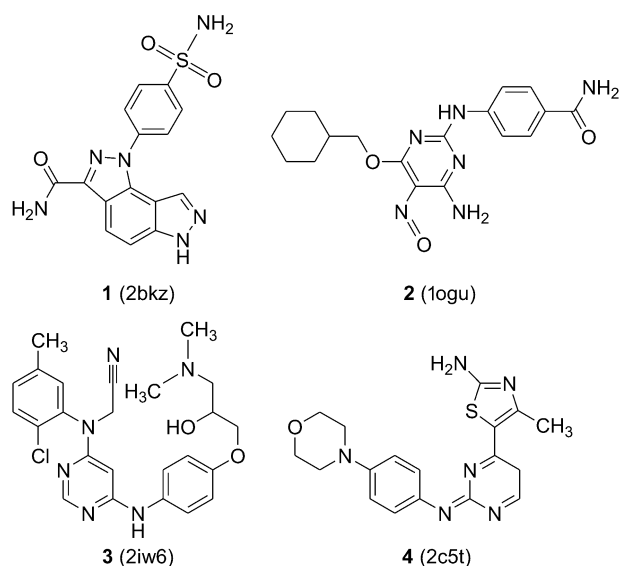


图 2 小分子配体的结构

Figure 2 Chemical structures of small molecules

### 1.2 自由能计算

本文中结合自由能  $\Delta G$  的计算使用 MM/PBSA 方法<sup>[18~21]</sup>, 计算公式如下

$$\Delta G = \Delta E_{MM} + \Delta \Delta G_{PBSA} - T\Delta S \quad (1)$$

其中,  $\Delta E_{MM}$  是真空下分子的内能差, 可通过分子力学方法获得,  $\Delta S$  是熵变, 可通过正则模方法计算获得,  $T$  是绝对温度.  $\Delta \Delta G_{PBSA}$  是溶剂化自由能差, 按下式计算

$$\Delta \Delta G_{PBSA} = \Delta \Delta G_{PB} + \Delta \Delta G_{SA} \quad (2)$$

$\Delta \Delta G_{PB}$  即极性溶剂化自由能差, 其计算是通过求解 Poisson-Boltzmann 方程获得<sup>[22]</sup>,  $\Delta \Delta G_{SA}$  即非极性溶剂化自由能差, 通过计算溶剂可及表面积拟合得到.

活性位点内溶剂水分子的自由能贡献采用下列公式计算:

$$\Delta \Delta G = \Delta G_W - \Delta G_{NW} \quad (3)$$

其中,  $\Delta G_W$  表示活性位点内溶剂水分子存在的情况下的结合自由能,  $\Delta G_{NW}$  表示活性位点内溶剂水分子不存在的情况下的结合自由能.

本文中的分子动力学模拟采用 Amber 9 程序包中的 Sander 及 PMEMD 模块<sup>[23]</sup>. CDK2 力场采用的是 Amber 9 中的 ff99<sup>[24]</sup>, 小分子配体力场采用 gaff<sup>[25]</sup>, 用 Amber 自带的 bcc 电荷拟合程序计算小分子部分电荷, 不存在的范德华非键参数用 Antechamber 程序<sup>[26]</sup>产生. 水溶液采

用显性的 TIP3P 模型, SHAKE 算法用来限制所有含氢的键的伸缩, 模拟步长为 2 fs. 非键相互作用的截断值 (cutoff) 为 0.8 nm. 在运行分子动力学之前, 500 步最陡下降法 (steepest descent method) 紧接着 500 步共轭梯度法 (conjugate gradient) 用来消除分子间的高能碰撞, 然后约束溶质部分而设定所有水分子及抗衡离子可以移动, 对体系升温到 300 K, 让密度在  $1 \text{ g/cm}^3$  平衡. 取消溶质约束, 每个体系运行 1 ns 的分子动力学模拟, 只对后 500 ps 进行采样, 每 10 ps 取一次构象, 利用采集的构象进行自由能计算. 自由能的计算采用 Amber9 程序包中的 mm\_pbsa.pl 模块. 熵的计算则采用正则模分析方法, 使用 NMODE 程序完成.

## 2 结果和讨论

### 2.1 动力学模拟中溶剂水分子进入活性口袋的调控

在蛋白质动力学模拟中, 传统的方法是首先将蛋白质晶体结构中所有结晶水剔除, 然后在其周围重新加入水分子并进行动力学模拟. 在本研究的开始, 我们也采用同样的方法, 首先将所有结晶水包括 CDK2-配体复合物外围以及活性口袋内的结晶水全部剔除, 再将其置入一个长方体水盒中 (使得复合物原子距长方体壁的最短距离为 0.8 nm), 然后按照前面介绍的方法进行动力学模拟. 在动力学模拟过程中发现, 虽然小分子配体与 CDK2 活性位点之间有一定的空隙, 并且实验已测得这些空隙中含有多个溶剂水分子, 但使用 Amber 程序 (取常规的默认参数值) 进行动力学模拟, 溶剂水分子不能自动进入空隙区. 经过多次摸索, 我们发现可通过调整 closeness 参数来控制溶剂水分子的进出. 若 closeness 参数大于或等于 1.0, 则溶剂水分子很难进入空隙, 若小于 1.0, 则水分子较容易进入. 特别将 closeness 参数设为 0.5 左右时, 水分子可以自由进出活性位点, 并且活性位点中水分子的位置与数量大致与晶体结构中结晶水的位置与数量一致.

但为了更逼近真实情况, 在实际计算活性位点内包含溶剂水分子情况下的结合自由能时, 活性位点内的溶剂水分子仍然取自晶体结构.

### 2.2 活性位点内溶剂水分子存在情况下的结合自由能

表 1 列出了活性位点内有溶剂水分子存在情况下的配体与 CDK2 的相互作用自由能  $\Delta G_W$  值及各能量项的贡献. 从表 1 可以看出, 配体 **1** (2bkz) 与 CDK2 的结合自由能最大, 其值为  $-127.82 \text{ kJ/mol}$ ; 配体 **4** (2c5t) 与 CDK2 的结合自由能最小, 其值为  $-87.11 \text{ kJ/mol}$ . 为检验计算得到的 CDK2-配体结合自由能, 我们也在表 1 中给出了实验测得的对应配体的  $IC_{50}$  值, 图 3(a) 给出了

$\Delta G$ -log  $IC_{50}$  之间的相关图, 其相关系数  $R^2$  为 0.920, 可见其  $IC_{50}$  与结合自由能具有很好的相关性. 另外, 表 1 中也列出了动力学模拟复合物的 RMS 均值, 可见各 RMS 均值都较小, 其范围在 0.140~0.168 nm 之间.

**表 1** 活性位点内有溶剂水分子存在时 CDK2-配体结合自由能 (单位:  $\text{kJ}\cdot\text{mol}^{-1}$ ) 及对应  $IC_{50}$  值 (单位:  $\text{nmol}\cdot\text{L}^{-1}$ )

**Table 1** Calculated binding free energies between CDK2 and ligands in the case that the solvent water molecules at the active site are considered together with the experimental  $IC_{50}$  values of the corresponding ligands

Ligand	$\Delta E_{MM}$	$\Delta\Delta G_{PBSA}$	$-T\Delta S$	$\Delta G_W$	RMS	$IC_{50}$
<b>1</b> (2bkz)	-360.11	167.58	64.71	-127.82	0.140	7
<b>2</b> (1ogu)	-380.63	159.05	95.89	-125.69	0.168	34
<b>3</b> (2iw6)	-367.71	174.77	87.86	-105.09	0.152	140
<b>4</b> (2c5t)	-348.95	171.97	89.87	-87.11	0.156	740

### 2.3 活性位点内无结晶水存在情况下的结合自由能

表 2 列出了活性位点内无结晶水存在的情况下的相互作用自由能  $\Delta G_{NW}$ , 从表 2 可以看出, 配体 **1** (2bkz) 与 CDK2 的结合自由能仍然最大, 其值为  $-104.12 \text{ kJ/mol}$ ; 而配体 **4** (2c5t) 与 CDK2 的结合自由能最小, 其值为  $-72.56 \text{ kJ/mol}$ . 图 3(b) 给出了  $\Delta G$ -log  $IC_{50}$  关系图, 其相关系数  $R^2$  为 0.837. 与前面包含溶剂水分子的情况相比, 如果不考虑活性位点内的溶剂水分子, 计算得到的 CDK2-配体结合自由能与配体活性的相关性大大降低.

另外, 活性位点内无结晶水存在的情况下, 动力学模拟中复合物的 RMS 均值也列于表 2, 其均值范围在 0.152~0.181 nm. 与考虑结晶水情况 (见表 1) 相比, 各 RMS 均值稍大, 说明在动力学模拟中, 考虑活性位点内溶剂结晶水所获得的 CDK2-配体复合物结构更接近晶体结构.

**表 2** 活性位点内无溶剂水分子存在情况下的 CDK2-配体相互作用自由能 ( $\text{kJ}\cdot\text{mol}^{-1}$ )

**Table 2** Calculated binding free energies between CDK2 and ligands in the case that not any water molecules at the active site are considered

Ligand	$\Delta E_{MM}$	$\Delta\Delta G_{PBSA}$	$-T\Delta S$	$\Delta G_{NW}$	RMS
<b>1</b> (2bkz)	-366.63	176.98	85.52	-104.12	0.152
<b>2</b> (1ogu)	-331.89	146.13	82.35	-103.41	0.170
<b>3</b> (2iw6)	-348.70	184.25	88.03	-76.41	0.163
<b>4</b> (2c5t)	-368.17	216.11	79.50	-72.56	0.181

### 2.4 活性位点内溶剂水分子对自由能的贡献

前面已经分别计算了活性位点内存在和不存在溶剂水分子的情况下的 CDK2-配体结合自由能  $\Delta G_W$  和  $\Delta G_{NW}$ , 于是, 活性位点内溶剂水分子对自由能的贡献

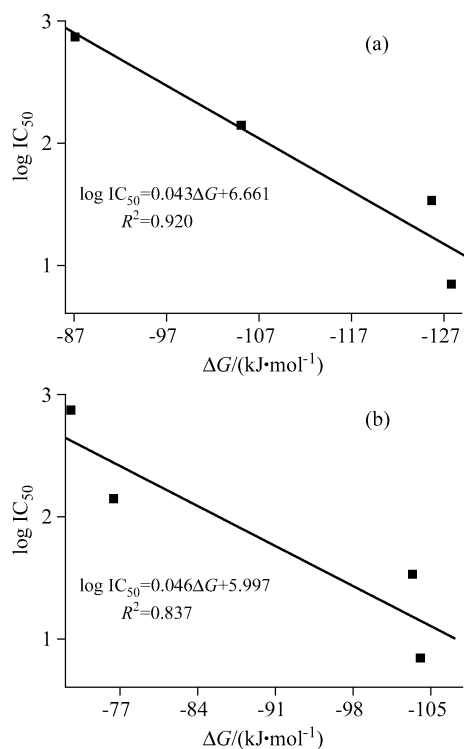


图3  $\log IC_{50}$  与  $\Delta G$  相关图

(a) 包含溶剂水分子; (b) 不包含溶剂水分子

Figure 3 Correlation between  $\log IC_{50}$  and  $\Delta G$

(a) Solvent water molecules were involved; (b) solvent water molecules were not involved

可以按公式  $\Delta\Delta G = \Delta G_W - \Delta G_{NW}$  计算. 计算得到的各能量项的差值,  $\Delta\Delta G$  及  $\Delta RMS$  见表 3. 从表 3 可以看出,  $\Delta\Delta G$  均为负值, 表明活性位点内溶剂水存在有助于降低蛋白质-配体的结合自由能值. 自由能降低最多的是 2iw6, 降低的能量值为  $-28.68$  kJ/mol; 自由能降低最少的是 2c5t, 降低的能量值为  $-14.55$  kJ/mol. 进一步分析表 3 中各能量项可以发现, 分子的内能, 溶剂化自由能和熵等各项能量项都可以对自由能的降低有贡献. 表 3 中  $\Delta RMS$  也均为负值, 这说明有溶剂水时模拟计算得到的结构更接近晶体结构, 因而偏差更小.

表 3 有溶剂水和无溶剂水时计算的各能量项的差值 ( $\text{kJ}\cdot\text{mol}^{-1}$ ) 以及  $\Delta RMS$  (nm)

Table 3 Differences of energy terms ( $\text{kJ}\cdot\text{mol}^{-1}$ ) and  $\Delta RMS$  (nm) between the cases with and without solvent water molecules included at the active site

Ligand	$\Delta\Delta E_{MM}$	$\Delta\Delta\Delta G_{PBSA}$	$-T\Delta\Delta S$	$\Delta\Delta G$	$\Delta RMS$
1 (2bkz)	6.52	-9.40	-20.81	-23.70	-0.012
2 (1ogu)	-48.74	12.92	13.54	-22.28	-0.002
3 (2iw6)	-19.01	-9.48	-0.17	-28.68	-0.011
4 (2c5t)	19.22	-44.14	10.37	-14.55	-0.025

4 个复合物中溶剂水分子对自由能贡献的差异主要

来自于在活性位点内这些水分子之间及它们与 CDK2 和配体之间形成的氢键网络差异, 而其中最重要的差异来自于桥接水分子(即活性位点内既与蛋白质残基之间形成氢键, 又与配体形成氢键的那些水分子). 图 4 给出了 CDK2-配体 1~4 活性位点内桥接水分子形成的氢键网络图. 从图 4 可以看出, 对于 2bkz, 溶剂水分子共介导了 5 个氢键[图 4(a)左], 在无溶剂水时, 有一个新的氢键生成[图 4(a)右], 则净的氢键总数减少了 4 个; 对于 1ogu, 溶剂水分子介导了 3 个氢键[图 4(b)左], 在无溶剂水时, 也没有新的氢键生成[图 4(b)右], 则净的氢键减少数目为 3 个; 对于 2iw6, 溶剂水分子介导了 4 个氢键[图 4(c)左], 在无溶剂水时, 同样也没有新的氢键生成[图 4(c)右], 净的氢键数目减少了 4 个; 最后, 对于 2c5t, 溶剂水分子介导了 3 个氢键[图 4(d)左], 在无溶剂水时, 有一个新的氢键生成[图 4(d)右], 则净的氢键总数减少了 2. 归纳起来, 2iw6 (CDK2-配体 3) 中溶剂水分子的贡献最大, 其氢键数目减少了 4 个, 并且减少的氢键中包含一个强的氢键(键长 0.222 nm), 而 2bkz (CDK2-配体 1) 虽然氢键减少的数目也为 4 个, 但均为较弱的氢键, 2c5t (CDK2-配体 4) 中溶剂水分子的贡献最小, 其氢键减少的数目仅为 2 个. 因此, 溶剂水分子对自由能的贡献大小大致可以用因去掉溶剂水分子而导致的氢键减少的数目来解释.

### 3 结论

采用 MM/PBSA 方法, 研究了 CDK2 活性口袋内溶剂水分子对 CDK2 与配体相互作用自由能的影响. 我们选取了分辨率较高 ( $<0.25$  nm) 并且已有对应配体的  $IC_{50}$  实验值的 4 个晶体结构. 计算结果表明: (1) 当考虑活性位点内的溶剂水分子时, 计算得到的蛋白质-配体相互作用自由能与对应配体的  $IC_{50}$  值具有更好的相关性. 这个结果一方面表明我们计算模型和方法的有效性, 另一方面说明考虑活性位点内溶剂水分子对准确计算蛋白质-配体相互作用自由能是非常重要的. (2) 活性位点内溶剂水分子对 CDK2-配体的结合自由能的贡献  $\Delta\Delta G$  均为负值, 这说明活性位点内溶剂水存在有助于增大蛋白质-配体的结合自由能值. 计算表明, 4 个复合物中溶剂水分子对结合自由能的贡献并不相同, 这主要与不同复合物中, 溶剂水分子与蛋白残基和配体之间形成的氢键相互作用网络差异有关. (3) 当考虑活性位点内的溶剂水分子时, 动力学得到的复合物结构与晶体结构的 RMS 值更低, 这说明溶剂水存在时的复合物结构更接近晶体结构.

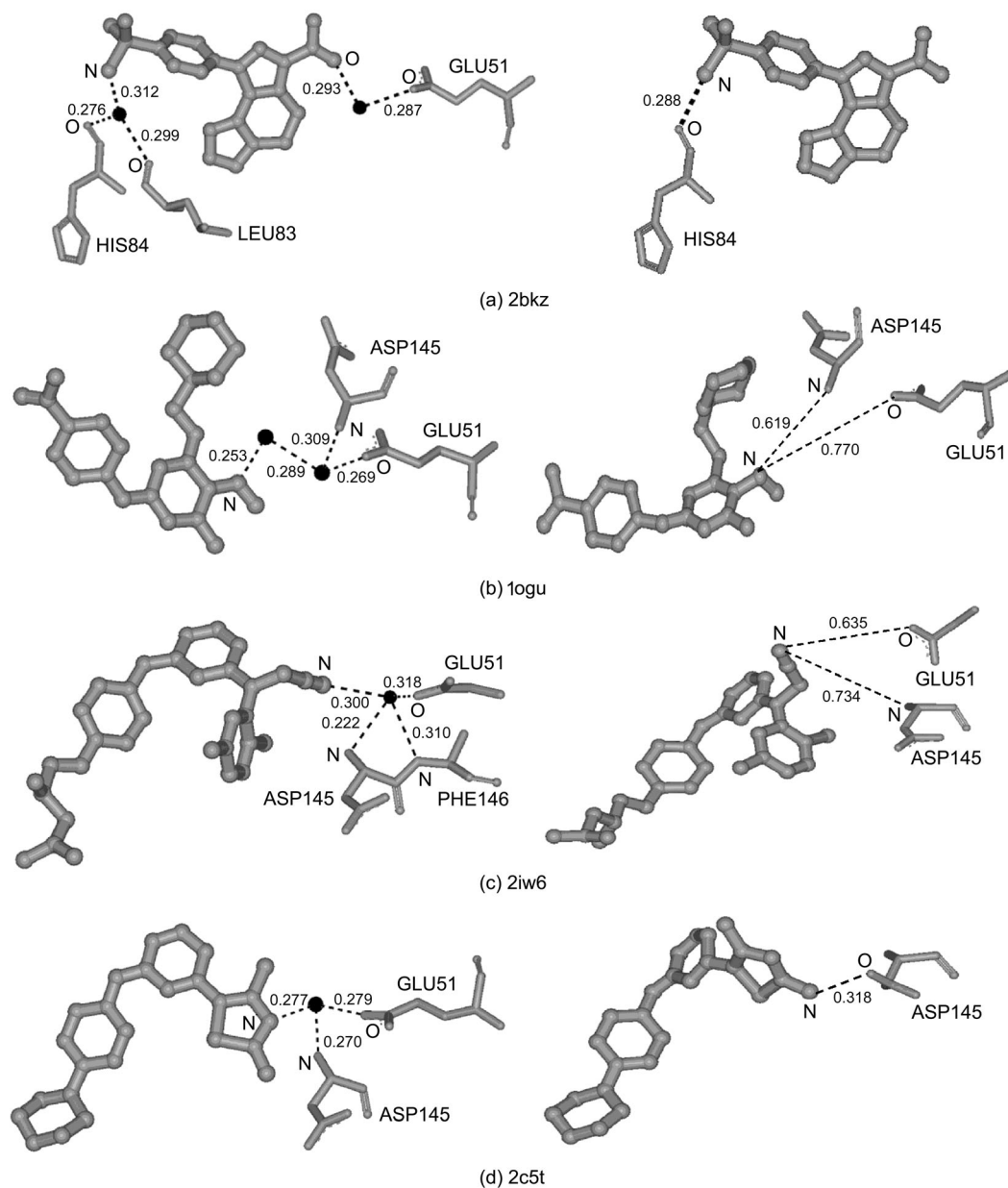


图 4 CDK2-配体复合物中活性位点内形成的与桥接水分子相关的氢键网络(键长单位: nm)

左: 考虑溶剂水分子, 右: 不考虑溶剂水分子

Figure 4 Networks of hydrogen bonding related to bridge water molecules at the active site of CDK2-ligand complexes  
Left: solvent water molecules were considered; right: solvent water molecules were not considered (Bond length in nm)

## References

- Hou, T.; Xu, X. *Curr. Pharm. Des.* **2004**, *10*, 1011.
- Congreve, M.; Murray, C. W.; Blundell, T. L. *Drug Discovery Today* **2005**, *10*, 895.
- Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151.
- Carlson, H. A.; McCammon, J. A. *Mol. Pharmacol.* **2000**, *57*, 213.
- McConkey, B. J.; Sobolev, V.; Edelman, M. *Curr. Sci.* **2002**, *83*, 845.
- Yang, J. M.; Chen, C. C. *Proteins* **2004**, *55*, 288.
- Pospisil, P.; Kuoni, T.; Scapozza, L.; Folkers, G. *J. Recept. Signal Transduction Res.* **2002**, *22*, 141.
- Verdonk, M. L.; Chessari, G.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Nissink, J. W.; Taylor, R. D.; Taylor, R. J. *Med. Chem.* **2005**, *48*, 6504.
- Jia, R.; Yang, L.-J.; Yang, S.-Y. *Chem. Phys. Lett.* **2008**, *460*, 300.
- Evan, G. I.; Vousden, K. H. *Nature* **2001**, *411*, 342.
- Shapiro, G. I. *J. Clin. Oncol.* **2006**, *24*, 1770.
- Slingerland, J.; Pagano, M. *J. Cell. Physiol.* **2000**, *183*, 10.
- Gao, C. Y.; Zelenka, P. S. *BioEssays* **1997**, *19*, 307.
- Harper, J. W.; Adams, P. D. *Chem. Rev.* **2001**, *101*, 2511.

- 15 Garrett, M. D.; Fattaey, A. *Curr. Opin. Genet. Dev.* **1999**, *9*, 104.
- 16 Webster, K. R. *Expert Opin. Invest. Drugs* **1998**, *7*, 865.
- 17 Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
- 18 Srinivasna, J.; Cheatbam, T. E.; Cieplak, P.; Kolman, P.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 9401.
- 19 Kolman, P. A.; Massova, I.; Reyes, C.; Kuhn, 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. *Acc. Chem. Res.* **2000**, *33*, 889.
- 20 Hou, T.-J.; Zhu, L.-L.; Chen, L.-R.; Xu, X.-J. *Acta Chim. Sinica* **2002**, *60*, 1023 (in Chinese).  
(侯廷军, 朱丽荔, 陈丽蓉, 徐筱杰, 化学学报, **2002**, *60*, 1023.)
- 21 Jiang, Y.-J.; Zeng, M.; Zhou, X.-B.; Zou, J.-W.; Yu, Q.-S. *Acta Chim. Sinica* **2004**, *62*, 1751 (in Chinese).  
(蒋勇军, 曾敏, 周先波, 邹建卫, 俞庆森, 化学学报, **2004**, *62*, 1751.)
- 22 Sharp, K. A.; Honig, B. *Annu. Rev. Biophys. Chem.* **1990**, *19*, 301.
- 23 Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A., *AMBER 9*, University of California, San Francisco, **2006**.
- 24 Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
- 25 Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
- 26 Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25*, 247.

(A0806101 Cheng, B.; Zheng, G.)