Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 15 (2005) 3364–3368

Active site binding modes of curcumin in HIV-1 protease and integrase

Opa Vajragupta,[a,*] Preecha Boonchoong,[b] Garrett M. Morris[c] and Arthur J. Olson[c]

[a]*Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Mahidol University, 447 Sri-Ayudhya Road, Bangkok 10400, Thailand*
[b]*Division of Pharmaceutical Chemistry and Technology, Faculty of Pharmaceutical Sciences, Ubon Rajathanee University, Ubon Ratchathani 34190, Thailand*
[c]*The Scripps Research Institute, Molecular Graphics Laboratory, Department of Molecular Biology, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA*

Received 9 March 2005; revised 7 May 2005; accepted 10 May 2005

Abstract—Structure models for the interaction of curcumin with HIV-1 integrase (IN) and protease (PR) were investigated using computational docking. Curcumin was found to bind preferentially in similar ways to the active sites of both IN and PR. For IN, the binding site is formed by residues Asp64, His67, Thr66, Glu92, Thr93, Asp116, Ser119, Asn120, and Lys159. Docked curcumin contacts the catalytic residues adjacent to Asp116 and Asp64, and near the divalent metal (Mg^{2+}). In the PR docking, the curcumin structure fitted well to the active site, interacting with residues Asp25, Asp29, Asp30, Gly27′, Asp29′, and Asp30′. The results suggest that *o*-hydroxyl and/or keto–enol structures are important for both IN and PR inhibitory actions. The symmetrical structure of curcumin seems to play an important role for binding to the PR protein, whereas the keto–enol and only one side of the terminal *o*-hydroxyl showed tight binding to the IN active site.

1. Introduction

Human immunodeficiency virus type 1 (HIV-1) is the etiological agent of acquired immunodeficiency syndrome (AIDS). The unique nature of the replication cycle of HIV-1 provides many potential targets for chemotherapeutic intervention. Following infection, this retrovirus uses three key enzymes to propagate its life cycle. (i) reverse transcriptase (RT), an RNA-dependent DNA polymerase that is used by the virus to transcribe the viral genomic RNA to proviral DNA for incorporation in the host DNA; (ii) integrase (IN), the enzyme responsible for catalyzing both 3′-preprocessing of the viral DNA and its insertion into the host DNA; and (iii) HIV protease (PR), the necessary enzyme for the processing of new virulent viral particles.[1] Combination of inhibitors of the RT and PR are currently the preferred clinical treatment for HIV infection and AIDS. However, targeting RT and PR still does not eliminate the virus from patients, making it necessary to explore other targets. Attention has recently been focused on the HIV IN enzyme because it is one of the earliest steps in the viral lifecycle and there is no native homologous process in the host cell.[2]

Curcuma longa Linn. or turmeric (Zingiberaceae) is a medicinal plant widely cultivated in tropical regions of Asia. Turmeric extract from the rhizomes commonly called curcumin is mainly composed of curcumin, about 75–95%, and a small amount of demethoxy curcumin, and bisdemethoxycurcumin.[3] In addition to its potent antioxidant activity, various related activities in biological systems have been extensively investigated.[4] Curcumin has been shown to exhibit anti-inflammatory,[5] gastrointestinal activity, and inhibitory activity against a variety of protein kinases including protein kinase C and phosphoryl kinase.[6] Curcumin is also reported to inhibit carcinogenesis and cancer growth.[7,8] More recently, curcumin has been reported to inhibit HIV replication and was claimed for anti-HIV-1 and HIV-2 activities in

Keywords: Curcumin; HIV-1; HIV integrase; HIV protease; 5-CITEP; Docking.
* Corresponding author. Tel.: +662 644 8677; fax: +662 644 8695; e-mail: pyovj@mahidol.ac.th

0960-894X/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2005.05.032

a recent patent application.[9] Data showed that curcumin inhibited the replication of HIV-1 integrase protein for both the stand transfer and 3′-processing activities with IC_{50} of 40 and 30 μM, respectively.[10–12] In addition, curcumin and related compounds have also been reported for inhibition of HIV-1 and HIV-2 proteases with IC_{50} of 100 and 250 μM, respectively.[13]

In an attempt to reveal the mechanism in molecular level of curcumin against HIV, the binding mode of curcumin with IN and PR was computed. Docking studies on the IN and PR catalytic core domains have been performed. Moreover, the structure of curcumin consists of two ortho methoxylated phenols linking with β-diketone function, and they are all conjugated; its rigid and electron-rich structure makes it an interesting candidate as a lead compound for further development of new inhibitors.

2. Docking studies

Docking was performed using version 3 of the AutoDock software package running on PC Intel-based Pentium IV, Linux platform. The crystal structure 1QS4.pdb of the core domain of HIV-1 IN (residues C56–Q209) complexed with inhibitor 5CITEP,[14] 1-(5-chloroindol-3-yl)-3-(tetrazolyl)1,3-propanedione enol, was used to prepare IN active binding site. Chain A which co-crystallized with 5-CITEP was selected, while chains B and C were deleted. The ligand within the active site and all waters were removed while a magnesium ion at the active site was maintained throughout the entire process with a charge of +2. The missing residues at positions 141–144 were incorporated from monomer B in crystal structure 1BIS.pdb[15] after superposition of the backbones of 136–140 and 145–149. For PR, the 1HSG.pdb obtained from X-ray structure of the L-735,524 complex was used.[16] All waters and inhibitor were deleted and both chains of protein were assigned for docking.

The structures of both IN and PR proteins were set up for docking as follows. Polar hydrogens were added using the PROTONATE utility distributed with AutoDock. Kollman united atom charges and solvation parameters were added to the final protein file. The grid maps representing the protein in the actual docking process were calculated with AutoGrid. The grids were chosen to be sufficiently large to include not only the active site but also significant portions of the surrounding surface. The dimensions of the grids for IN (1QS4) protein docking were thus $60 \times 65 \times 50$ Å, with a spacing of 0.375 Å between the grid points and the center close to the Cβ of the catalytic residue ASP 64. For PR, the center of the cavity between chains A and B was used as the center of grid, with the grid points of $60 \times 60 \times 60$ Å and the spacing of 0.375 Å. AutoDock 3 with a Lamarckian genetic search algorithm was chosen for all dockings.[17] The optimized AutoDocking run parameters are as follows: the maximum number of energy evaluations was increased to 2,500,000 per run; the maximum number of generation in the genetic algorithm was increased to 100,000; and the number of GA run was 100. All other run parameters were maintained at their default settings.

The docking program AutoDock 3 utilizes a fully flexible ligand in its docking algorithm while targeting a rigid protein. It has been shown to reproduce successfully many crystal structure complexes.[17,18] As far as the search methods are concerned, search efficiency in version 3 is significantly improved by the development of a so-called Lamarckian genetic algorithm, which is the combination of a traditional genetic algorithm with a local search method.[17–19] For all investigation, the resulting ligand orientations from 100 independent docking runs were grouped into clusters of similar configuration with the RMSD of 1.5 Å. 5-CITEP and L-745,524 were re-docked in order to validate the prepared IN and PR active binding sites. The result in Table 1 shows that 98% of docked conformations of L-745,524 grouped into a single cluster, and the docked orientation is close to that of the crystal structure with RMSD of 0.556. The re-docking result of L-745,524 indicated that the prepared HIV-1 PR protein is a good model for docking study of curcumin binding mode.

For HIV-1 IN, a highly favorable bound conformation of 5-CITEP in the active site is found 62 times out of the 100 independent runs (Table 1). In this most favorable docked orientation 5-CITEP was found in the center of the active site, almost perpendicular to the orientation of 5-CITEP found in the crystal structure. The different orientation from the crystal poses results in the high RMSD value of 8.066 Å. The molecule aligns horizontally within the active site providing the large surface area for van der Waals and H-bond forming (Figs. 1 and 2). The chloroindole moiety is located between residues Asn155 and Glu152, and the indole NH forms H-bond with Lys159. Therefore, the indole ring is placed in the strand transfer cavity.[21,22] The keto–enol oxygens and the tetrazole ring are positioned in the 3′-processing cavity. The keto–enol oxygens move far away from the Glu152 and point toward His67 and Glu92, and the

Table 1. Automated docking results for crystallographic ligands

HIV-1 enzyme	Ligand	Cluster	Members in cluster	Lowest energy (kcal/mol)	RMSD (Å)[d]
IN	5-CITEP[a]	1[b]	17	−8.25	7.046
		3[c]	62	−8.00	8.066
PR	L-745,524	1[b,c]	98	−18.4	0.556

[a] 5-CITEP was docked as anion.
[b] The cluster with lowest docking energy of a total 100 runs.
[c] The cluster with highest conformations of a total 100 runs.
[d] RMSD from crystal orientation.

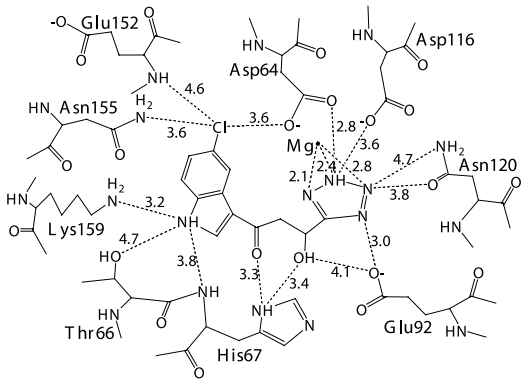

Figure 1. A schematic drawing of 5-CITEP showing interacting residues obtained from docking to HIV-1 IN protein.

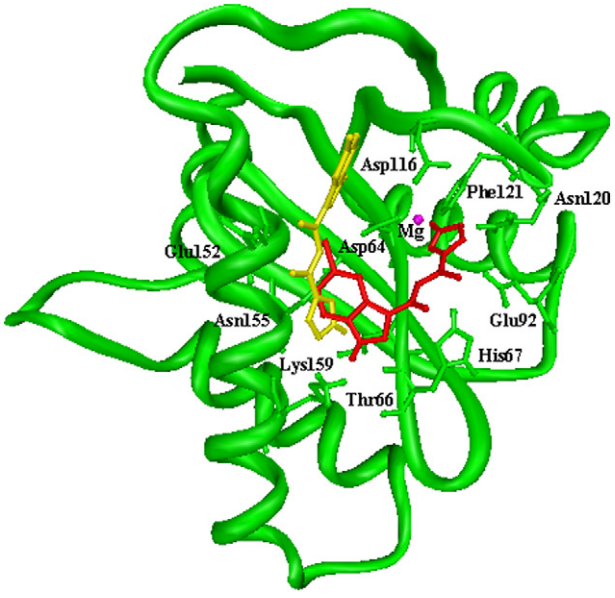

Figure 2. The comparison of bound conformation of 5-CITEP in the active site of HIV-1 IN: the yellow model shows the crystal orientation and the docking result is shown in red.

tetrazole ring locates in proximity to Asp116, Asp64, and Glu92, and forms co-ordinate bonds with Mg^{2+} ion.

The difference between docked orientation and crystal orientation was due to crystal packing effects and the flexible nature of HIV-1 IN. The crystal packing effect arises from the close linking between subunit A in the dimer of IN catalytic core domain that contains three asymmetric subunits (A, B, and C) in each monomer. It was reported that in the crystal environment, the ligand bound to subunit A has the influence on the binding mode of ligand bound to the linking subunit A′.[21] We looked at the crystal packing, and in fact saw that the ligand CITEP interacts with another CITEP ligand in an adjacent subunit across a 2-fold axis. It even looks like the two CITEP molecules may coordinate with an unreported Mg^{2+} sitting on the 2-fold. Thus, the crystal structure of the complex (1QS4) is clearly not representative of the ligand binding in a single, isolated active site. This is why the determined binding energy of the crystal configuration of 5-CITEP was found to be −4.38 kcal/mol, worse than the docked configuration (−8.00 kcal/mol). This poor binding energy supports the existence of the additional interaction due to differences in the crystal environment with respect to the solution phase as mentioned in the earlier report. As the additional interaction was not accounted in our docking procedure which used monomer alone, it led to high RMSD.

Many attempts have been made to decrease the RMSD as well as to prepare the precise IN template from 5-CITEP and other IN inhibitors, namely, using a sophisticated induced-fit protocol model[23] allowing fully flexible docking of both ligand and IN protein (the substrate-free form of the subunit C, 1BL3), using B chain without Mg^{2+} and a different docking tool (Genetic Optimization for Ligand Docking, GOLD).[21] All obtained 5-CITEP orientations from docking were different from the crystal structure. Although the docked orientation obtained from induced-fit model gave the best RMSD (0.272 Å), the orientation was still different from the crystal orientation; moreover, the methodology is sophisticated, not practical. In addition to the RMSD, experimental activity data are essential in the verification of the template derived from the substrate-free form. According to our simple docking, the result found was similar to the docked orientations previously reported by Sotriffer et al.[20] and Dayam et al.[21] In addition, Sotriffer et al.[20] reported that the estimated K_i (at 298 K) for their docking position was 8.7 μM, which compared well with the experimental IC_{50} value of 2.1–2.3 kcal/mol.[14] Chains B and C were not selected in our study because both chains are empty of HIV-1 IN and there is no Mg^{2+} in chain B. As 5-CITEP/HIV-1 IN complex (1QS4) is the only X-ray structure and the active site is not completely resolved, the suitability of the IN template verified by 5-CITEP crystal orientation alone would be crucial. Consideration of docking results could provide better indications of which regions should be explored to achieve a tighter binding.

Curcumin was docked to HIV-1 IN and HIV-1 PR. Docking of curcumin to HIV-1 proteins was carried out using the X-ray structure of curcumin obtained from Cambridge database and optimized with Insight II program (Accelrys Inc.) running on Silicon Graphics Indigo Elan workstation. The deprotonated form at the keto–enol position was assigned and atomic charges were added using Gasteiger–Marsili formalism, which was the charge method used in the calibration the AutoDock empirical free energy function. The ligand was set up for docking with the help of Autotors and the number of flexible torsions to be considered during the docking process was defined to 4, the hydroxyl and phenyl rotors. Docking results are shown in Table 2.

The binding modes are shown in Figure 3, which depicts the amino acid residues contacting curcumin. A schematic drawing of curcumin within the active site is shown in Figure 4. The keto–enol oxygens are located adjacent to the catalytic residues Asp116, Asp64, and

Table 2. Docking of curcumin into HIV-1 IN and PR proteins

HIV-1 enzyme	Cluster	Members in cluster	Lowest energy (kcal/mol)
IN	1[a,b]	48	−8.79
PR	1[a,b]	39	−9.77

[a] The cluster with lowest docking energy of a total 100 runs.
[b] The cluster with highest conformations of a total 100 runs.

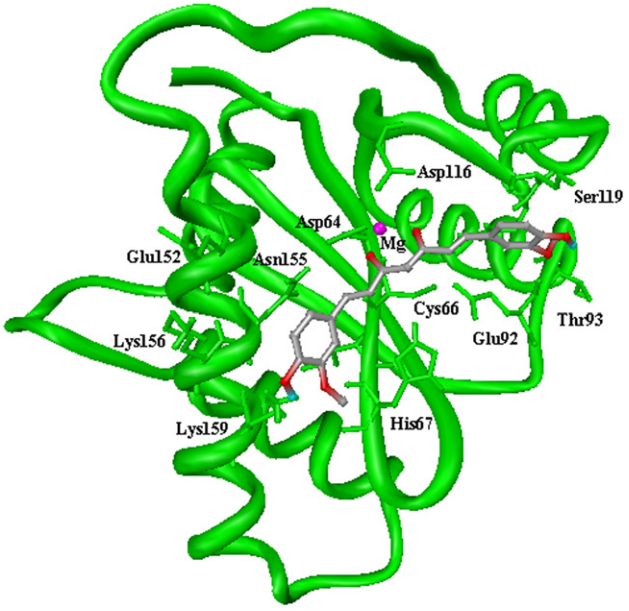

Figure 3. The bound conformation of curcumin in the active site of HIV-1 IN. The ribbon model shows the backbone of HIV-1 IN catalytic domain with all interacting amino acid residues shown as stick models. Mg^{+2} is in magenta.

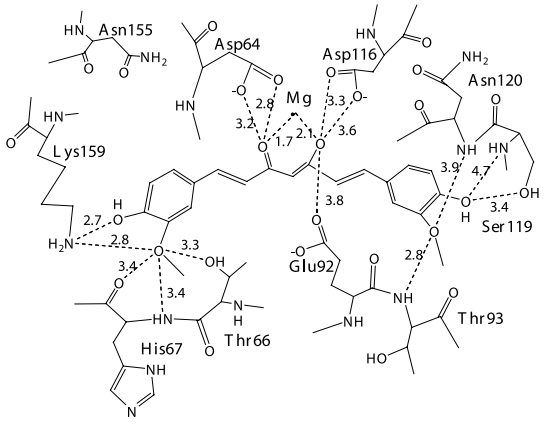

Figure 4. A schematic drawing of curcumin showing interacting residues obtained from docking to HIV-1 IN protein.

Glu92, and within coordination range of Mg^{2+}. One of the terminal phenol moiety forms H-bonds with His67, Thr66, and Lys155 inside a cavity with a large surface area. In addition, the other phenol ring is within hydrogen bonding range of Asn120, Ser119, and Thr93.

It is apparent that the orientation of docked curcumin involves both cavities of the IN active site, the areas for strand transfer and 3′-processing. Curcuma's phenolic oxygen interacts with Lys159 the ending residues for strand transfer function, and *o*-methoxy group interacts with His67 and Thr66, the amino acid residues in the bottom cavity involving in 3′-processing activity. Curcumin interacts with Asp116, Asp64, Glu92, Asn120, and the divalent Mg^{2+}, the amino acid residues in the cavity for 3′-processing contributing to its 3′-processing activity. The binding mode from this docking study is in accordance with the experimental activities of curcumin. Curcumin was reported to inhibit HIV-1 IN by both strand transfer and 3′-processing mechanisms. IC_{50} values against HIV-1 IN were found to be 40 and 30 μM for both strand transfer and 3′-processing activity, respectively.

The docking results of HIV-1 PR indicate that the curcumin structure fits well to the binding site of HIV-1 PR protein as demonstrated in Figure 5. The keto–enol oxygens of curcumin are in proximity to the carboxylic acid function of Asp25 (chain A) and the carbonyl oxygen (backbone) of Gly27 of chain B (Fig. 6). Importantly, the interaction of the oxygens of the phenolic and methoxy functions in one of the terminal phenyl rings forms H-bonds to residues Asp 29 and Asp 30 of chain A. The other ring is directly pointed to Asp29′ and Asp30′, however, only the phenolic oxygen is able

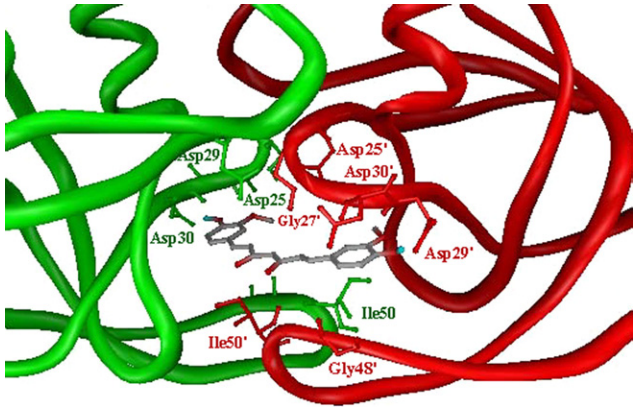

Figure 5. The bound conformation of curcumin in the active site of HIV-1 PR. The ribbon model shows chain A (green) and chain B (red) of HIV-1 PR catalytic domain with all interacting amino acid residues shown as stick models.

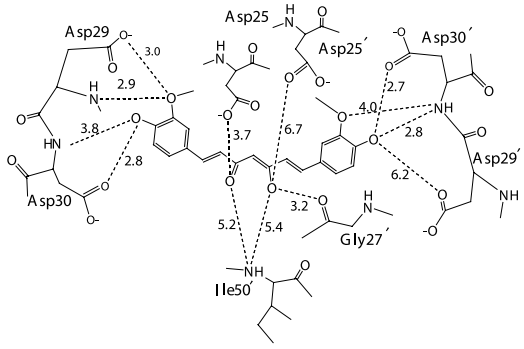

Figure 6. A schematic drawing of curcumin showing interacting residues obtained from docking to HIV-1 PR protein.

to form hydrogen bonding with Asp30′ at the proper distance of 2.7 Å, while the methoxyl oxygen is not close enough to interact with any residues.

Experimentally, curcumin shows a stronger inhibitory effect against IN than PR. The curcumin IC_{50} value against HIV-1 PR was reported as 100 μM. Our docking results support the superior inhibition of curcumin against HIV-1 IN, as curcumin is well targeted to the active binding sites of the IN protein. It is located and bound to multiple key functional residues. The *o*-hydroxyl and/or keto–enol structures are very important for both IN and PR inhibitory actions. In terms of PR binding modes, the symmetrical structure of curcumin appears to play an important role for binding to the homodimeric PR protein. In contrast, only keto–enol and one side of the terminal *o*-hydroxylated phenyl ring show tight binding to the IN catalytic domain.

3. Conclusion

Although the current knowledge about structure and function of the IN enzyme in vivo is still rather limited, the results mentioned above indicate that targeting the active site in the catalytic core domain should be a useful strategy for inhibitor design. Based on the crystal structures of inhibitors bound to the active sites of IN and PR of HIV-1, the docking results presented here suggest molecular mechanisms of curcumin inhibition which are in agreement with experimental activities. The information from these docking studies may be useful for further structure based drug design. Our docking results are the first detailed molecular binding models for curcumin interacting with HIV targets. The coherent picture that they present of its interactions at the active sites of PR and IN should be helpful in future inhibitor design.

References and notes

1. Weissbrich, B.; Heinkelein, M.; Jassoy, C. *Adv. Virus Res.* **2002**, *58*, 157.
2. Craigie, R. *J. Biol. Chem.* **2001**, *276*, 23213.
3. Jayaprakasha, G. K.; Rao, L. J. M.; Sakariah, K. K. *J. Agric. Food Chem.* **2002**, *50*, 3668.
4. Sreejayan, S.; Rao, M. N. A. *J. Pharm. Pharmacol.* **1994**, *46*, 1013.
5. Montterini, R. *Free Radic. Biol. Med.* **2000**, *28*, 1303.
6. Araujo, C. A. C.; Leon, L. L. *Mem. Inst. Oswaldo Cruz. Rio de Janeiro* **2001**, *96*, 723.
7. Pipier, J. T.; Singhal, S. S.; Salameh, M. S. *Int. J. Biochem. Cell. Biol.* **1998**, *30*, 445.
8. Elanttar, T. M. A.; Virji, A. S. *Anticancer Res.* **2000**, *20*, 1733.
9. Pardee, A. B.; Li, J.; Crumpacker, C.; Zang, L. US patent US933470, WO9404139, 1994, pp 1–31.
10. Mazumder, A.; Neamati, N.; Sunjay, S. *J. Med. Chem.* **1997**, *40*, 3057.
11. Mazumder, A.; Raghavan, K.; Weinstein, J.; Kohn, K. W.; Pommier, Y. *Biochem. Pharmacol.* **1995**, *49*, 1165.
12. Artico, M.; Santo, R. D.; Costi, R.; Novllino, E.; Greco, G.; Massa, S.; Tramontano, E.; Marongiu, M. E.; Dontis, A. D.; Colla, P. L. *J. Med. Chem.* **1998**, *41*, 3948.
13. Sui, Z.; Salto, R.; Li, J.; Craik, C.; Montellano, P. R. *Bioorg. Med. Chem.* **1993**, *1*, 415.
14. Goldgur, Y.; Craigie, R.; Cohen, G. H.; Fujiwara, T.; Yoshinaga, T.; Fujishita, T.; Sugimoto, H.; Endo, T.; Murai, H.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13040.
15. Goldgur, Y.; Dyda, F.; Hickman, A. B.; Jenkins, T. M.; Craigie, R.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 9150.
16. Chen, Z.; Li, Y.; Chen, E.; Hall, D. L.; Darke, P. L.; Culberson, C.; Shafer, J. A.; Kuo, L. C. *J. Biol. Chem.* **1994**, *269*, 26344.
17. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639.
18. Hetenyi, C.; van der Spoel, D. *Protein Sci.* **2002**, *11*, 1729.
19. Sotriffer, C. A.; Ni, H.; McCammon, J. A. *J. Med. Chem.* **2000**, *43*, 4109.
20. Sotriffer, C. A.; Ni, H.; McCammon, J. A. *J. Am. Chem. Soc.* **2000**, *122*, 6136.
21. Dayam, D.; Neamati, N. *Bioorg. Med. Chem.* **2004**, *12*, 6371.
22. Heuer, T. S.; Brown, P. O. *Biochemistry* **1997**, *36*, 10655.
23. Keseru, G. M.; Kolossva'ry, I. *J. Am. Chem. Soc.* **2001**, *123*, 12708.