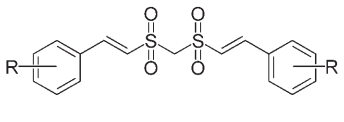
DOI: 10.1002/cmdc.200600040

Correlation of Biological Activity with Active Site Binding Modes of Geminal Disulfone HIV-1 Integrase Inhibitors

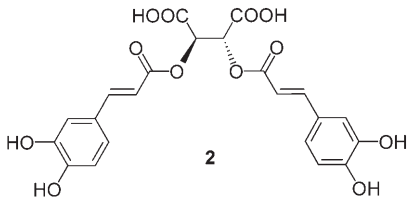
D. Christopher Meadows, Dean J. Tantillo, and Jacquelyn Gervay-Hague*[a]

HIV infection affects more than 40 million people worldwide, and infection rates continue to rise in all areas of the world except the Caribbean.[1] Current treatment generally involves a combination of inhibitors of viral reverse transcriptase (RT) and viral protease (PR). While combination therapy has been highly effective at lowering viral loads and extending lives, problems with toxicity and especially drug resistance remain major obstacles to HIV chemotherapy. Indeed, almost 25% of new infections are found to be drug resistant, clearly demonstrating the need for new drugs to combat HIV infection.[2] The viral replication cycle offers many potential sites for interruption using chemotherapy; one of the most promising is integrase (IN). IN is a 32 kDa protein that processes proviral DNA in a step termed 3′-processing (3P) which is followed by strand transfer (ST), whereby IN inserts the processed proviral DNA into the host genome. Several compounds have been reported to inhibit IN, yet the lack of structural information for the intact protein and questions regarding its oligomeric nature have impeded the discovery of a clinically useful IN inhibitor.[3] These difficulties have made molecular modeling a critical component in both the design of new IN inhibitors and the identification of important protein–ligand interactions.[4]

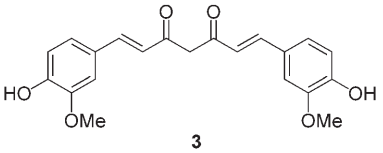
We recently reported a series of potent disulfone-containing anti-IN and antiviral compounds of general structure **1**.[5] The compounds were originally designed to be neutral chicoric acid analogues (L-chicoric acid shown, **2**), reasoning that they could more easily traverse the cell membrane to interact with cytoplasmic IN. Having demonstrated that these compounds were active against IN in the low-micromolar range, it became important to identify the bioactive conformations in order to design more potent inhibitors. To this end, docking studies on the catalytic core of IN were performed. While there are 14 crystal structures of IN available from the Protein Data Bank (PDB), only one of these has an inhibitor bound in the active site: 1QS4.[6] Of the remaining 13, only two have completely resolved active sites. Because we endeavored to design more potent inhibitors of IN, we felt the crystal structure containing the inhibitor, 1QS4, would be the most relevant active site conformation on which to conduct the docking experiments, despite previously reported crystal-packing effects associated with this structure.[7]



1



2



3

Further impetus for the docking studies stem from previous modeling studies independently conducted by the research groups of McCammon and Olson. McCammon and co-workers docked chicoric acid, among other IN inhibitors, into the active site of 1QS4 to compare binding modes of various compounds.[8] Later, Olson and co-workers docked structurally related curcumin **3** into the 1QS4 active site to explore its binding modes.[9] Curiously, these molecules were found to have rather different active site conformations (see below). We were interested to determine which, if any, similarities existed between our disulfone-containing compounds and either of the other structurally analogous inhibitors **2** and **3**.

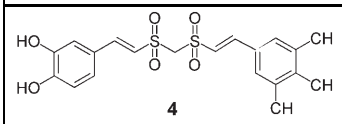
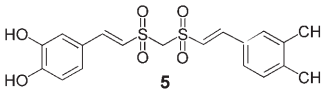
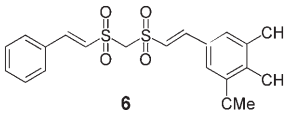
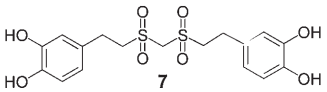
The docking studies were performed using AutoDock 3.0.[10,11] This program uses a genetic algorithm-based approach to sample various conformations of small molecules inside binding cavities of protein hosts. Docking parameters began with 50 individuals, a maximum of 2500000 energy evaluations, and a maximum of 27000 generations. The crossover rate was set to 0.80, and the mutation rate was 0.02 with Cauchy distribution parameters $\alpha = 0$ and $\beta = 1$. Elitism was set to 1. In the local search, the pseudo-Solis–Wets algorithm was applied with a maximum of 300 iterations per search. The probability of performing the local search on an individual was 0.06, and the maximum number of successes or failures before changing the local search step size was 4. The number of initial torsions varied depending on the ligand docked. A total of 100 independent docking runs were then performed for each compound, and the results differing by less than 1.5 Å rmsd were clustered into groups. The best docked conformations were those found to have the lowest binding energy and also the greatest number of members in the cluster, which is an indication of good convergence.

The initial structures of the small molecules (structures and IC_{50} values shown in Table 1) were determined using molecular

[a] *Dr. D. C. Meadows,[+] Prof. D. J. Tantillo, Prof. J. Gervay-Hague*
Department of Chemistry
University of California, Davis
One Shields Ave., Davis, CA 95616 (USA)
Fax: (+1) 530-752-8995
E-mail: jgervayhague@ucdavis.edu

[+] *Recipient of the R. B. Miller Dissertation Award, 2005, UC Davis*

Table 1. Structures and IC_{50} values for test compounds.

Compound	IC_{50} (3P) [μM]	IC_{50} (ST) [μM]		No. in Cluster[a]	ΔG [kcal mol^{-1}][b]	Hydrogen Bonding Residues[c]
 **4**	1.8	0.9	U	45	−8.8	K159, T66, E92, D64, N155
			L	31	−7.8	K159, T66, E152, Q146, Q148
 **5**	4	5	U	49	−8.1	K159, D116, N155, Q148
			L	40	−7.8	K159, T66, D64, E152, Q146
 **6**	80	80	U	34	−7.6	K159, T66, D64, D116, Q146, Q148
			L	38	−7.6	K159, T66, Q148
 **7**	40	17	U	42	−8.2	K159, D116, C65, Q148
			L	36	−7.9	K159, T66, Q146, Q148

[a] Number of individuals in top two clusters. [b] Free energy of binding for top two clusters. [c] Residues with hydrogen bond donor/acceptor groups within 4 Å of the bound ligand are shown.

mechanics methods to conduct the energy minimizations with the AMBER force field in MacroModel.[12,13] All atomic charges were assigned using the Gasteiger–Marsili formalism, nonpolar hydrogen atoms and lone pairs merged, aromatic carbons identified, rigid root defined, and all formal single bonds except those to H were allowed to rotate.

The catalytic domain of crystal structure 1QS4 was used with domains B and C, crystallographic waters, and the bound ligand removed. The active site Mg ion was maintained with a charge of +2 throughout. Polar hydrogen atoms were added using PROTONATE and distributed with AutoDock. Histidine residues were maintained unprotonated as previously determined through computational methods to be appropriate.[14,15] Lastly, Kollman united-atom partial charges and solvent parameters were added. The grid maps representing the protein were then calculated using AutoGrid, and the dimensions were set to 23×23×23 Å^3 with a spacing of 0.375 Å centered on catalytic residue Glu 152.

The docking studies revealed two low-energy binding modes common to all of the test compounds: a "U-shaped" conformation and an "L-shaped" conformation (Figures 1–4). The U-shaped conformer makes use of almost the entire enzyme active site, while the L-shaped conformer fills half of the active site with half of the molecule, the other half pointing away from the catalytic residues.

As can be seen from Figure 1, the U conformer of **4** makes extensive contacts between one of the sulfones and K159 and T66, with the other sulfone situated between K159 and K156. One of the catechols coordinates the active site Mg ion and also shows interactions with catalytic residue D64, while the other ring has contacts with E92 and N155.[16] Alternatively, in the L conformer it is one of the catechols showing interactions between K159 and T66, while one sulfone contacts Q148 and the other, E152; the remaining catechol interacts with Q146.

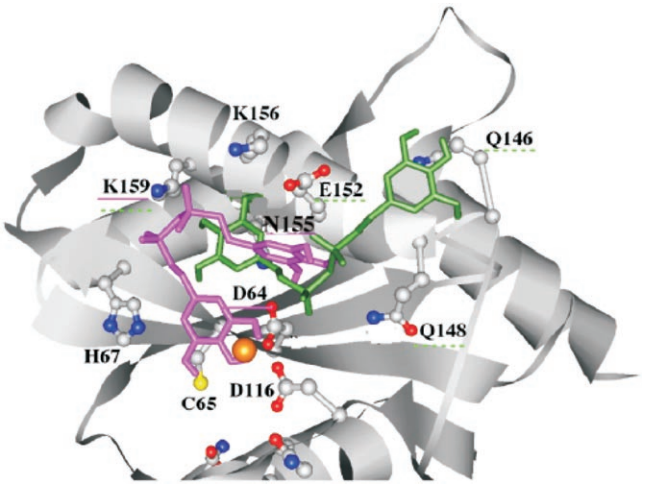


Figure 1. Two low-energy binding modes for **4**. Magenta underscore indicates hydrogen bonding to U-shaped conformer. Green dashed underscore indicates hydrogen bonding to L-shaped conformer.

The predicted binding energy for the U conformer (−8.8 kcal mol^{-1}) is slightly more favorable than that for the L conformer (−7.8 kcal mol^{-1}). This is intuitive given the more extensive active site interactions made by the U conformer.

The U conformer of **5** (Figure 2), not surprisingly, makes active site contacts similar to **4**. One of the sulfones makes contacts with K159, and the Mg ion is again coordinated by one of the catechol rings which also interacts with D116. The other sulfone interacts with N155, and the second catechol contacts Q148. The L conformer of **5** also makes contacts similar to **4** with one of the catechols interacting with K159 and T66, while the other catechol is involved in H bonding with Q146. The sulfones are favorably situated with one interacting

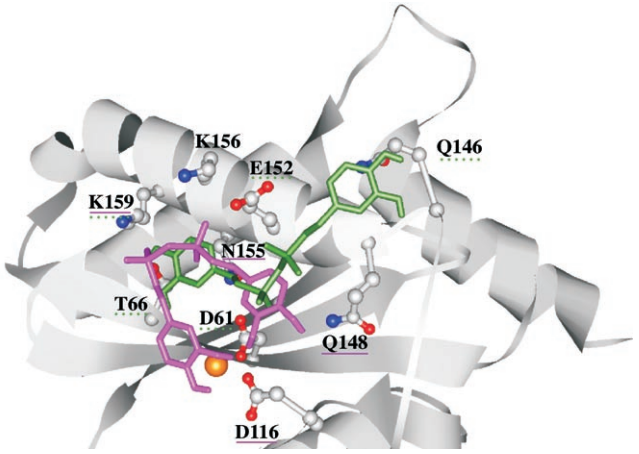


Figure 2. Two low-energy binding modes for **5**. Magenta underscore indicates hydrogen bonding to U-shaped conformer. Green dashed underscore indicates hydrogen bonding to L-shaped conformer.

with catalytic residue D64 and the other with catalytic residue E152. Here, as with **4**, the predicted binding energy of the U conformer (-8.1 kcal mol^{-1}) is again more favorable than that predicted for the L conformer (-7.8 kcal mol^{-1}).

The U conformer of **6** (Figure 3) also shows common albeit fewer interactions than the previous inhibitors. The catechol ring coordinates the Mg ion and also interacts with catalytic residues D64 and D116. One of the sulfones makes contacts with K159 and T66. It is interesting in the case of the L conformer that it is the functionalized ring that is inserted into the active site while the non-functionalized ring is pointing away. Here the catechol interacts with K159 and T66, and one of the sulfones makes contacts with Q148. It is intuitive that **6**, which lacks hydrogen bonding capability on half of the molecule, would result in the observed reduced activity. In this case the U and L conformers have identical binding energies: $\Delta G_U = \Delta G_L = -7.6$ kcal mol^{-1}.

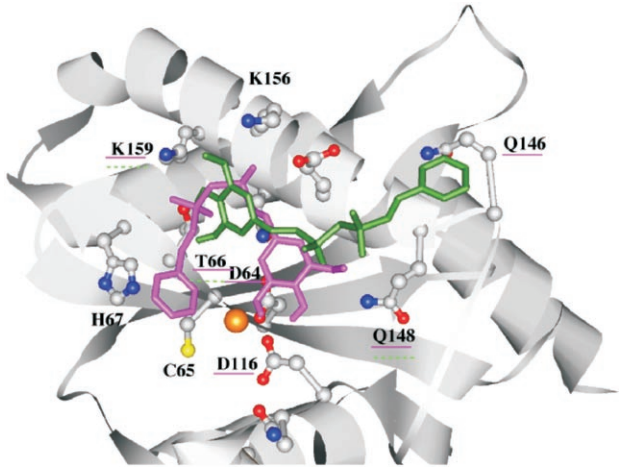


Figure 3. Two low-energy binding modes for **6**. Magenta underscore indicates hydrogen bonding to U-shaped conformer. Green dashed underscore indicates hydrogen bonding to L-shaped conformer.

Compound **7** was made to help identify whether the vinyl sulfone moieties in the molecules were responsible for their activity. While the anti-IN activity of this molecule is lower than that of its unsaturated analogue, it still shows significant activity against both catalytic functions performed by the enzyme. Figure 4 shows that the saturated derivative **7** still possesses

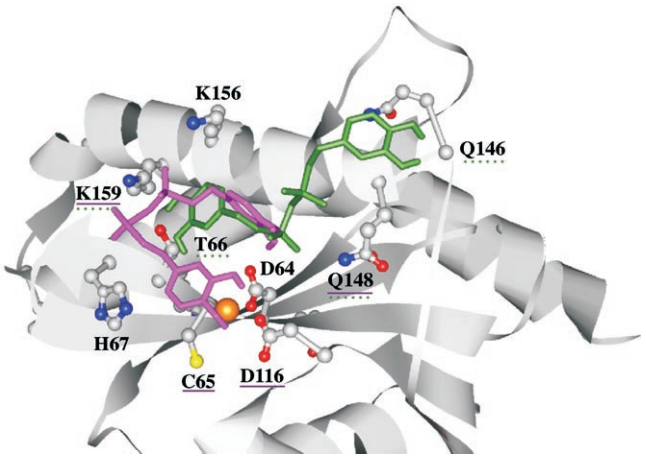


Figure 4. Two low-energy binding modes for **7**. Magenta underscore indicates hydrogen bonding to U-shaped conformer. Green dashed underscore indicates hydrogen bonding to L-shaped conformer.

two low-energy binding modes. The U conformer coordinates the Mg ion through one of the catechol rings, while the rest of the molecule interacts with many of the same residues as **5,** adding a hydrogen bond to C65. The L conformer shows one of the catechol rings interacting with K159 and T66, the other catechol within H-bonding distance to Q146, and one of the sulfones making contacts with Q148. As with the symmetric molecules **4** and **5**, the U conformer of **7** possesses a more favorable binding energy of -8.2 kcal mol^{-1} as compared with -7.9 kcal mol^{-1} for the L-shaped conformation.

It is interesting to compare the binding modes of the disulfone-containing compounds with previously reported binding modes of analogous inhibitors. In the case of curcumin, the "west" half of the molecule occupies a similar position to that taken by the L conformer, with the phenolic group forming hydrogen bonds to K159 and T66.[9] The ketone/enol functionalities of curcumin interact in the center of the active site with all three catalytic residues and the Mg ion, whereas the L conformer of **4** and **5** shows interaction of one of the sulfones and the catalytic residues. The other half of curcumin extends almost linearly to interact with T93, S119, and N120, interactions not observed for either low-energy binding modes identified in our docking. A plausible reason for the difference in active site interactions is that the ketone/enol form of curcumin adopts a more planar conformation than the disulfone compounds.

The L conformer also shows some of the same interactions observed for another small-molecule inhibitor of IN docked by McCammon.[8] The anthraquinone, quinalizarin **8**, forms hydro-

gen bonds to K159 and T66 through its catechol ring, just as the disulfone compounds do. One of the phenols on the other aromatic ring of the anthraquinone interacts with the center of the active site, a similar placement to the sulfone (Figure 5).

Conversely, the U conformer of the disulfone compounds is very similar to the most favorable binding mode obtained from the docking of L-chicoric acid (Figure 6).[8] Here the carboxylates of chicoric acid occupy a similar position to the sulfones, with one between K156 and K159 and the other between K159 and H67. The catechol of the "east" half of chicoric acid also has a very similar position to that of the analogous catechol of the disulfone compounds, forming hydrogen bonds with Q148. Due to the more extended nature of chicoric acid, however, the catechol of the "west" half of the molecule is pushed further away from the catalytic residues and interacts with E92 rather than the Mg ion and D116.

As noted previously, 1QS4 was chosen because it has a ligand bound in the active site. However, we felt it might be instructive to dock the small molecule into crystal structure 1BL3, which contains a fully resolved active site, to compare what, if any, common binding modes are present. In pilot studies, **5** was docked into the active site of 1BL3 using the same docking protocol.[17] Interestingly, this also results in two low-energy binding modes: the familiar U-shaped conformation and a new binding mode not observed with 1QS4 (Figure 7). The new binding mode shows the one catechol coordinating

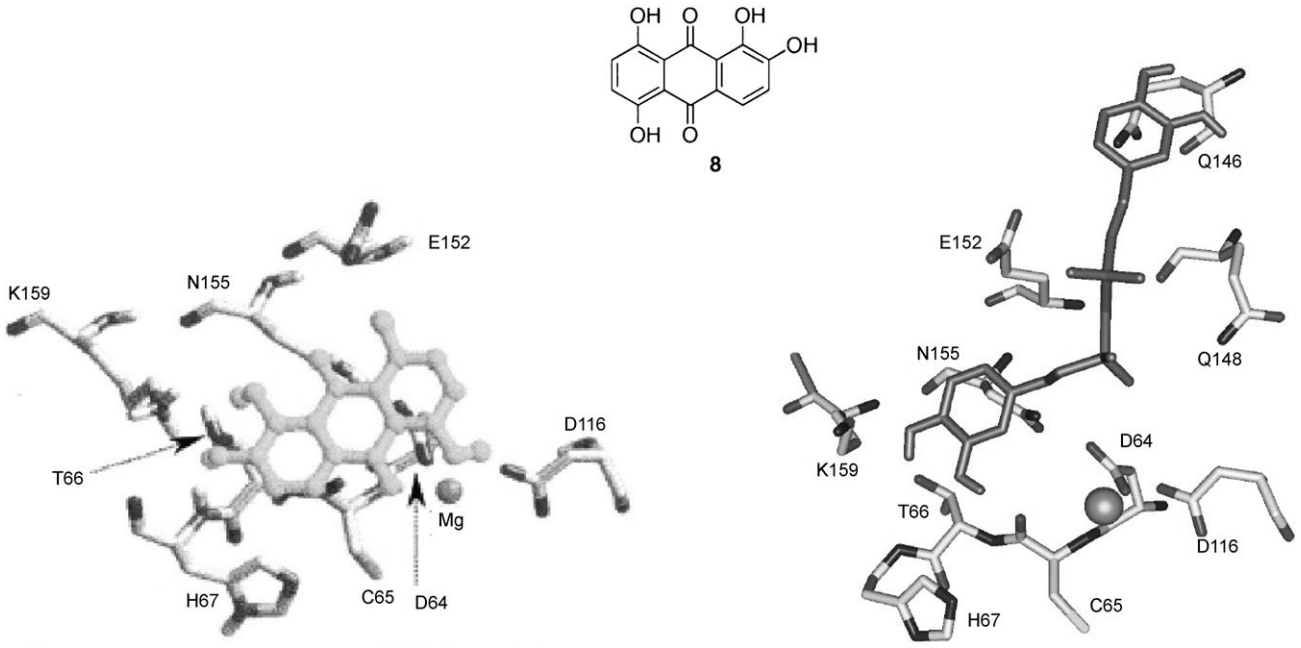


Figure 5. Comparison of binding modes between quinalizarin (left) and **5** (right). (Reproduced with permission, Copyright 2000 American Chemical Society.)[8]

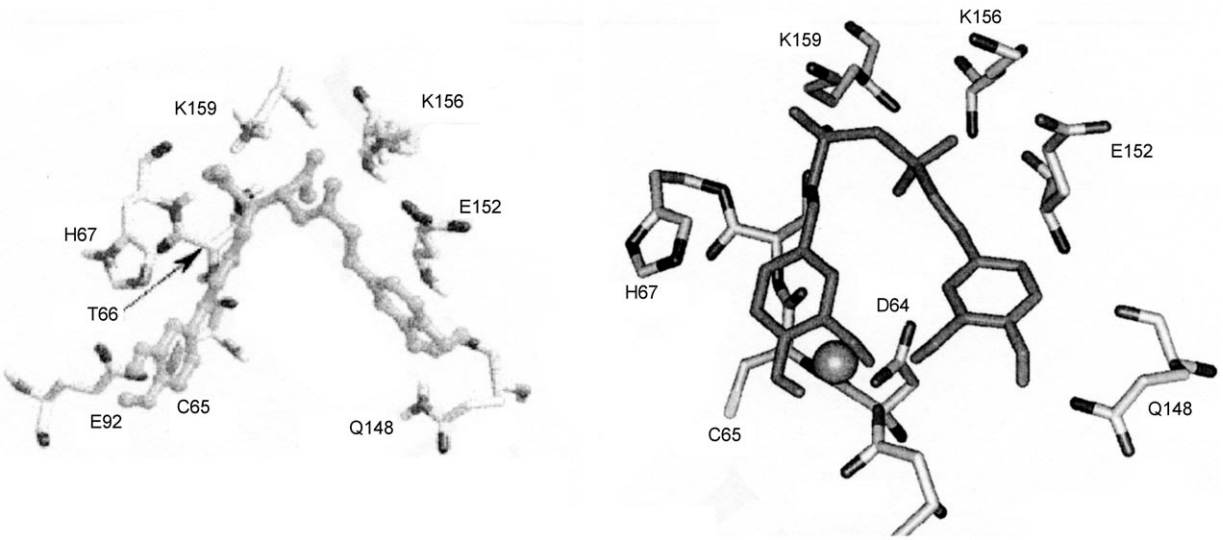


Figure 6. Comparison of binding modes between chicoric acid (left) and **5** (right). (Reproduced with permission, Copyright 2000 American Chemical Society.)[8]

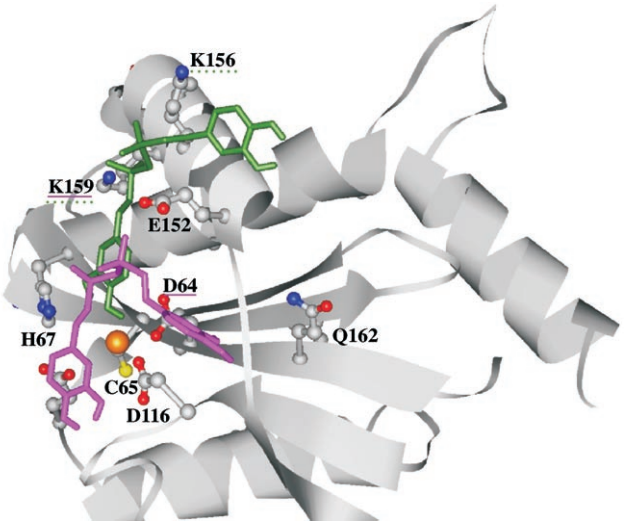


Figure 7. Docking **5** into the active site of 1BL3. Magenta underscore indicates hydrogen bonding to U-shaped conformer. Green dashed underscore indicates hydrogen bonding to L-shaped conformer.

the Mg ion and the sulfones interacting with K156 and K159. With this, as in previous examples, the U-shaped conformation displays a better binding energy at -6.6 kcal mol^{-1} compared with -6.1 kcal mol^{-1} for the new binding mode. The difference in binding modes is not altogether unexpected given the large differences in active site conformations between the crystal structures. Diverse binding modes were also found previously for the bis-diketo acids depending on the crystal structure used.[18]

The docking studies reported herein suggest that tight interactions between K156, K159, Q148, and the catalytic residues are important for potent inhibition of IN. Previous studies have substantiated this binding through both molecular dynamics simulations and photo-cross-linking studies.[19,20] Furthermore, Lee and Robinson recently showed through site directed mutagenesis that in addition to the above-listed residues, mutations of E92, G140, or G149 (the latter two probably reduce flexibility in the disordered loop) also led to reduced susceptibility to chicoric acid, implicating these residues in the binding of the inhibitor.[21]

The enzymatic data is more consistent with the energies associated with the U-shaped binding mode than the L, suggesting that this conformation is more biologically relevant. Perhaps molecules that are structurally biased to a more bent conformation would enable better exploitation of these important interactions and lead to more potent inhibitors of IN.

Docking can be useful in the design of more potent inhibitors; however, IN represents an especially challenging subject and its complexity cannot be overstated. Ambiguities associated with the flexible loop, different active site conformations depending on crystal structure, and the questionable oligomeric nature of IN complicate any computational studies performed on the enzyme. Nevertheless, the docking results reported herein show good correlation with experimental data and provide a valuable tool for both evaluating compounds and designing more potent inhibitors.

Acknowledgements

This work was supported by the R. B. Miller Fellowship, National Institutes of Health (Grant GM60917), and the University of California AIDS Research Program (Grant D02D400). We appreciate the thoughtful and helpful comments provided by the reviewers of this manuscript.

Keywords: AutoDock · binding modes · HIV · inhibitors · integrase

[1] UNAIDS, AIDS Epidemic Update, December **2005**.
[2] a) L. Menendez-Arias, *Trends Pharmacol. Sci.* **2002**, *23*, 381; b) E. Poveda, B. Rodes, J.-L. Labernardiere, J. M. Benito, C. Toro, J. Gonzalez-Lahoz, J.-L. Faudon, F. Clavel, J. Schapiro, V. Soriano, *J. Med. Virol.* **2004**, *74*, 21; c) S. A. Wegner, S. K. Brodine, J. R. Mascola, S. A. Tasker, R. A. Shaffer, M. J. Starkey, A. Barile, G. J. Martin, N. Aronson, W. W. Emmons, K. Stephan, S. Bloor, J. Vingerhoets, K. Hertogs, B. Larder, *AIDS* **2000**, *14*, 1009.
[3] D. C. Meadows, J. Gervay-Hague, *ChemMedChem* **2006**, *1*, 16.
[4] a) M. C. Nicklaus, N. Neamati, H. Hong, A. Mazumder, S. Sunder, J. Chen, G. W. A. Milne, Y. Pommier, *J. Med. Chem.* **1997**, *40*, 920; b) H. Hong, N. Neamati, S. Wang, M. C. Nicklaus, A. Mazumder, H. Zhao, T. R. Burke, Y. Pommier, G. W. A. Milne, *J. Med. Chem.* **1997**, *40*, 930; c) A. L. Perryman, J. A. McCammon, *J. Med. Chem.* **2002**, *45*, 5624; d) S. Guenther, V. Nair, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2233; e) R. Dayam, N. Neamati, *Bioorg. Med. Chem.* **2004**, *12*, 6371; f) J. R. Schames, R. H. Henchman, J. S. Siegel, C. A. Sotriffer, H. Ni, J. A. McCammon, *J. Med. Chem.* **2004**, *47*, 1879; g) C. A. Sotriffer, H. Ni, J. A. McCammon, *J. Am. Chem. Soc.* **2000**, *122*, 6136; h) H. M. Zhu, W. Z. Chen, C. X. Wang, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 475; i) C. Marchand, A. A. Johnson, R. G. Karki, G. C. G. Pais, X. Zhang, K. Cowansage, T. A. Patel, M. C. Nicklaus, T. R. Burke, Jr., Y. Pommier, *Mol. Pharmacol.* **2003**, *64*, 600; j) F. Aiello, A. Brizzi, A. Garofalo, F. Grande, G. Ragno, R. Dayam, N. Neamati, *Bioorg. Med. Chem.* **2004**, *12*, 4459.
[5] D. C. Meadows, T. B. Matthews, T. W. North, M. J. Hadd, C. L. Kuo, N. Neamati, J. Gervay-Hague, *J. Med. Chem.* **2005**, *48*, 4526.
[6] Y. Goldgur, R. Craigie, G. H. Cohen, T. Fujiwara, T. Yoshinaga, T. Fujishita, H. Sugimoto, T. Endo, H. Murai, D. R. Davies, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13040.
[7] C. A. Sotriffer, H. Ni, A. J. McCammon, *J. Am. Chem. Soc.* **2000**, *122*, 6136.
[8] Chicoric acid was docked according to the described experimental protocol and the most favorable binding mode exhibited the same conformation described in C. A. Sotriffer, H. Ni, J. A. McCammon, *J. Med. Chem.* **2000**, *43*, 4109.
[9] Curcumin was docked according to the described experimental protocol and the same binding mode was found albeit as the second most favorable conformation described in O. Vajragupta, P. Boonchoong, G. M. Morris, A. J. Olson, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3364.
[10] AutoDock 3.0:http://www.scripps.edu/pub/olson-web/doc/autodock/
[11] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19*, 1639–1662.
[12] F. Mohamadi, N. G. J. Richard, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, W. C. Still, *J. Comput. Chem.* **1990**, *11*, 440.
[13] D. A. Case, T. E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K. M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang, R. Woods, *J. Comput. Chem.* **2005**, *26*, 1668.
[14] R. D. Lins, J. M. Briggs, T. P. Straatsma, H. A. Carlson, J. Greenwald, S. Chloe, J. A. McCammon, *Biophys. J.* **1999**, *76*, 2999–3011.
[15] C. A. Sotriffer, H. Ni, J. A. McCammon, *J. Med. Chem.* **2000**, *43*, 4109.
[16] For examples of catechols complexing active site metal ions a) C. Steegborn, T. N. Litvins, K. C. Hess, A. B. Capper, R. Taussig, J. Buck, L. R. Levin, H. Wu, *J. Biol. Chem.* **2005**, *280*, 31754; b) C. Lerner, A. Ruf, V. Gramlich, B. Masjost, G. Zurcher, R. Jakob-Roetne, E. Borroni, F. Diederich, *Angew. Chem.* **2001**, *113*, 4164; *Angew. Chem. Int. Ed.* **2001**, *40*, 4040.
[17] S. Maignan, J. P. Guilloteau, Q. Zhou-Liu, C. Clement-Mella, V. Mikol, *J. Mol. Biol.* **1998**, *282*, 359.

[18] R. Dayam, N. Neamati, *Bioorg. Med. Chem.* **2004**, *12*, 6371.
[19] A. A. Adesokan, V. A. Roberts, K. W. Lee, R. D. Lins, J. M. Briggs, *J. Med. Chem.* **2004**, *47*, 821.
[20] T. M. Jenkins, D. Esposito, A. Engelman, R. Craigie, *EMBO J.* **1997**, *16*, 6849.
[21] D. J. Lee, W. E. Robinson, Jr., *Antimicrob. Agents Chemother.* **2006**, *50*, 134.

Received: February 17, 2006
Revised: March 27, 2006
Published online on July 26, 2006

