J Mol Model (2013) 19:1125–1142
DOI 10.1007/s00894-012-1660-4

ORIGINAL PAPER

Can cyclic HIV protease inhibitors bind in a non-preferred form? An ab initio, DFT and MM-PB(GB)SA study

Daniel P. Oehme · Robert T. C. Brownlee · David J. D. Wilson

Received: 21 May 2012 / Accepted: 21 October 2012 / Published online: 13 November 2012

Abstract X-ray crystallography studies have identified that most cyclic inhibitors of HIV protease (including cyclic ureas) bind in a symmetric manner, however some cyclic inhibitors, such as cyclic sulfamides, bind in a non-symmetric manner. This raises the question as to whether it is possible for cyclic sulfamides to bind symmetrically and conversely for cyclic ureas to bind non-symmetrically. Herein we report an analysis of the conformational preference of cyclic ureas and sulfamides both free in solution and bound to HIV protease, including an investigation of the effect of branching. Quantum chemical calculations (B3LYP, M06-2X, MP2, CCSD(T)) predict the cyclic urea to prefer a symmetric conformation in solution, with a large activation barrier towards inter-conversion to the non-symmetric conformation. This differs from the cyclic sulfamides, which marginally prefer a non-symmetric conformation with a much smaller barrier to inter-conversion making it more likely for a non-preferred conformation to be observed. It is predicted that the cyclic scaffold itself favours a symmetric form, while branching induces a preference for a non-symmetric form. MD simulations on the free inhibitors identified inter-conversion with the cyclic sulfamides but not the cyclic ureas, in support of the quantum chemical results. MM-PB(GB)SA calculations on the cyclic inhibitors bound to HIV protease corroborate the X-ray crystallography studies, identifying the cyclic ureas to bind symmetrically and the cyclic sulfamides in a non-symmetrical manner. While the non-preferred form of the sulfamide may well be present as a free molecule in solution, our results suggest that it is unlikely to bind to HIV protease in a symmetric manner.

Keywords Ab initio · Cyclic inhibitors · DFT · HIV protease · MM-PB(GB)SA · Molecular dynamics (MD)

Introduction

The search for potent and selective small molecule inhibitors of HIV protease has led to a focus on cyclic classes of compounds. Cyclic compounds offer advantages over extended 'straight-chain' inhibitors due to their smaller size and constrained flexibility [1]. As free molecules in solution, cyclic inhibitors possess reduced flexibility in comparison with extended (non-cyclic) inhibitors, which is largely responsible for limiting the entropy loss of the cyclic inhibitors upon binding to HIV protease. Further reduction of the entropy loss is expected by including inhibitor functionality to mimic the conserved 'structural' water present in the binding site of HIV protease [2, 3]. Moreover, cyclic compounds may be expected to yield improved bioavailability over extended chain compounds of similar potency.

To date a number of cyclic inhibitors of varying potency have been developed [4–6]. A cyclic urea compound was the first cyclic inhibitor co-crystallized in HIV protease [7], which subsequently lead to the development of numerous urea analogues [1, 8–12]. Most known cyclic inhibitors have been developed from a common template structure; a central seven-member ring that allows for seven-fold substituent functionality (Fig. 1). The functionality at the top of the cyclic scaffold (W) mimics the hydrogen bonding features of a key structural water molecule that interacts with the HIV protease flap tips (β-hairpin structures). The four side chains (R_1, R_1', R_2, R_2') branch off the scaffold and bind to

Electronic supplementary material The online version of this article (doi:10.1007/s00894-012-1660-4) contains supplementary material, which is available to authorized users.

D. P. Oehme · R. T. C. Brownlee · D. J. D. Wilson (✉)
Department of Chemistry, La Trobe Institute for Molecular Sciences (LIMS), La Trobe University,
Bundoora, Victoria 3086, Australia
e-mail: david.wilson@latrobe.edu.au

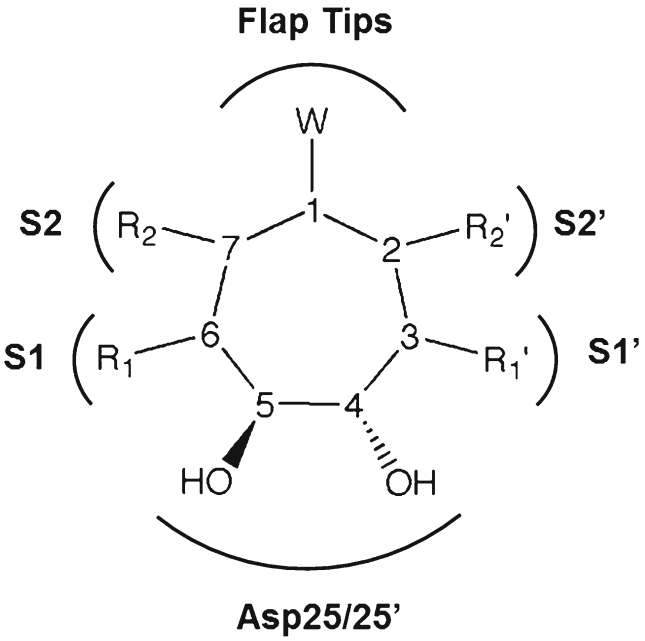

Fig. 1 Common seven-member ring template structure of cyclic inhibitors of HIV protease. The W signifies functionality to mimic the structural water molecule, $R_2/R_1/R_1'/R_2'$ are branches that bind to the S2/S1/S1′/S2′ subsites of HIV protease, and the diol functionality mimics the catalytic water molecule

the four major subsites of HIV protease (S1, S1′, S2, S2′ using standard nomenclature where S*i* and S*i*' sites are structurally equivalent), while the branches can also be extended to fill the S3 and S3′ subsites. The bottom of the scaffold generally contains diol functionality to interact with the catalytic aspartates and their flanking glycine residues. For the interested reader, the review of Brik and Wong [13] gives a full description of the binding site of HIV protease.

The binding site of HIV protease exhibits C_2 symmetry, and so the C_2-symmetric cyclic inhibitors may be expected to bind in a symmetric fashion with the axes of symmetry of the inhibitor and enzyme being in alignment (positional symmetry) [14]. Conformational symmetry of the inhibitor is also important, for which binding of an inhibitor in the active site of HIV protease is considered to be symmetric if the R_1/R_1' branches bind into the S1/S1′ subsites and the R_2/R_2' branches bind into the S2/S2′ subsites, respectively (Fig. 2). Indeed, the majority of cyclic inhibitors adopt a symmetric conformation when bound to HIV protease [15]. However, it should be noted that symmetric inhibitors are able to bind in both symmetric and non-symmetric modes [16].

The cyclic sulfamides provide an exception to this preference for symmetric binding. Bäckbro and co-workers [15] have demonstrated that cyclic sulfamides adopt a non-symmetric conformation in the binding site with the R_1' branch binding in the S2′ subsite while the R_2' branch binds in the S1′ subsite (Fig. 2). It has been hypothesized that the change in symmetry preference of the cyclic sulfamides is due to a change in the geometry of the central scaffold in comparison with other cyclic inhibitors [15].

The question arises as to why the cyclic sulfamides bind in a non-symmetric conformation when the vast majority of cyclic inhibitors bind in a symmetric manner. Hulten and co-workers [17] have previously investigated the conformer preference of a cyclic sulfamide inhibitor (with methyl branches) and calculated the non-symmetric form to be only 10.0 $kJmol^{-1}$ lower in energy at the B3LYP/6-31G(d) level of theory. It was suggested that this barrier is small enough that appropriate substitution of the branches could overcome the symmetry preference of the free molecule, leading to the possibility that the cyclic sulfamide inhibitor may bind in a symmetric conformation. Schaal and co-workers [18] have tested this hypothesis by making and testing a series of potential sulfamide inhibitors with both symmetric and non-symmetric substituted branches, however they were unsuccessful in producing sulfamide inhibitors that bind symmetrically. To date there is no reported evidence of cyclic sulfamide inhibitors binding in a symmetric manner.

Although there are no available crystal structures with the inhibitors bound in a non-preferred manner, the possibility that these inhibitors can bind in a non-preferred manner cannot be excluded. While it is common to assume that the preferred mode of binding is provided by the available crystal structures, it should be remembered that crystal structures themselves are often effectively force-field fitted snapshots of the measured electron density [19]. Indeed, it is possible to consider crystal structures as the protein-ligand structure that is the kinetically least soluble form of the protein, and moreover, that the X-ray structure of the protein-ligand complex may consist of an ensemble of conformations [19]. The X-ray structure is a single representation of a system over the conformational energy landscape and so there remains the distinct possibility that more than one conformation of an inhibitor could be bound. Moreover, the preferred conformation of the bound inhibitor may not be the same as when the inhibitor is free in solution.

To answer the title question it is critical to accurately model the conformer preference of these cyclic inhibitors as free molecules in solution, as well as when they are complexed with HIV protease. In their study of sulfamide inhibitors, Hulten and co-workers [17] only reported B3LYP/6-31G(d) calculated energies, which may not be sufficiently accurate for the reliable treatment of such molecular systems. Here we report a detailed theoretical investigation of the relative stability of the symmetric and non-symmetric conformations of cyclic ureas and sulfamides free in solvent and bound in HIV protease. Ab initio and density functional theory (DFT) calculations have been carried out to determine the favoured (lowest energy) conformations of the inhibitors when free in solution, in both a non-polar (diethyl ether) and polar (water) solvent with an implicit solvent

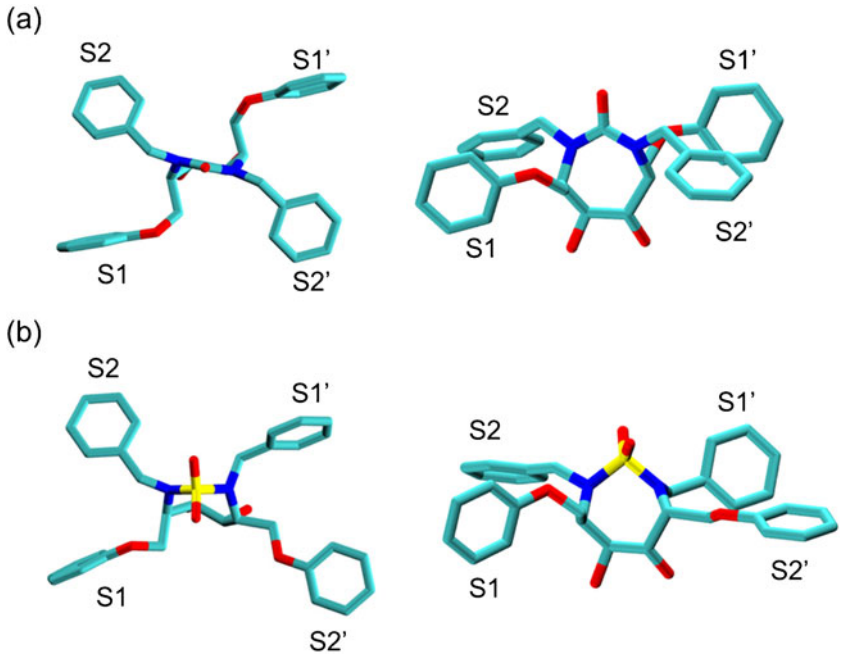

Fig. 2 **a** Symmetric conformation of cyclic urea inhibitor (heavy atoms only) and **b** non-symmetric conformation of cyclic sulfamide inhibitor (heavy atoms only). View from above on left, side view on right. R_2 branches from N at scaffold positions 2 and 7, while R_1 branches from C at positions 3 and 6 (see Fig. 1)

model. Importantly, the barrier to inter-conversion between the symmetric and non-symmetric forms of this class of inhibitors has been investigated.

Molecular dynamics (MD) calculations have additionally been carried out to identify the preferred conformation when the inhibitors are free in solution and when bound to HIV protease. Simulated annealing modelling additionally provides a measure of the ease of inter-conversion of the symmetric and non-symmetric forms of the free inhibitors. Results from the ab initio and DFT calculations are also combined with molecular mechanics-Poisson-Boltzmann surface area (MM-PBSA) and analogous generalized-Born (MM-GBSA) results in order to include an estimate of inhibitor strain energy (difference in energy of symmetric and non-symmetric forms) in binding, which is not included in the single trajectory MM-PB(GB)SA approach.

A number of inhibitor scaffold branches were considered in an attempt to identify whether specific conformations are favoured due to the scaffold itself, or as a result of branching. Hulten and co-workers [17] only considered methyl branches with a cyclic sulfamide inhibitor. Here we have considered a range of branched cyclic inhibitors, which has allowed the conformational preference of the bare seven-membered cyclic scaffold (de-branched) itself to be determined. In this manner we have tested the efficacy of employing simplified model systems in place of the fully branched inhibitor, which has ramifications for future predictive studies. The availability of a model inhibitor system could be useful for understanding the process of binding (conformational requirements for the inhibitors) and could aid further development of new cyclic inhibitors of HIV protease.

Computational methods

Structure preparation

Crystal structures of HIV protease complexed with symmetric urea (PDB ID: 1ajx) and non-symmetric sulfamide (PDB ID: 1ajv) inhibitors were used as templates for the generation of all symmetric and non-symmetric inhibitors, respectively (both bound and free in solution). These inhibitors were selected since they have the same side chains, such that variation in the inhibitor binding energies could more readily be attributed to either the conformational symmetry or scaffold type. For both the urea and sulfamide, two phenoxymethyl branches sit in the S1 subsites, with benzyl branches in the S2 subsite (Fig. 1). Three-dimensional structures were generated for four classes of inhibitors: symmetric and non-symmetric cyclic urea inhibitors, and symmetric and non-symmetric cyclic sulfamide inhibitors. Each inhibitor was generated from the appropriate template structure within Maestro [20], which were subsequently energy minimized using the universal force field (UFF) [21]. These inhibitor structures were used as starting structures in quantum chemical calculations.

The three model inhibitor systems considered throughout this work were: (i) the complete inhibitors (labelled fully branched), (ii) inhibitors whereby the side chains were replaced with methyl groups (labelled methyl branched, R, R' = CH_3 in Fig. 1), and (iii) inhibitors with side chains removed (labelled de-branched, R, R' = H in Fig. 1). For the DFT investigation of the free inhibitors in solution an additional isopropyl branched model system was considered (labelled isopropyl branched, R, R' = $CH(CH_3)_2$). The

model systems retain the scaffold symmetry of the complete inhibitors, as illustrated in Fig. 3.

The template PDB structures were used to generate all HIV protease-ligand complexes. Complexes with the model inhibitor systems were produced within Maestro [20] by first superimposing the new inhibitor onto the template structure of the same inhibitor symmetry. For example, the HIV protease complex with the symmetric form of the sulfamide inhibitor was generated from the 1ajx structure (with a bound symmetric urea inhibitor) by superimposing the sulfamide molecule onto the urea, before deleting the urea and leaving the new HIV protease complex with the symmetric sulfamide inhibitor docked into the same position as the template structure.

Ab initio and density functional theory calculations

Geometries of all inhibitors were optimized using B3LYP [22, 23] and M06-2X [24] functionals with a self-consistent reaction field (SCRF) in order to provide a representation of the free inhibitor in solution. Default optimization parameters were used, with the root-mean square (RMS) of the force required to be less than 3.0×10^{-3} E_h/Bohr. Stationary points were characterized as minima by calculating the Hessian matrix analytically at the same level of theory. Thermodynamic corrections, including zero-point energies (ZPE), were taken from these calculations (standard state of T=298.15 K and p=1 atm). Two basis sets were employed for a comparison of optimized geometries; 6–31G(d) [25] for comparison with the work of Hulten et al. [17], and cc-pVTZ [26], which corresponds to the basis set used to parameterize the partial charges in the ff03 force field that was used in the molecular dynamics (MD) study. Two solvent environments were also considered; diethyl ether with a dielectric constant of 4.2 to represent a protein-like environment [27], and water with a dielectric constant of 78 to represent a polar solvent environment. Solvent effects were modelled with the integral equation formulation of the polarizable continuum model (IEFPCM) [28–30]. Geometry optimization of transition states employed the quadratic synchronous transit (QST) approach [31]. Single-point energy calculations, including IEFPCM-SCRF, were carried out at the respective geometries with B3LYP, M06-2X [24], MP2, SCS-MP2 [32] and SOS-MP2 [33] methods. For a subset of inhibitors, dispersion-corrected B3LYP (B3LYP-D3) [34] calculations were carried out within GAMESS (version 2012, 1 May) [35]. Additional CCSD and CCSD(T) calculations were carried out at the B3LYP/cc-pVTZ optimized structures for the de-branched and methyl branched inhibitors. The 6-311+G(d,p) [36, 37], TZVP [38, 39], def2-TZVPP [40] and def2-QZVP [40] basis sets were used for single point energies. For simplicity, the def2- label is omitted in the following, with the def2-TZVPP and def2-QZVP basis sets labelled TZVPP and QZVP, respectively. For the systems considered in this work, the def2-QZVP and def2-QZVPP basis sets are equivalent. Unless noted, all calculations were carried out in Gaussian 09 [41] with default SCRF parameters (for ether and water solvents). For technical reasons, atomic charge parameters for MM calculations were taken from geometries optimized within Gaussian 03 with default SCRF parameters for water and ether solvents [42].

MD simulations

For MD simulations of the full inhibitor bound to HIV protease, inhibitor geometries were optimized at both the B3LYP/6-31G(d) and B3LYP/cc-pVTZ levels of theory. In each case an implicit solvent model (SCRF) was employed, with separate ether and water type solvent models. Atomic charge parameters were determined by first generating Merz-Kollman [43] electrostatic potentials (ESP) at the same level of theory as used in the geometry optimization, which were subsequently employed in the restrained ESP (RESP) methodology [44] as implemented in AMBER 10 [45]. All other inhibitor parameters were defined within the gaff force field [46]. Parameters for HIV protease were defined within the ff03 force field [47]. As recommended [48–50], one of the catalytic aspartic acids (Asp25) of HIV protease was protonated while the other (Asp25') was deprotonated. Each HIV protease-inhibitor system was solvated in an octahedral SPC/E box that extended at least 12 Å

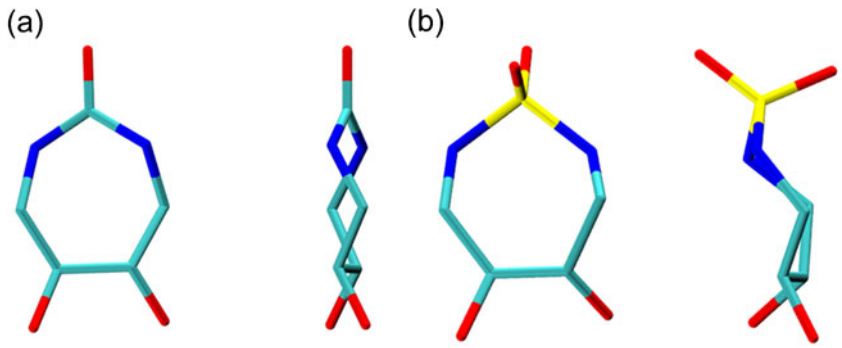

Fig. 3 De-branched **a** cyclic urea structure highlighting its symmetric nature, and **b** cyclic sulfamide structure highlighting its non-symmetric nature. In each case the inhibitors are illustrated with a side view (*left*) and front view (*right*). Only non-hydrogen atoms are shown

from the complex, and the system was charge neutralized with the inclusion of five chloride ions. For MD simulations of the free inhibitors in solution, solvation was accomplished by extending the octahedral SPC/E box at least 12 Å from the inhibitor.

The MD simulations were performed within AMBER (*sander* and *pmemd*) using the SHAKE [51] and PME [52] methodologies with periodic boundary conditions and a cut-off of 12 Å. All simulations were performed in duplicate and followed a six-stage protocol. The first three stages served to systematically minimize the energy of the system. An initial 2000-step minimization was carried out (1000 steepest-descent (SD) then 1000 conjugate gradient (CG)) with 2.0 $kcal\,mol^{-1}$ restraints placed on all heavy atoms. A 2000-step minimization followed with restraints (2.0 $kcal\,mol^{-1}$) placed only on non-hydrogen atoms from the protein. The third minimization stage involved a 2500-step minimization (1000 SD and 1500 CG) with all restraints removed. The systems were then heated to 300 K using Langevin dynamics [53] in a slow, stepwise fashion (60 K every 3 ps). During heating, 2.0 $kcal\,mol^{-1}$ restraints were placed on all atoms in both the protein and inhibitor. Equilibration was then performed in the NPT ensemble for 50,000 steps with all restraints removed (a total of 100 ps with a 2 fs timestep). The temperature was controlled using a Langevin dynamics temperature scheme and pressure by an isotropic position-scaling algorithm. Finally, a production phase of 5 ns was performed using the same conditions as the equilibration phase. To ensure that duplicate simulations did not synchronize, random seeds were used to initialize velocities, while snapshots were output every 10 ps to ensure that the snapshots were not correlated [54].

For MD simulations of the methyl branched and full inhibitor free in solution, atomic partial charges were obtained from B3LYP/cc-pVTZ optimized geometries in a water-like solvent environment (SCRF), employing the same ESP and RESP methodologies. The MD simulations followed the general protocol outlined above for the complex, although with a production phase of 50 ns. Separate simulations at higher temperatures were also carried out, with heating to 400, 500 and 600 K implemented in the same slow, stepwise fashion (60 K every 3 ps) using Langevin dynamics [53]. A production phase of 5 ns was utilized for the simulations at temperatures greater than 300 K.

MM-PB(GB)SA binding free energies

The MM-PB(GB)SA methodologies [55, 56] were utilized to calculate inhibitor binding free energies for the receptor-ligand complex,

$$\Delta G_{bind} = \langle G_{complex} \rangle - \langle G_{receptor} \rangle - \langle G_{ligand} \rangle. \quad (1)$$

Free energies for each molecular component (complex, receptor, ligand) are calculated according to,

$$G = E_{mm} + G_{solv} + TS_{mm}, \quad (2)$$

where S_{mm} is the configurational entropy, G_{solv} is the solvation free energy and E_{mm} is the MM energy, which is calculated as,

$$E_{mm} = E_{internal} + E_{elec} + E_{vdW}. \quad (3)$$

Here $E_{internal}$ is the bond, angle and dihedral energies, E_{elec} is the electrostatic energy and E_{vdW} is the van der Waals (vdW) energy. In the single trajectory approach, $E_{internal}$ will cancel between ligand, receptor and complex, simplifying Eq. 3. The G_{solv} term may be separated into polar (electrostatic) and non-polar components,

$$G_{solv} = G_{pol} + G_{np}. \quad (4)$$

The polar component (G_{pol}) is calculated using either a PB or GB method while the non-polar component (G_{np}) is calculated via a linear relationship to the solvent-accessible surface area (SASA),

$$G_{np} = \gamma SASA + \beta, \quad (5)$$

where γ and β are coefficients. The GB model is an approximation to the more rigorous PB model for which it is generally assumed that MM-PBSA energies are superior to MM-GBSA energies.

The PB polar solvation free energies (G_{pol}) were calculated with DelPhi [57] using parse radii [58], a cubic lattice that was 80 % filled by the complex, a grid spacing of 0.5 Å, and 500 linear steps to solve the PB equation. The GB estimates of G_{pol} were calculated using the GB model developed by Onufriev et al. [59]. Both the PB and GB calculations used dielectric constants of 1 and 80 for the interior and exterior of the complex, respectively. For PB non-polar solvation energies (Eq. 5), γ was set to 0.00542 $kcal\,mol^{-1}\,Å^{-2}$ and β to 0.92 $kcal\,mol^{-1}$ [60], while GB non-polar solvation energies were calculated with coefficients of 0.0072 $kcal\,mol^{-1}\,Å^{-2}$ for γ and 0.0 $kcal\,mol^{-1}$ for β [61].

Entropy contributions (S_{mm}) to the binding free energy were not calculated, since it has previously been identified that entropy differences for inhibitors of similar structure should be negligibly small [55, 62, 63]. All counterions and water molecules were removed before calculating binding energies, which were averaged over duplicate simulations in an effort to increase the reliability of calculated binding energies.

Results

When bound to HIV protease, the preferred conformation of cyclic urea and sulfamide inhibitors is the symmetric and

non-symmetric scaffold conformations, respectively. It has been claimed that "cyclic inhibitors have a high degree of pre-organization, and thus there should be minimal differences between the bound and unbound forms" [1, 64]. In the following computational investigation of free inhibitors, we start from the premise that the preferred inhibitor conformations are the same when bound or free in solution, and subsequently analyse the validity of this assumption on the basis of theoretical modelling.

Urea - free inhibitor in solution: ab initio and DFT energies

Conformer relative energies

The smaller de-branched model inhibitor system (R = H) provides an opportunity to investigate the effect of branching as well as to benchmark calculation methods against higher-level computational methods that are not feasible for the larger molecular systems. Relative energies (between the symmetric and non-symmetric conformations) of the DFT optimized structures of the de-branched inhibitors are given in Table 1. For reference, the CCSD(T)/TZVPP results may be considered the most accurate of those presented in Table 1. Additional results are provided in the Online resource (energies at B3LYP/6-31G(d) and M06-2X/6-31G(d) optimized geometries, and additional single-point energies).

All levels of theory predict the symmetric conformation to be the most stable form of the de-branched cyclic urea. The calculated conformer energy differences are relatively independent of the DFT functional (B3LYP or M06-2X) and basis set employed in the geometry optimization (6-31G(d) or cc-pVTZ), and are also relatively independent of the method and basis set employed for single point energy calculations. Solvent effects are more noticeable. Compared to the ether solvent environment, the water solvent preferentially stabilizes the non-symmetric form of the inhibitors (relative energy is less positive). The Δ(ZPE) and $\Delta(\Delta G)$ corrections are consistent in magnitude between methods and basis sets; they are negative in sign and thus reduce the preference of the symmetric form.

Basis set convergence is consistent for the MP2 and coupled cluster methods, with the relative energies converging from above (becomes less positive). With the QZVP basis set, there are only minimal differences between the MP2, SCS-MP2 and SOS-MP2 methods in comparison with the CCSD(T)/TZVPP results (~1.5 kJmol^{-1}). Basis set limit CCSD(T) values for urea may be estimated as 19.5±1.0 kJ mol^{-1} (ether) and 15.0±1.0 kJmol^{-1} (water), where the error bar represents two standard deviations. It is interesting to note that the CCSD(T) difference between TZVP and TZVPP relative energies is within 0.2 kJmol^{-1} of the corresponding MP2 (and SCS-MP2, SOS-MP2) differences. This trend is promising, as it suggests that for branched inhibitors the computationally efficient MP2 method may provide sufficiently accurate relative energies.

Relative energies for the methyl branched inhibitor (R = Me) are also given in Table 1 (additional results are provided in the Online resource). For reference, the CCSD(T)/TZVP results may be considered the most accurate of those presented. At all levels of theory the symmetric form of urea is predicted to be the most stable. The magnitude of the relative energy difference is smaller for the methyl branched inhibitor than for the de-branched case, although for the methyl branched form the observed trends in method, basis set and geometry effect are consistent with those of the de-branched inhibitor. The effect of implicit solvent is significantly reduced in comparison with the de-branched form, with the average difference between water and ether solvent calculated energies being only 0.1 kJmol^{-1}. The Δ(ZPE) and $\Delta(\Delta G)$ corrections are positive in sign, which merely serves to reinforce the preference for the symmetric form.

Basis set convergence is similar to that of the de-branched case, with the MP2 relative energies converging from above (becomes less positive). The TZVP and TZVPP basis set results are estimated to be within 2 and 1 kJmol^{-1} of the basis set limit, respectively. It is possible to estimate CCSD(T)/TZVPP relative energies by addition of the MP2 calculated difference between TZVP and TZVPP results, to the CCSD(T)/TZVP results, yielding relative energies of 6.5 (ether) and 6.4 (water) kJmol^{-1}, respectively. An appropriate error bar would be 1.5 kJmol^{-1}. The SCS-MP2 and SOS-MP2 results are in closer agreement to those from CCSD(T) than are conventional MP2 results, suggesting that the scaled MP2 methods may perform better as the molecular system increases in size.

For the fully branched (complete) inhibitor, calculated relative energies are presented in Table 2. There are several notable trends in comparison with the methyl and de-branched cases. Firstly, the DFT and MP2 methods yield different results: the DFT methods produce energies that are relatively independent of basis set, and preferentially stabilize the non-symmetric form (relative energies are less positive) in comparison to the MP2 methods. The MP2 results for the full inhibitor are in closer agreement to the de-branched inhibitor than the methyl inhibitor, while the DFT relative energies resemble those of the methyl branched inhibitor. It is noted that the 6-311++G(d,p) basis set produces results that are far from the basis set limit, and so it is recommended that this basis set should be used with caution. Secondly, there are significant geometry effects – both the basis set and DFT functional used in the geometry optimization produce striking effects on the calculated energies. The fact that the DFT calculated relative energies differ dramatically whether a B3LYP or M06-2X optimized geometry was employed, suggests that dispersion effects may be considerable in the larger inhibitor (M06-2X includes an account of dispersion).

Table 1 Energy difference between symmetric and non-symmetric forms of de-branched and methyl branched cyclic urea inhibitors (kJ mol^{-1}). Energy is defined relative to the symmetric conformation (positive values indicate that the symmetric form is lower in energy)

	Ether	Water	Ether	Water
De-branched inhibitor	B3LYP/cc-pVTZ geometry		M06-2X/cc-pVTZ geometry	
Geometry opt method [b]	21.1	16.7	19.6	14.8
B3LYP/TZVPP	19.5	14.8	19.9	15.3
M06-2X/TZVPP	18.6	13.6	18.2	13.2
MP2/TZVPP	20.2	15.1	19.7	14.7
SCS-MP2/TZVPP	20.1	15.0	19.9	14.8
SOS-MP2/TZVPP	20.0	14.9	20.0	14.8
MP2/QZVP	19.2	14.1	18.7	13.6
SCS-MP2/QZVP	19.3	14.1	19.1	13.9
SOS-MP2/QZVP	19.3	14.1	19.2	14.0
CCSD/TZVP	21.3	15.8	21.2	15.7
CCSD/TZVPP	20.3	15.3	–	–
CCSD(T)/TZVP	21.6	16.2	21.4	16.0
CCSD(T)/TZVPP	20.7	15.8	–	–
Δ(ZPE)	−1.9	−1.5	−1.4	−1.4
Δ(ΔG)	−3.4	−2.8	−2.3	−2.2
Methyl branched inhibitor	B3LYP/cc-pVTZ geometry		M06-2X/cc-pVTZ geometry	
Geometry opt method [b]	7.5	7.6	4.2	4.0
B3LYP/TZVPP	7.1	7.2	7.0	7.1
M06-2X/TZVPP	3.7	3.5	4.0	3.8
MP2/TZVPP	5.4	5.3	5.8	5.8
SCS-MP2/TZVPP	7.0	6.5	6.9	6.9
SOS-MP2/TZVPP	7.6	7.1	7.5	7.4
MP2/QZVP	5.0	5.0	5.3	5.4
SCS-MP2/QZVP	6.3	6.2	6.5	6.5
SOS-MP2/QZVP	7.0	6.8	7.1	7.0
CCSD/TZVP	8.4	8.3	8.6	8.6
CCSD(T)/TZVP	8.0	8.0	8.3	8.4
Δ(ZPE)	0.6	0.2	1.0	1.4
Δ(ΔG)	1.7	0.6	2.7	3.5

Electronic energies (E_e) unless noted. Geometries and subsequent thermodynamic quantities were calculated at the same level of theory, in the presence of implicit solvent. The ZPE correction, Δ(ZPE), refers to the *change* in relative conformer energies as a result of including ZPE; the ZPE-corrected energy is defined as $E_0 = E_e + \Delta(\mathrm{ZPE})$. The Gibbs' free energy correction Δ(ΔG) refers to the *change* in relative conformer energies if Gibb's free energies are used rather than electronic energies, with the relative conformer Gibb's free energy given as $\Delta G = E_e + \Delta(\Delta G)$

[b] Electronic energy at the level of theory and basis set employed in the geometry optimization

To further test the effect of branching and the efficacy of these DFT functionals in the calculation of relative energies, an isopropyl branched inhibitor was considered with results presented in Table 2. In the same manner as for the full inhibitor, there is a marked difference between energies calculated at the B3LYP and M06-2X optimized geometries. The effect of dispersion was investigated with gas-phase B3LYP and dispersion-corrected B3LYP-D3 calculations (at the B3LYP/6-31G(d) optimized geometries with ether solvent), for which the dispersion effect is estimated to be +6 kJmol^{-1}. That value is similar to the energy difference between the B3LYP and M06-2X optimized geometry energies.

We can conclude that for the larger inhibitors, dispersion effects are significant and that the M06-2X functional provides more reliable geometries than does the standard B3LYP functional. This conclusion is of further consequence for the consideration of thermodynamic corrections, since the B3LYP and M06-2X functionals produce quite different results (6-31G(d) basis set). For the isopropyl branched inhibitor, the M06-2X thermal corrections are positive in sign, which enhances the preference for a symmetric conformation. For the full inhibitor, the M06-2X method yields Δ(ΔG) corrections that are negative in sign, which is opposite that predicted by B3LYP. However, since the electronic energies calculated at the M06-2X geometries are somewhat more positive than those at the B3LYP geometries, then the ΔG relative energies ($E_0 = E_e + \Delta(\Delta G)$) are quite similar, independent of the geometry. With a cc-pVTZ basis set, the B3LYP calculated Δ(ΔG) correction is negative in sign, which suggests that basis set effects may also be important in modelling the larger inhibitor.

Barriers to inter-conversion

An important aspect of considering the energy differences between the symmetric and non-symmetric forms is the

Table 2 Energy difference between symmetric and non-symmetric forms of isopropyl and fully branched cyclic urea inhibitors (kJ mol^{-1}). Energy is defined relative to the symmetric conformation (positive values indicate that the symmetric form is lower in energy)

	Ether	Water	Ether	Water
Isopropyl branched inhibitor	B3LYP/6-31G(d) geometry		M06-2X/6-31G(d) geometry	
Geometry opt method [b]	14.3	14.9	24.5	24.9
B3LYP/TZVPP	10.4	11.3	15.9	16.6
M06-2X/TZVPP	19.5	19.8	22.7	23.2
MP2/TZVP	18.4	19.1	23.2	24.0
SCS-MP2/TZVP	17.5	18.1	22.1	22.8
SOS-MP2/TZVP	17.0	17.6	21.6	22.2
Δ(ZPE)	−0.3	−0.3	2.5	2.2
Δ(Δ*G*)	−1.4	−1.2	3.6	3.2
Full inhibitor	B3LYP/6-31G(d) geometry		M06-2X/6-31G(d) geometry	
Geometry opt method [b]	13.7	13.5	19.1	23.8
B3LYP/TZVP	15.4	15.0	−0.7	5.6
B3LYP/TZVPP	14.2	13.6	−2.5	3.5
M06-2X/TZVP	14.8	14.3	20.6	24.2
M06-2X/TZVPP	13.3	12.6	18.6	21.9
MP2/TZVP	22.0	21.6	36.7	41.2
SCS-MP2/TZVP	21.3	20.8	30.2	34.7
SOS-MP2/TZVP	20.9	20.4	26.9	31.5
MP2/TZVPP	20.4	19.9		
SCS-MP2/TZVPP	19.9	19.2		
SOS-MP2/TZVPP	19.6	18.9		
Δ(ZPE)	−1.4	−0.5	0.3	−0.8
Δ(Δ*G*)	0.2	2.9	−6.8	−7.7
Full inhibitor	B3LYP/cc-pVTZ geometry		M06-2X/cc-pVTZ geometry	
Geometry opt method [b]	7.5	12.2	16.6	20.2
B3LYP/TZVP	9.4	14.5	5.6	6.7
B3LYP/TZVPP	8.5	12.8	3.5	4.3
M06-2X/TZVP	9.4	13.3	20.1	23.9
M06-2X/TZVPP	8.1	11.2	17.4	21.1
MP2/TZVP	15.6	19.9	34.4	40.0
MP2/TZVPP	14.1	18.4		
SCS-MP2/TZVPP	13.9	18.0		
SOS-MP2/TZVPP	13.8	17.8		
Δ(ZPE)	−1.3	−2.2	1.2	0.5
Δ(Δ*G*)	0.4	−1.0	−1.7	−4.2

See Table 1 footnotes

barrier to inter-conversion. Transition state calculations for the methyl branched inhibitor (Table 3) indicate that the barrier height is substantial, being of the order of 50–60 kJ mol^{-1}. Importantly, it is expected that this barrier height is sufficient to preclude any significant degree of inter-conversion between symmetric and non-symmetric conformers of the urea inhibitor when free in solution.

In summary, all methods predict the symmetric conformation of the cyclic urea inhibitor to be more stable (with or without ZPE and Δ*G* corrections), independent of branching. This result is consistent with inhibitor conformations determined from HIV protease-inhibitor crystal structures, although there is no available experimental evidence for the preferred conformation in solution. The effect of branching on conformer preference may be estimated by comparison with the basis set limit result for the de-branched case, which are 20 and 15 kJmol^{-1} for ether and water implicit solvent models, respectively. With an ether solvent, the conformer energy difference for branched inhibitors is smaller in magnitude than that of the de-branched by 2–6 kJmol^{-1}, which suggests that the effect of branching in the cyclic urea is to favour a non-symmetric conformation by 2–6 kJmol^{-1}. With an implicit water solvent, methyl branching favours the non-symmetric form while the larger branched species marginally favour the symmetric conformer, from which it is difficult to estimate a consistent effect of branching.

Table 3 Barrier heights (kJ mol^{-1}) relative to the symmetric form of the methyl branched urea conformer, for inter-conversion between symmetric and non-symmetric conformers

	B3LYP/6-31G(d) geometry		M06-2X/6-31G(d) geometry	
	Ether	Water	Ether	Water
Geometry opt method	50.2	56.6	52.4	63.3
B3LYP/TZVPP	52.8	49.9	54.6	65.1
M06-2X/TZVPP	55.3	60.6	55.0	65.0
MP2/TZVPP	55.5	63.9	55.3	65.0
SCS-MP2/TZVPP	56.8	62.1	55.8	64.1
SOS-MP2/TZVPP	57.5	61.1	56.1	63.7
MP2/QZVP	55.4	59.7	54.8	64.9
SCS-MP2/QZVP	56.7	58.0	55.4	64.1
SOS-MP2/QZVP	57.3	57.1	55.7	63.7
Δ(ZPE)	−0.3	−1.0	0.3	1.3
Δ(Δ*G*)	−0.4	−3.8	−0.2	3.1

See Table 1 footnotes

Since the energy difference and barrier height between the symmetric and non-symmetric forms is relatively large, it is not expected that the non-symmetric form will be present in solution to any great extent. Referring to the title question of whether cyclic HIV protease inhibitors can bind in a non-preferred form, it is expected that the non-symmetric form is not readily available for binding to HIV protease. For the inhibitor to bind in a non-symmetric conformation, the barrier to inter-conversion would need to be overcome before a complex could be formed. Here we have defined the difference in Gibbs' free energies (ΔG) of the symmetric and non-symmetric forms as the strain energy, which is discussed further below. For urea, the strain energy is 15–20 kJmol^{-1}.

Comparison of the model systems suggests that any of these systems may be considered a reasonable model of the full inhibitor system. Moreover, our results indicate that for accurate representations of these inhibitors, one may restrict modelling to the cyclic ureas in the symmetric conformation. For the purpose of predicting the favoured conformation of new cyclic inhibitors, it is recommended that model inhibitor scaffolds contain a minimum of methyl branching for reliable modelling of cyclic scaffolds. It should be noted though, that basis sets of at least TZ quality (with appropriate polarization functions) are still required to produce reliable results, even with such model inhibitor systems.

Sulfamide - free inhibitor in solution: ab initio and DFT energies

Conformer relative energies

The de-branched model sulfamide inhibitor system (R = H) was used to benchmark the performance of the computational methods in the same manner as for the cyclic urea inhibitor. Relative energies (between the symmetric and non-symmetric conformations) of the DFT optimized structures are given in Table 4. Additional results are provided in the Online resource (energies at B3LYP/6-31G(d) and M06-2X/6-31G(d) optimized geometries, and additional single point energies). For reference, the CCSD(T)/TZVPP results may be considered the most accurate of those presented.

For the de-branched sulfamide the calculated energy differences are relatively independent of the DFT functional and basis set employed in the geometry optimization, and are also relatively independent of the method and basis set employed in the single-point energy calculations. Basis set effects are more pronounced than in the case for the urea inhibitor, although solvent effects are smaller in comparison with the urea inhibitor results. The Δ(ZPE) and Δ(Δ*G*) corrections are consistent in magnitude between methods and basis sets; they are typically negative in sign and thus reduce the preference for the symmetric form.

The sulfamide relative energies consistently converge from below (become more positive) with respect to the basis set, which is opposite to the trend observed for the urea inhibitor. The rate of basis set convergence is slower than that of the cyclic urea, with the basis set effect between TZVP and TZVPP being more pronounced (4–5 kJmol^{-1}). It would appear that for the sulfur-containing species, addition of an extra *d* polarization function to the TZVP basis set (i.e. TZVPP) is important, for which it is concluded that the TZVPP basis set is necessary to adequately describe the sulfamides considered in this work. It is expected that the MP2 calculated results with TZVP and TZVPP basis sets are converged to within 5.0 and 0.5 kJmol^{-1} of the basis set limit, respectively. Basis set limit CCSD(T) values for the de-branched sulfamide may be estimated to be 23±2 kJ mol^{-1} (ether) and 21±2 kJmol^{-1} (water).

The difference between TZVP and TZVPP calculated relative energies with CCSD(T) is within 0.5 kJmol^{-1} of the corresponding MP2 (and SCS-MP2, SOS-MP2) differences,

Table 4 Energy difference between symmetric and non-symmetric forms of de-branched and methyl branched cyclic sulfamide inhibitors (kJ mol^{-1}). Energy is defined relative to the symmetric conformation (positive values indicate that the symmetric form is lower in energy)

	Ether	Water	Ether	Water
De-branched inhibitor	B3LYP/cc-pVTZ geometry		M06-2X/cc-pVTZ geometry	
Geometry opt method	23.3	22.2	23.4	22.0
B3LYP/TZVPP	23.7	22.4	23.9	22.6
M06-2X/TZVPP	24.0	22.3	23.8	22.3
MP2/TZVPP	22.6	21.1	22.8	21.5
SCS-MP2/TZVPP	21.6	20.2	21.9	20.7
SOS-MP2/TZVPP	21.1	19.7	21.5	20.2
MP2/QZVPP	22.7	21.3	22.8	21.5
SCS-MP2/QZVP	21.8	20.4	22.0	20.7
SOS-MP2/QZVP	21.3	19.9	21.6	20.3
CCSD/TZVP	17.6	16.1	17.9	16.4
CCSD(T)/TZVP	16.6	15.2	17.0	15.6
CCSD/TZVPP	22.1	20.7	–	–
CCSD(T)/TZVPP	21.3	20.0	–	–
Δ(ZPE)	−0.8	0.2	1.2	0.4
Δ(Δ*G*)	−1.9	0.1	1.4	0.3
Methyl branched inhibitor	B3LYP/cc-pVTZ geometry		M06-2X/cc-pVTZ geometry	
Geometry opt method	−6.4	−4.8	−6.1	−5.0
B3LYP/TZVPP	−5.7	−4.0	−6.1	−4.4
M06-2X/TZVPP	−6.2	−4.4	−4.3	−3.0
MP2/TZVPP	−6.9	−5.1	−7.0	−5.5
SCS-MP2/TZVPP	−6.3	−4.5	−7.9	−6.3
SOS-MP2/TZVPP	−6.0	−4.2	−8.4	−6.7
MP2/QZVPP	−6.9	−5.0	−6.2	−4.6
SCS-MP2/QZVP	−6.3	−4.4	−7.2	−5.5
SOS-MP2/QZVP	−6.0	−4.1	−7.8	−6.0
CCSD/TZVP	−6.4	−4.8	−11.1	−9.5
CCSD(T)/TZVP	−7.6	−6.0	−12.2	−10.6
Δ(ZPE)	0.8	0.5	0.3	0.6
Δ(Δ*G*)	3.6	2.8	1.0	1.7

See Table 1 footnotes

which again suggests that for branched inhibitors addition of the MP2 difference between TZVP and TZVPP energies to the CCSD(T)/TZVP energies could provide an approximate CCSD(T)/TZVPP relative energy. Significant variation is noted in the deviation of MP2 results from those of CCSD(T)/TZVPP: the SCS-MP2 differences (0.5 kJmol^{-1}) are less than half that of MP2 (1.4 kJmol^{-1}), while SOS-MP2 is within 0.1 kJmol^{-1}. It is suggested that for the larger methyl and fully branched inhibitors that the scaled MP2 methods (SCS-MP2 and SOS-MP2) will provide more reliable energies than standard MP2.

In contrast to the observed experimental preference for a non-symmetric (branched) form when bound, for the de-branched sulfamide inhibitor the symmetric form is predicted to be the most stable at all levels of theory. In the case of the de-branched urea inhibitor, the preference for the symmetric conformer is approximately 20 and 15 kJmol^{-1} with an implicit ether and water solvent, respectively. For the de-branched sulfamide the preference for the symmetric form is approximately 20 kJmol^{-1} with either solvent model. It is concluded that the ring scaffold itself has a preference for a symmetric conformation, for which branching is required to produce a non-symmetric conformation, possibly as a result of reduced steric interactions. It has previously been hypothesized that the change in symmetry preference of the cyclic sulfamides is due to a change in the geometry of the central scaffold in comparison with other cyclic inhibitors [15], however our results and analysis for the de-branched scaffold do not support that hypothesis.

Results for the methyl branched sulfamide are presented in Table 4. At all levels of theory considered the non-symmetric form is favoured, which differs from the de-branched case. Solvent polarity effects are minimal, with an implicit water model always favouring a more positive energy difference (favours the symmetric form). The Δ(ZPE) and Δ(Δ*G*) corrections are positive in sign, which favours the symmetric conformation. While the Δ(Δ*G*)

correction is opposite in sign to the relative electronic energy, it is generally not large enough to alter the preference for the non-symmetric form. An exception is with the M06-2X/6-31G(d) result, although with the larger cc-pVTZ basis set the $\Delta(\Delta G)$ correction is significantly reduced in magnitude to the extent that the ΔG relative energy remains negative in sign.

The effect of geometry on DFT calculated relative energies is minimal, however for MP2 and CCSD(T) results the effect is significant. Although the CCSD(T) calculated results may be expected to be the most accurate, limitations of the TZVP basis set results in the de-branched case suggests that the spin-scaled MP2/QZVP results may be more accurate. This is particularly so in light of the effect of DFT functional on the geometry – the B3LYP and M06-2X optimized geometries produce a significant effect on subsequent MP2 and CCSD(T) relative energies. With a B3LYP optimized geometry the TZVP and TZVPP results are estimated to be within 0.8 and 0.3 $kJmol^{-1}$ of the basis set limit, respectively. With an M06-2X optimized geometry, the rate of convergence is much slower. Moreover, there is greater variation amongst the scaled MP2 methods when using the M06-2X optimized geometry. It is not clear (especially with the M06-2X optimized geometries) whether simply adding the MP2 calculated TZVP to TZVPP difference to the CCSD(T)/TZVP result will produce a valid estimate of CCSD(T)/TZVPP energies. Such an approach appears most valid with the B3LYP optimized geometry, for which an approximate CCSD(T)/TZVPP result is -7.5 ± 2.0 (ether) and -5.8 ± 2.0 (water) $kJmol^{-1}$.

Hulten et al. [17] have reported B3LYP/6-31G(d) results for the methyl branched sulfamide inhibitor, from which it was predicted that the non-symmetric conformation was favoured by 10.0 $kJmol^{-1}$. That is 3–5 $kJmol^{-1}$ greater than the estimates in the present work for a B3LYP optimized geometry. All geometry optimizations in the present work have been performed in the presence of implicit solvent, whereas no solvent effects were included in the work of Hulten and co-workers [17]. Solvent effects may well account for the difference between results from the two studies, and moreover gives an indication of the quantitative effect of solvation on the relative energies (*ca*. 4–5 $kJmol^{-1}$).

Results for the fully branched inhibitor are presented in Table 5. Again, the DFT method used to optimize the geometry has a significant effect on calculated energetics, which is most notable in the B3LYP calculated energies. To further investigate the effect of branching an isopropyl branched inhibitor was considered, with results presented in Table 5. Surprisingly, for the isopropyl case there is little effect of geometry on the calculated energies. Dispersion effects (calculated with B3LYP-D3 in the gas phase) are estimated to be *ca*. +5 $kJmol^{-1}$ (favouring the symmetric form), which is greater than the geometry effect on the B3LYP and M06-2X calculated relative energies for the isopropyl case. For the fully branched inhibitor, for which dispersion effects may be considered to be more important, the M06-2X results are generally more positive in comparison with the B3LYP results.

For the full inhibitor the calculated results are not as consistent as in the previous cases. Here the DFT and MP2 results differ significantly, with the DFT methods preferentially stabilizing the non-symmetric form (relative energies are less positive), however there is no uniform prediction of the preferred conformation. The MP2 methods favour the symmetric form more than do the DFT methods, although it should be noted that the scaled MP2 methods reduce the preference. Again, the relative energies converge with basis set to more negative values, so it may be expected that in the basis set limit the non-symmetric form is favoured. It would appear that in the case of the fully branched sulfamide that we cannot definitively state which conformation is preferred when free in solution. While extensive CCSD(T) calculations with TZVPP or QZVP basis sets would shed light on the situation, such calculations are not presently feasible.

The effect of branching on conformer preference is greater for sulfamide than for urea. With B3LYP optimized geometries, the relative conformer energy in the basis set limit is consistently about −7 and −2 $kJmol^{-1}$ with an implicit ether and water solvent, respectively. That is, branching favours the non-symmetric conformation by 20–25 $kJmol^{-1}$. With M06-2X optimized geometries the basis set limit values are not as uniform and the slower basis set convergence makes it difficult to extrapolate with any certainty. Nevertheless, the effect of branching can be estimated to be 20–30 $kJmol^{-1}$ in favour of the non-symmetric conformation. Comparison of the model systems suggests that the debranched system is not an appropriate model of the full system. It is recommended that model inhibitor scaffolds contain a minimum of methyl branching, and that basis sets of at least TZ quality (with appropriate polarization functions) are employed.

Barriers to inter-conversion

Transition state calculations for the methyl branched sulfamide inhibitor (Table 6) indicate that the barrier height is much smaller than for the urea, being of the order of 10–20 $kJmol^{-1}$. Importantly, this lower barrier height may allow inter-conversion between symmetric and non-symmetric conformers of the sulfamide inhibitor when free in solution.

With regard to the question as to why the cyclic sulfamides bind in a non-symmetric conformation when the vast majority of cyclic inhibitors bind in a symmetric manner [15] our analysis indicates that it is the effect of branching rather than the scaffold itself that is responsible for the non-symmetric conformation of the cyclic sulfamide. Branching reduces the relative energy difference between the symmetric and non-

Table 5 Energy difference between symmetric and non-symmetric forms of isopropyl and fully branched and cyclic sulfamide inhibitors (kJ mol^{-1}). Energy is defined relative to the symmetric conformation (positive values indicate that the symmetric form is lower in energy)

	Ether	Water	Ether	Water
Isopropyl branched inhibitor	B3LYP/cc-pVTZ geometry		M06-2X/cc-pVTZ geometry	
Geometry opt method	−2.7	−2.1	6.3	6.9
B3LYP/TZVP	−5.5	−4.6	−5.7	−4.8
B3LYP/TZVPP	−6.0	−5.1	−6.3	−5.5
M06-2X/TZVP	4.6	5.4	5.0	5.9
M06-2X/TZVPP	4.0	4.7	4.2	5.0
MP2/TZVPP	1.8	2.8	2.6	3.7
SCS-MP2/TZVPP	0.2	1.1	0.9	1.9
SOS-MP2/TZVPP	−0.6	0.3	0.0	0.9
Δ(ZPE)	−2.0	−2.1	−2.7	−3.4
Δ(Δ*G*)	−4.2	−3.8	−6.3	−7.5
Fully branched inhibitor	B3LYP/6-31G(d) geometry		M06-2X/6-31G(d) geometry	
Geometry opt method	−11.4	−3.2	1.3	6.6
B3LYP/TZVP	−12.2	−4.7	−23.2	−15.7
B3LYP/TZVPP	−14.5	−7.1	−28.5	−21.0
M06-2X/TZVP	−5.0	−1.1	−1.3	−1.3
M06-2X/TZVPP	−7.6	−3.7	−6.3	−1.5
MP2/TZVP	1.6	5.9	17.0	22.0
SCS-MP2/TZVP	−0.5	4.0	9.1	14.2
SOS-MP2/TZVP	−1.6	3.1	5.1	10.4
MP2/TZVPP	−3.0	1.4		
SCS-MP2/TZVPP	−4.8	−0.1		
SOS-MP2/TZVPP	−5.8	−0.9		
Δ(ZPE)	−0.3	−0.4	−0.5	−6.1
Δ(Δ*G*)	2.6	1.7	−1.3	−6.0
Fully branched inhibitor	B3LYP/cc-pVTZ geometry		M06-2X/cc-pVTZ geometry	
Geometry opt method	−13.8	−6.5	−3.4	−2.1
B3LYP/TZVP	−13.1	−5.7	−20.9	−19.1
B3LYP/TZVPP	−14.4	−7.6	−25.7	−24.4
M06-2X/TZVP	−5.9	−2.0	−1.2	0.5
M06-2X/TZVPP	−7.5	−4.0	−6.0	−4.8
MP2/TZVP	−0.3	4.1	15.9	18.3
SCS-MP2/TZVP	−2.5	2.2	8.6	10.9
SOS-MP2/TZVP	−3.6	1.3	5.0	7.1
MP2/TZVPP	−3.9	0.9		
SCS-MP2/TZVPP	−5.7	−0.8		
SOS-MP2/TZVPP	−6.6	−1.6		
Δ(ZPE)	−0.2	−1.0	0.4	−0.5
Δ(Δ*G*)	2.0	−1.6	−2.2	−4.8

See Table 1 footnotes

symmetric forms favouring the non-symmetric conformation to the extent that with extended branches, the non-symmetric form is favoured.

The title question of this work remains: can cyclic HIV protease inhibitors bind in a non-preferred form? Firstly, when free in solution the ab initio and DFT results suggest that the non-symmetric conformation of the cyclic sulfamide is favoured, which is consistent with inhibitor conformations determined from HIV protease-inhibitor crystal structures. Since the energy difference and barrier height between the symmetric and non-symmetric forms of sulfamide is relatively small, both forms may be present in solution and thus available for binding to HIV protease. That is, it may be possible that the cyclic sulfamide can bind in a non-preferred manner. However, previous experimental investigations have not found any evidence of non-preferred binding of cyclic sulfamides in HIV protease. It is suggested that with appropriate side-chain branches, binding of the symmetric (non-preferred) form might be possible.

Table 6 Barrier heights (kJ mol^{-1}) relative to the symmetric form of the methyl branched sulfamide conformer, for inter-conversion between symmetric and non-symmetric conformers

	B3LYP/6-31G(d) geometry		M06-2X/6-31G(d) geometry	
	Ether	Water	Ether	Water
Geometry opt method [b]	7.7	8.8	8.9	14.8
B3LYP/TZVPP	7.2	9.0	9.7	19.0
M06-2X/TZVPP	7.6	9.3	10.2	17.7
MP2/TZVPP	9.6	9.6	9.2	17.6
SCS-MP2/TZVPP	10.0	10.0	9.0	18.1
SOS-MP2/TZVPP	10.2	10.2	8.9	18.3
Δ(ZPE)	−0.6	−0.8	0.5	2.4
Δ(Δ*G*)	2.9	2.4	4.4	8.5

See Table 1 footnotes

Free inhibitor in solution: MD simulations

While quantum chemical methods are feasible for modelling the free inhibitors in solution, modelling the protease-inhibitor complex necessitates the use of MM methods. It is therefore important to consider MD simulations on the free inhibitors to fill the gap between quantum-chemical calculations of inhibitors in solution and MD simulations of bound inhibitors. Here, 50 ns MD simulations have been performed on the fully branched and methyl branched inhibitors in solution at 300 K. To corroborate the quantum chemical transition state (barrier height) calculations, we have carried out 5 ns simulated annealing simulations on the methyl branched inhibitors at 400, 500 and 600 K. The RMSD of the scaffold atoms to both the starting symmetric and non-symmetric structures were calculated for each simulation as a measure of stability and conformational inter-conversion.

The MD simulations of the fully branched inhibitors free in solution indicated no inter-conversion of scaffold conformations; the starting scaffold conformation (symmetric or non-symmetric) was the only conformation sampled throughout the simulation. The only case where a motion away from the initial structure was observed was for the symmetric sulfamide. Throughout this simulation, a small number of structures were identified in a modified symmetric conformation, however they were instantly followed by structures in the initial symmetric conformation. The RMSDs between the non-symmetric sulfamide simulation and its initial structure are much greater than the other simulations and indicate greater structural variation. During the simulation the non-symmetric sulfamide adopts a conformation that while being non-symmetric, is different from the initial structure (if the scaffold is viewed from the side it has a concave (u-) shape, compared to the stretched z-shaped conformation of the initial structure). When the non-symmetric sulfamide simulation is compared to the structure from the final step of equilibration, the RMSDs are much more comparable to the other simulations. Plots of RMSD over the length of the simulations are provided in the Online resource.

Simulations on the methyl branched urea inhibitors showed no inter-conversion between conformations, independent of the temperature at which the simulations were performed. As the temperature was increased the RMSD to the initial structure (e.g. symmetric conformation) also increased, however the RMSD to the alternative conformation (e.g. non-symmetric conformation) did not decrease. This indicated that inter-conversion did not occur but rather that the simulated movement was to a modified form of the inhibitor with the same general symmetry.

For the sulfamide, inter-conversion was observed when the temperature of the sulfamide simulations was increased to 600 K. For the simulations of the symmetric conformation at 600 K, the greatest proportion of time was spent in a non-symmetric conformation. However, while inter-conversion was also observed in the non-symmetric simulation, the symmetric conformation was only sampled once and for a very short time frame (~0.1 ns). It is suggested that the non-symmetric sulfamide is the more stable form. The increase in the number of inter-conversions observed for the sulfamide simulations compared to the urea simulations correlates with the DFT transition state calculations, where it was found that the activation barrier between the sulfamide conformers was much smaller than that of the urea.

Bound inhibitors: MM-PB(GB)SA binding energies

Knowledge of the favoured conformation of cyclic inhibitors when free in solution is important for an understanding of inhibitor behaviour, however it is equally important to understand the favoured conformations of the inhibitors upon binding to HIV protease. In the bound form only the fully branched inhibitors were considered. Here MD simulations and MM-PB(GB)SA calculations were utilized, as ab initio and DFT methods are not feasible for such a large system, and moreover, MD and MM-PB(GB)SA methods are commonly used with inhibitor-protein systems.

Cyclic urea

MM-PB(GB)SA binding free energies (ΔG) of the cyclic urea inhibitor bound to HIV protease in both the symmetric and non-symmetric conformation are given in Table 7. Crystal structures suggest that only the symmetric form binds to HIV protease. It is relevant to the title question that no interconversion of conformers was observed in any of the simulations. The single trajectory approach of the MM-PB(GB)SA protocol relies on the assumption that the inhibitor conformations sampled in the MD simulations (bound inhibitor) are similar to those that will be sampled when free in solution. In this context, it is important to consider the ab initio and DFT calculated energy difference between the symmetric and non-symmetric conformers (strain energy) when free in solution.

The MM-PBSA binding energies are reasonably similar for the symmetric and non-symmetric forms of the cyclic urea inhibitors. Simulations using the B3LYP/6-31G(d) ligand charges and geometries favoured the experimentally observed symmetric form of the inhibitor, while those with B3LYP/cc-pVTZ ligand charges and geometries favoured the non-symmetric form. Addition of strain energy (conformational preference for a symmetric conformation for the free inhibitor) serves to emphasize the preference for a symmetric form when bound for the B3LYP/6-31G(d) ligand parameters, and changes the preference with the B3LYP/cc-pVTZ ligand parameters to the symmetric form.

With MM-GBSA, ΔBE results suggest the non-symmetric form to be more strongly bound, which is opposite what is expected on the basis of available crystal structures. With B3LYP/6-31G(d) ligand parameters the MM-GBSA binding energy is insufficient to overcome the strain energy, which results in the symmetric form being favoured once strain energy is included. However, with B3LYP/cc-pVTZ ligand parameters the preference for the non-symmetric form is greater than with the 6-31G(d) basis set (i.e. ΔBE is more negative), and coupled with the smaller strain energy, this results in the non-symmetric form being favoured.

Analysis of the components of the binding energies indicates that E_{elec} for the symmetric conformation are always greater in magnitude than that of the non-symmetric conformation, while E_{vdW} of the symmetric conformation are smaller in magnitude. It is of interest to note that ΔG_{solv}(PB) is greater in magnitude for the non-symmetric inhibitors, even though the E_{elec} energy is smaller. However, this is countered by the corresponding increase in E_{vdW} for the non-symmetric inhibitors, which yields final MM-PBSA binding energies that are similar in magnitude.

The inability of the MM-GBSA method to correctly identify the preferred conformation of the bound inhibitor is due to ΔG_{solv}(GB) preferentially stabilizing the non-symmetric conformation. In all four calculations the PB solvation approach yields greater solvation energies for the non-symmetric conformation by 8–16 kJmol^{-1}. In contrast the GB solvation

Table 7 Components of MM-PB(GB)SA calculated binding energies (kJ mol^{-1}) for the fully branched cyclic urea inhibitor bound to HIV protease

Component	Inhibitor basis set and solvent							
	6-31G(d), ether		6-31G(d), water		cc-pVTZ, ether		cc-pVTZ, water	
	sym	non-sym	sym	non-sym	sym	non-sym	sym	non-sym
E_{elec}	−210.3	−209.0	−257.7	−230.4	−205.7	−200.9	−251.5	−241.0
E_{vdW}	−259.1	−274.4	−257.3	−281.2	−257.4	−281.2	−254.9	−278.9
ΔG_{solv}(PB)	327.0	342.3	372.1	380.5	324.4	341.9	368.3	378.5
ΔG_{bind}(PB)	−142.4	−141.1	−142.9	−131.1	−138.7	−140.3	−138.1	−141.4
ΔBE(PB) [b]		1.3		11.8		−1.6		−3.3
ΔG_{solv}(GB)	221.6	225.9	264.4	242.0	222.0	217.3	264.5	248.1
ΔG_{bind}(GB)	−247.7	−257.5	−250.6	−269.5	−241.1	−264.9	−241.9	−271.8
ΔBE(GB) [b]		−9.2		−18.9		−23.8		−29.9
Strain energy [c]		20.0		22.2		14.4		17.0
ΔBE(PB) + strain		21.3		34.0		12.8		13.7
ΔBE(GB) + strain		10.9		3.2		−9.4		−12.9

Inhibitor atomic charges are calculated with the RESP procedure from B3LYP calculations using the tabled basis set and solvent, at B3LYP optimized geometries with the same basis set and solvent. Here 'sym' and 'non-sym' refer to symmetric and non-symmetric conformations, respectively. $\Delta G_{bind} = E_{elec} + E_{vdW} + \Delta G_{solv}$. BE refers to binding energy

[b] Δ(BE) relative binding free energies are given relative to the symmetric conformation

[c] Strain energies (ΔG) calculated with SCS-MP2/TZVPP electronic energies and B3LYP calculated $\Delta(\Delta G)$ corrections ($\Delta G = E_e + \Delta(\Delta G)$) using the appropriate solvent and B3LYP optimized geometry. SOS-MP2/TZVPP values are on average 0.2 kJmol^{-1} smaller

approach yields greater solvation energies for the non-symmetric conformation only with the B3LYP/6-31G(d) ether solvent inhibitor parameters; in all other cases the GB solvation energy is greater for the symmetric conformation. This preferential stabilization effect is of the order of 5–22 kJmol^{-1} for the four approaches considered. When combined with a decrease in E_{elec} and increase in E_{vdW}, this allows for a significant favouring of the non-symmetric form, especially when the charges were calculated in a water-type solvent.

It is observed that the magnitudes of the electrostatic and solvation energies are greater for the inhibitors produced with the water solvent model than for the ether solvent model, while there was little difference between the different basis sets. It is expected that this is a result of variation in the atomic charges that arises from using different solvent models. We have recently reported a detailed study on the significant effect that atomic charge methods may have on subsequent MM-PB(GB)SA binding energies [65].

It is possible to hypothesize that improved binding energies may be achieved by taking MM-PB(GB)SA binding energy estimates from MD simulations performed with atomic charges generated in an ether-type solvent environment, and including ab initio calculated strain energies carried out in an implicit water solvent environment. That is, the different solvent environments of the bound and free inhibitor are being considered. Resultant MM-PBSA energies are 23.4 and 15.4 kJmol^{-1} for the 6-31G(d) and cc-pVTZ geometries, respectively, while MM-GBSA energies are 13.0 and −6.8 kJmol^{-1}, respectively. With this approach, all calculations apart from MM-GBSA with B3LYP/cc-pVTZ geometries correctly predict the symmetric conformation to be most stable when bound to HIV protease. It would appear that the solvent environment (while modelled implicitly) has a significant effect on calculated binding energies.

Cyclic sulfamide

Calculated binding energies for the different conformations of the cyclic sulfamide inhibitor bound to HIV protease are presented in Table 8. Both MM-PBSA and MM-GBSA methods predict the non-symmetric conformation to be favoured, in agreement with experiment. The difference between the two conformations is greater with the MM-GBSA method. As for the cyclic urea simulations, a greater E_{elec} contribution was noted for the symmetric conformation, while a greater E_{vdW} contribution was predicted for the non-symmetric form. These opposing trends largely cancel leaving the solvation energies as the determining factor of stability, which favours the non-symmetric conformation. The strain energy is generally negative in sign and typically two orders of magnitude smaller than the calculated binding energy itself (ΔG_{bind}), which results in no change of preferred conformation with the inclusion of strain energy.

If a mixed solvent approach is considered for sulfamide in the same manner as for the urea, the MM-PBSA energies are −13.9 and −18.4 kJmol^{-1} for the 6-31G(d) and cc-pVTZ

Table 8 Components of MM-PB(GB)SA calculated binding energies (kJ mol^{-1}) for the cyclic sulfamide inhibitor bound to HIV protease

Component	Inhibitor basis set and solvent							
	6-31G(d), ether		6-31G(d), water		cc-pVTZ, ether		cc-pVTZ, water	
	sym	non-sym	sym	non-sym	sym	non-sym	sym	non-sym
E_{elec}	−270.7	−230.3	−299.5	−282.0	−243.9	−197.9	−275.7	−233.3
E_{vdW}	−269.5	−290.4	−264.4	−285.3	−273.5	−290.9	−263.9	−285.4
ΔG_{solv}(PB)	387.3	352.4	412.9	400.8	371.7	327.1	386.1	362.8
ΔG_{bind}(PB)	−152.9	−168.4	−151.0	−166.5	−145.7	−161.7	−153.5	−155.9
ΔBE(PB) [b]		−15.5		−15.5		−16.0		−2.3
G_{solv}(GB)	252.2	212.1	282.6	252.0	242.6	188.1	264.9	220.2
ΔG_{bind}(GB)	−287.9	−308.7	−281.3	−315.4	−274.7	−300.6	−274.6	−298.5
ΔBE(GB) [b]		−20.8		−34.0		−25.9		−23.9
Strain energy [c]		−2.2		1.6		−3.7		−2.4
ΔBE(PB) + strain		−17.7		−13.9		−19.7		−4.7
ΔBE(GB) + strain		−23.0		−32.4		−29.6		−26.3

Inhibitor atomic charges are calculated with the RESP procedure from B3LYP calculations using the tabled basis set and solvent, at B3LYP optimized geometries with the same basis set and solvent. Here 'sym' and 'non-sym' refer to symmetric and non-symmetric conformations, respectively. $\Delta G_{bind} = E_{elec} + E_{vdW} + \Delta G_{solv}$. BE refers to binding energy

[b] Δ(BE) relative binding free energies are given relative to the symmetric conformation

[c] Strain energies (ΔG) calculated with SCS-MP2/TZVPP electronic energies and B3LYP calculated $\Delta(\Delta G)$ corrections ($\Delta G = E_e + \Delta(\Delta G)$) using the appropriate solvent and B3LYP optimized geometry. SOS-MP2/TZVPP values are on average 0.9 kJmol^{-1} more negative

geometries, respectively, while the MM-GBSA energies are −19.1 and −28.3 kJmol^{-1}, respectively. This follows the same trend for the standard MM-PB(GB)SA energies, and reinforces the conclusion that it is the non-symmetric form of the sulfamide inhibitor that binds to HIV protease.

It appears that employing the GB approach in the calculation of solvation energies (ΔG_{solv}) favours the non-symmetric forms of both inhibitors. As a result, it would appear to be important to include strain energy in the binding energies, and moreover, that consideration should be given to the inhibitor environment when selecting the method (or parameters) used to calculate the strain and binding energies.

Conclusions

The work here presents a thorough analysis of conformer energetics of cyclic urea and cyclic sulfamide inhibitors of HIV protease in different environments using quantum chemical calculations, as well as MD simulations followed by MM-PBSA calculations. The primary conclusions that arise from this work are:

- For the cyclic urea inhibitor, quantum chemical calculations predict that the symmetric conformation is favoured when free in solution, independent of branching. For the cyclic sulfamide inhibitor, there is a fine balance between the symmetric and non-symmetric forms, with the non-symmetric form predicted to be marginally more stable in the case of the full inhibitor.
- Transition state modelling of methyl-branched inhibitors suggests that the barrier to inter-conversion of conformations for urea is 50–60 kJmol^{-1}, while for the sulfamide it is 10–20 kJmol^{-1}. It is concluded that inter-conversion will not occur in the case of the urea, but that inter-conversion may well occur with the sulfamide.
- In both cases the central scaffold (de-branched inhibitor) is predicted to strongly favour a symmetric conformation. Branching typically favours the non-symmetric conformation, for which an appropriate choice of branching may result in the non-symmetric conformation becoming favoured.
- For free inhibitors, a methyl-branched system appears to be an effective model system for the full inhibitor for both cyclic urea and sulfamide inhibitors, which provides an efficient and effective avenue for future theoretical studies. It is concluded that a de-branched cyclic sulfamide is not an appropriate model system for studies of cyclic sulfamide inhibitors.
- MM-PBSA predicts the sulfamide to bind to HIV protease in the non-symmetric form, and cyclic urea to bind in the symmetric form, consistent with available X-ray data. MM-GBSA predicts the sulfamide to bind in a non-symmetric conformation, with urea also binding in a non-symmetric conformation. The GB method preferentially favours non-symmetric binding relative to the PB approach (for both urea and sulfamide) as a result of the GB calculated ΔG_{solv} component.
- Inclusion of strain energy to MM-PB(GB)SA binding energies is significant in determining the preferred form of bound inhibitors. Inclusive of strain energy, the MM-PBSA predicts the bound form to be the same as the X-ray structure (for all ligand charges and geometries). With MM-GBSA, all except the B3LYP/cc-pVTZ ligand parameters predict the bound form to be the same as the X-ray structure.

Can these inhibitors bind in a non-preferred conformation? Our results indicate that when free in solution, the cyclic sulfamide inhibitor may well be present in the non-preferred symmetric conformation in addition to the non-symmetric form, while this is not likely in the case of the urea. Subsequently, the cyclic sulfamide is predicted to be the more likely candidate to bind in the non-preferred manner, although MM-PB(GB)SA calculations suggest that the non-symmetric form is preferentially bound. To date, the hypothesis of non-preferred binding to HIV protease has only been tested (unsuccessfully) with cyclic sulfamides [18]. It is hoped that the present results will stimulate further experimental investigation.

Acknowledgments D.O. was supported by an Australian Postgraduate Award (APA) scholarship. The authors acknowledge support from the National Computational Infrastructure National Facility (NCI-NF), Victorian Partnership for Advanced Computing (VPAC), Victorian Life Science Computing Initiative (VLSCI) and the high-performance computing facility of La Trobe University.

References

1. Lam PYS, Ru Y et al (1996) Cyclic HIV protease inhibitors: synthesis, conformational analysis, P2/P2′ structure-activity relationship, and molecular recognition of cyclic ureas. J Med Chem 39:3514–3525
2. Cram DJ (1986) Preorganization—from solvents to spherands. Angew Chem Int Ed 25:1039–1057
3. Cram DJ (1988) The design of molecular hosts, guests and their complexes. Science 240:760–767
4. Jadhav PK, Woerner FJ et al (1998) Nonpeptide cyclic cyanoguanidines as HIV-1 protease inhibitors: synthesis, structure-activity relationships, and x-ray crystal structure studies. J Med Chem 41:1446–1455
5. Abbenante G, March DR, Bergman DA, Hunt PA, Garnham B, Dancer RJ, Martin JL, Fairlie DP (1995) Regioselective structural and functional mimicry of peptides. Design of hydrolytically-stable cyclic peptidomimetic inhibitors of HIV- 1 protease. J Am Chem Soc 117:10220–10226
6. Zhao C, Sham HL et al (2005) Synthesis and activity of N-acyl azacyclic urea HIV-1 protease inhibitors with high potency against multiple drug resistant viral strains. Bioorg Med Chem Lett 15:5499–5503

7. Lam PYS, Jadhav PK et al (1994) Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. Science 263:380–384
8. Ala PJ, DeLoskey RJ et al (1998) Molecular recognition of cyclic urea HIV-1 protease inhibitors. J Biol Chem 273(20):12325–12331
9. Ala PJ, Huston EE, Klabe RM, Jadhav PK, Lam PYS, Chang C-H (1998) Counteracting HIV-1 protease drug resistance: structural analysis of mutant proteases complexed with XV638 and SD146, cyclic urea amides with broad specificities. Biochemistry 37:15042–15049
10. Ala PJ, Huston EE et al (1997) Molecular basis of HIV-1 protease drug resistance: structural analysis of mutant proteases complexed with cyclic urea inhibitors. Biochemistry 36:1573–1580
11. Jadhav PK, Ala PJ, Woerner FJ, Chang C-H, Garber SS, Anton ED, Bacheler LT (1997) Cyclic urea amides: HIV-1 protease inhibitors with low nanomolar potency against both wild type and protease inhibitor resistant mutants of HIV. J Med Chem 40:181–191
12. Ishima R, Gong Q, Tie Y, Weber IT, Louis JM (2010) Highly conserved glycine 86 and arginine 87 residues contribute differently to the structure and activity of the mature HIV-1 protease. Proteins 78:1015–1025
13. Brik A, Wong C-H (2003) HIV-1 protease: mechanism and drug discovery. Org Biomol Chem 1:5–14
14. Hosur MV, Bhat TN et al (1994) Influence of stereochemistry on activity and binding modes for C2 symmetry-based diol inhibitors of HIV-1 protease. J Am Chem Soc 116(3):847–855. doi:10.1021/ja00082a004
15. Backbro K, Lowgren S et al (1997) Unexpected binding mode of a cyclic sulfamide HIV-1 protease inhibitor. J Med Chem 40:898–902
16. Erickson J (1993) Design and structure of symmetry-based inhibitors of HIV-1 protease. Perspect Drug Discov Des 1(1):109–128. doi:10.1007/bf02171658
17. Hulten J, Andersson HO et al (1999) Inhibitors of the c2-symmetric HIV-1 protease: nonsymmetric binding of a symmetric cyclic sulfamide with ketoxime groups in the P2/P2′ side chains. J Med Chem 42:4054–4061
18. Schaal W, Karlsson A et al (2001) Synthesis and comparative molecular field analysis (CoMFA) of symmetric and nonsymmetric cyclic sulfamide HIV-1 protease inhibitors. J Med Chem 44:155–169
19. Clark T (2011) Predictive modeling of molecular properties: can we go beyond interpretation? In: Comba P (ed) Modeling of molecular properties. Wiley-VCH , Heidelberg, pp 91–106. doi:10.1002/9783527636402.ch7
20. Schrodinger L (2007) New York, NY Maestro, version 8
21. Rappe AK, Casewit CJ, Colwell KS, Goddard WA III, Skiff WM (1992) UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. J Am Chem Soc 114:10024–10035
22. Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B 37(2):785–789
23. Becke AD (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. Phys Rev A 38 (6):3098–3100
24. Zhao Y, Truhlar DG (2008) The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. Theor Chem Acc 120:215–241
25. Hariharan PC, Pople JA (1973) The influence of polarization functions on molecular orbital hydrogenation energies. Theor Chim Acta 28(3):213–222. doi:10.1007/bf00533485
26. Dunning TH Jr (1971) Gaussian basis functions for use in molecular calculations. III. Contraction of (10s6p) atomic basis sets for the first-row atoms. J Chem Phys 55(2):716–723
27. Schutz CN, Warshel A (2001) What are the dielectric "Constants" of proteins and how to validate electrostatic models? Proteins 44:400–417
28. Pomelli C, Tomasi J, Barone V (2001) An improved iterative solution to solve the electrostatic problem in the polarizable continuum model. Theor Chim Acta 105:446–451
29. Cances E, Mennucci B, Tomasi J (1997) A new integral equation formalism for the polarizable continuum model: theoretical background and applications to isotropic and anisotropic dielectrics. J Chem Phys 107:3032–3041
30. Tomasi J, Mennucci B, Cances E (1999) The IEF version of the PCM solvation method: an overview of a new method addressed to study molecular solutes at the QM ab initio level. J Mol Struct (THEOCHEM) 464:211–226
31. Peng C, Ayala PY, Schlegel HB, Frisch MJ (1996) Using redundant internal coordinates to optimize equilibrium geometries and transition states. J Comput Chem 17(1):49–56. doi:10.1002/(sici)1096-987x(19960115)17:1<49::aid-jcc5>3.0.co;2-0
32. Gerenkamp M, Grimme S (2004) Spin-component scaled second-order Møller–Plesset perturbation theory for the calculation of molecular geometries and harmonic vibrational frequencies. Chem Phys Letts 392:229–235
33. Jung Y, Lochan RC, Dutoi AD, Head-Gordon M (2004) Scaled opposite-spin second order Møller–Plesset correlation energy: an economical electronic structure method. J Chem Phys 121:9793–9802
34. Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys 132 (15):154104–154119
35. Schmidt MW, Baldridge KK et al (1993) General atomic and molecular electronic structure system. J Comput Chem 14 (11):1347–1363. doi:10.1002/jcc.540141112
36. Raghavachari K, Binkley JS, Seeger R, Pople JA (1980) Self-consistent molecular orbital methods. 20. Basis set for correlated wave-functions. J Chem Phys 72:650–654
37. Frisch MJ, Pople JA, Binkley JS (1984) Self-consistent molecular orbital methods. 25. Supplementary functions for gaussian basis sets. J Chem Phys 80:3265–3269
38. Schaefer A, Horn H, Ahlrichs R (1992) Fully optimized contracted Gaussian basis sets for atoms Li to Kr. J Chem Phys 97:2571–2577
39. Schaefer A, Huber C, Ahlrichs R (1994) Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. J Chem Phys 100:5829–5835
40. Weigend F, Ahlrichs R (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy. Phys Chem Chem Phys 7:3297
41. Frisch MJ, Trucks GW et al. (2009) Gaussian 09, revision A.1. Gaussian, Inc, Wallingford, CT
42. Frisch MJ, Trucks GW et al (2003) Gaussian 03. Gaussian, Inc., Wallingford CT
43. Singh UC, Kollman PA (1984) An approach to computing electrostatic charges for molecules. J Comput Chem 5(2):129–145
44. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. J Phys Chem 97:10269–10280
45. Case DA, Darden TA et al (2008) AMBER 10. University of California, San Francisco
46. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. J Comput Chem 25:1157–1174
47. Duan Y, Wu C et al (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24:1999–2012

48. Lepsik M, Kriz Z, Havlas Z (2004) Efficiency of a second-generation HIV-1 protease inhibitor studied by molecular dynamics and absolute binding free energy calculations. Proteins 57:279–293
49. Wittayanarakul K, Aruksakunwong O, Saen-oon S, Chantratita W, Parasuk V, Sompornpisut P, Hannongbua S (2005) Insights into saquinavir resistance in the G48V HIV-1 protease: quantum calculations and molecular dynamic simulations. Biophys J 88:867–879
50. Wittayanarakul K, Aruksakunwong O, Sompornpisut P, Sanghiran-Lee V, Parasuk V, Pinitglang S, Hannongbua S (2005) Structure, dynamics and solvation of HIV-1 protease/saquinavir complex in aqueous solution and their contributions to drug resistance: molecular dynamic simulations. J Chem Inf Model 45:300–308
51. Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23:327–341
52. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092
53. Loncharich RJ, Brooks BR, Pastor RW (1992) Langevin dynamics of peptides: the frictional dependence of isomerization rates on N-Acetylalanyl-N'-Methylamide. Biopolymers 32:523–535
54. Cerutti DS, Duke R, Freddolino PL, Fan H, Lybrand TP (2008) A vulnerability in popular molecular dynamics packages concerning langevin and andersen dynamics. J Chem Theory Comput 4 (10):1669–1680
55. Kollman PA, Massova I et al (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 33:889–897
56. Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA (1998) Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. J Am Chem Soc 120 (37):9401–9409
57. Rocchia W, Sridharan S, Nicholls A, Alexov E, Chiabrera A, Honig B (2002) Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. J Comput Chem 23:128–137
58. Sitkoff D, Sharp KA, Honig B (1994) Accurate calculation of hydration free energies using macroscopic solvent models. J Phys Chem 98:1978–1988
59. Onufriev A, Bashford D, Case DA (2000) Modification of the generalized born model suitable for macromolecules. J Phys Chem B 104:3712–3720
60. Massova I, Kollman PA (2000) Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. Perspect Drug Discov Des 18:113–135
61. Gohlke H, Case DA (2004) Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. J Comput Chem 25:238–250
62. Wang J, Morin P, Wang W, Kollman PA (2001) Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of Efavirenz by docking and MM-PBSA. J Am Chem Soc 123:5221–5230
63. Wang W, Lim WA, Jakalian A, Wang J, Wang JM, Luo R, Bayly CI, Kollman PA (2001) An analysis of the interactions between the Sem-5 SH3 domain and its ligands using molecular dynamics, free energy calculations, and sequence analysis. J Am Chem Soc 123:3986–3994
64. Chang C-E, Trylska J, Tozzini V, McCammon JA (2007) Binding pathways of ligands to HIV-1 protease: coarse-grained and atomistic simulations. Chem Biol Drug Des 69:39946
65. Oehme DP, Brownlee RTC, Wilson DJD (2012) Effect of atomic charge, solvation, entropy, and ligand protonation state on MM-PB (GB)SA binding energies of HIV protease. J Comput Chem. doi:10.1002/jcc.23095