J|A|C|S
ARTICLES

Published on Web 05/12/2009

Conformational Variability of Benzamidinium-Based Inhibitors

Xue Li, Xiao He, Bing Wang, and Kenneth Merz, Jr.*

Department of Chemistry, Quantum Theory Project, 2328 New Physics Building, P.O. Box 118435, University of Florida, Gainesville, Florida 32611-8435

Received February 11, 2009; E-mail: merz@qtp.ufl.edu

Abstract: Determining the structure of a small molecule bound to a biological receptor (*e.g.*, a protein implicated in a disease state) is a necessary step in structure-based drug design. The preferred conformation of a small molecule can change when bound to a protein, and a detailed knowledge of the preferred conformation(s) of a bound ligand can help in optimizing the affinity of a molecule for its receptor. However, the quality of a protein/ligand complex determined using X-ray crystallography is dependent on the size of the protein, the crystal quality, and the realized resolution. The energy restraints used in traditional X-ray refinement procedures typically use "reduced" (*i.e.*, neglect of electrostatics and dispersion interactions) Engh and Huber force field models that, while quite suitable for modeling proteins, often are less suitable for small molecule structures due to a lack of validated parameters. Through the use of *ab initio* QM/MM-based X-ray refinement procedures, this shortcoming can be overcome especially in the active site or binding site of a small-molecule inhibitor. Herein, we demonstrate that *ab initio* QM/MM refinement of an inhibitor/protein complex provides insights into the binding of small molecules beyond what is available using more traditional refinement protocols. In particular, QM/MM refinement studies of benzamidinium derivatives show variable conformational preferences depending on the refinement protocol used and the nature of the active-site region.

Introduction

The majority of protein−ligand structures that are used in structure-based drug design (SBDD)[1] efforts are obtained using X-ray crystallography. A crystal structure of a protein−ligand complex provides a detailed view of the spatial arrangement as well as the interactions present between the ligand and the enzyme active site. For more reactive, short-lived species, like enzyme substrates, time-resolved crystallography is a unique tool that identifies structural changes that occur as a reaction proceeds, thereby facilitating the study of reaction intermediates.[2−7]

After the structure of a protein−ligand complex has been solved, the resultant information can be used to assist in the design of new molecules. In what is termed "induced fit", both the ligand and the protein target adjust their conformations (if necessary) in order to form the best complex possible given their mutual structural constraints.[8−11] This information can be critical in many instances because both the protein and the ligand can undergo significant conformational changes with respect to their free or uncomplexed forms. Hence, the accurate prediction of protein-bound conformations of small molecules is an essential aspect of SBDD. However, due to the nature of the small molecules studied in SBDD efforts, typical methods used to refine these structures are less well evolved than those for the protein structures themselves.[12,13] Hence, developing methods that can enhance our understanding of small molecules bound to proteins is of great contemporary interest.

Importantly, the available structures of protein−ligand complexes determined by X-ray crystallography are often used as reference structures to generate and validate pharmacophore models, docking algorithms, and force fields.[14] Recently, protein−ligand structures have been used to estimate the strain induced when a small molecule binds to a receptor, yielding surprisingly large strain energies.[15−17] However, an accurate understanding of the quantitative aspects of the conformational changes a ligand undergoes when it binds to a protein is still surprisingly elusive.[18] The conformation of small molecules may

(1) Brown, D. G.; Flocco, M. M. *Structure-Based Drug Discovery* **2006**, *32*, 32–53.
(2) Schlichting, I. *Acc. Chem. Res.* **2000**, *33*, 532–538.
(3) Moffat, K. *Chem. Rev.* **2001**, *101*, 1569–1582.
(4) Moffat, K. *Nat. Struct. Biol.* **1998**, *5*, 641–643.
(5) Moffat, K. *Acta Crystallogr., Sect. A* **1998**, *54*, 833–841.
(6) Schmidt, M.; Ihee, H.; Paul, R.; Srajer, V. *Methods Mol. Biol.* **2005**, *305*, 115–154.
(7) Petsko, G. A.; Ringe, D. *Curr. Opin. Chem. Biol.* **2000**, *4*, 89–94.
(8) Nicklaus, M. C.; Wang, S. M.; Driscoll, J. S.; Milne, G. W. A. *Bioorg. Med. Chem.* **1995**, *3*, 411–428.
(9) Mooij, W. T. M.; Hartshorn, M. J.; Tickle, I. J.; Sharff, A. J.; Verdonk, M. L.; Jhoti, H. *ChemMedChem* **2006**, *1*, 827–838.
(10) Bostrom, J.; Norrby, P. O.; Liljefors, T. *J. Comput.-Aided Mol. Design* **1998**, *12*, 383–396.
(11) Tirado-Rives, J.; Jorgensen, W. L. *J. Med. Chem.* **2006**, *49*, 5880–5884.
(12) Jack, A.; Levitt, M. *Acta Crystallogr., Sect. A* **1978**, *34*, 931–935.
(13) Tronrud, D. E.; Teneyck, L. F.; Matthews, B. W. *Acta Crystallogr., Sect. A* **1987**, *43*, 489–501.
(14) Grootenhuis, P. D. J.; Vanboeckel, C. A. A.; Haasnoot, C. A. G. *Trends Biotechnol.* **1994**, *12*, 9–14.
(15) Perola, E.; Charifson, P. S. *J. Med. Chem.* **2004**, *47*, 2499–2510.
(16) Benfield, A. P.; Teresk, M. G.; Plake, H. R.; DeLorbe, J. E.; Millspaugh, L. E.; Martin, S. F. *Angew. Chem., Int. Ed.* **2006**, *45*, 6830–6835.
(17) Chang, C. E.; Chen, W.; Gilson, M. K. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1534–1539.
(18) Houk, K. N.; Leach, A. G.; Kim, S. P.; Zhang, X. *Angew. Chem., Int. Ed.* **2003**, *42*, 4872–4897.

10.1021/ja9010833 CCC: $40.75

change significantly when binding to a protein.[19] Ideally, small molecules should not undergo significant conformational changes in order to avoid free energy of binding penalties upon complex formation.

The Cambridge Structural Database (CSD)[20,21] can be used to assist in validating protein−ligand complexes in the Protein Data Bank (PDB)[22−24] through a comparison of the "free" crystalline molecules relative to the structure observed in the complexed state.[25] For example, conformational analyses based on the structures from the CSD have been used to validate conformational minima of molecules observed in protein active sites.[26,27] The use of crystallographic data in this way complements gas-phase *ab initio* calculations since it gives insights into conformational preferences in condensed-phase situations.[27−29] Crystallographic data also underpin many molecular-fragment libraries and programs for generating three-dimensional models from two-dimensional chemical structures. When modeling ligand binding to metalloenzymes, the information from the CSD can be used for guidance regarding the preferred coordination number(s) and geometries for metal-binding warheads.[30] Crystallographically derived information has contributed to many life science software applications, including programs for locating binding "hot spots" on proteins, docking ligands into enzyme active sites, de novo ligand design, molecular alignment, and three-dimensional QSAR.[31,32]

Given the critical importance of understanding both the structure of a protein−ligand complex and the interaction(s) between the binding partners, it is remarkable how little work has focused on accurately refining these complexes. For example, the topologies and parameters for ligands and inhibitors used in the current generation of X-ray refinement programs, such as X-PLOR/CNS,[33] REFMAC,[34] and SHELX,[35] generally do not have enough information to properly model small molecules during refinement against experimental crystallographic data. In general, the ligand electron density is often difficult to fit unambiguously in protein crystallographic studies using low- to medium-resolution data, which is the resolution range where a large portion of protein−ligand complexes are being solved today. Because of these issues, distorted geometries for the ligand and significant clashes between protein and ligand atoms are not uncommon.[36] Hence, unusual ligand conformations in the PDB should be treated with caution and need to be confirmed using multiple experimental and theoretical methods.

Benzamidinium (protonated form of the parent benzamidine) and benzamidinium derivatives are competitive inhibitors widely found in inhibitors of serine proteinases like trypsin, thrombin, and factor Xa.[37,38] Given the similarity in this class of inhibitors and in the target proteins themselves, benzamidinium derivatives offer an interesting compound series to develop and validate QSAR models and our understanding of protein−ligand interactions.[39−42] The amidine group in benzamidinium is chemically similar to the guanidinium moiety of the arginine (Arg) side chain.[43] The interaction between the guanidinium from Arg and the carboxylate from Glu/Asp plays an important role in receptor−substrate and receptor−drug interactions seen in serine proteinases. Indeed, benzamidinium derivatives are known to be potent, nonpeptide antagonists or antimetastatic agents.[44,45] In order to better understand the preferred conformation(s) of benzamidinium ions, we have examined a series of inhibitor/protein complexes involving this class of compounds using *ab initio* QM/MM X-ray refinement protocols and have shown that the originally identified conformations are not always the preferred ones. Indeed, we find that there is a rich "bound" conformational preference for this deceivingly simple molecule and that the nature of the enzyme active site can stabilize conformations found to be unstable in the gas phase.

Methods

1. Selection and Analysis of Benzamidinium-Containing Protein−Ligand Complexes. Prior to refining proteins complexed with benzamidinium and benzamidinium derivatives (shown in Figure 1), we carried out an analysis of the PDB for this class of protein−ligand complex. For this analysis the ligand must contain a benzamidinium moiety without multiple binding conformations due to crystallographic disorder, defined as partial occupancy for any atom from the ligand. The structures with crystallographic resolution below 1.5 Å, with hydrogen atoms defined in the PDB, were also excluded from our selection. It is important to note that the benzamidinium moiety can occur multiple times in a ligand, and these occurrences were counted as independent observations.

2. QM/MM and MM X-ray Refinement of Select Protein−Ligand Complexes. The deposited PDB coordinates of 1RKW, 1RPW, and 1Y3X served as the starting geometries for our QM/

(19) Khan, A. R.; Parrish, J. C.; Fraser, M. E.; Smith, W. W.; Bartlett, P. A.; James, M. N. *Biochemistry* **1998**, *37*, 16839–16845.
(20) Allen, F. H.; Motherwell, W. D. S. *Acta Crystallogr., Sect. B* **2002**, *58*, 407–422.
(21) Allen, F. H. *Acta Crystallogr., Sect. B* **2002**, *58*, 380–388.
(22) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.
(23) Berman, H. M.; et al. *Acta Crystallogr., Sect. D* **2002**, *58*, 899–907.
(24) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *Eur. J. Biochem.* **1977**, *80*, 319–324.
(25) Taylor, R. *Acta Crystallogr., Sect. D* **2002**, *58*, 879–888.
(26) Conklin, D.; Fortier, S.; Glasgow, J. I.; Allen, F. H. *Acta Crystallogr., Sect. B* **1996**, *52*, 535–549.
(27) Allen, F. H.; Harris, S. E.; Taylor, R. *J. Comput.-Aided Mol. Design* **1996**, *10*, 247–254.
(28) Bohm, H. J.; Brode, S.; Hesse, U.; Klebe, G. *Chem−Eur. J.* **1996**, *2*, 1509–1513.
(29) Bohm, H. J.; Klebe, G. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2589–2614.
(30) Harding, M. M. *Acta Crystallogr., Sect. D* **1999**, *55*, 1432–1443.
(31) Tanrikulu, Y.; Schneider, G. *Nat. Rev. Drug Discovery* **2008**, *7*, 667–677.
(32) Tropsha, A.; Golbraikh, A. *Curr. Pharm. Des.* **2007**, *13*, 3494–3504.
(33) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr., Sect. D* **1998**, *54*, 905–921.
(34) Vagin, A. A.; Steiner, R. A.; Lebedev, A. A.; Potterton, L.; McNicholas, S.; Long, F.; Murshudov, G. N. *Acta Crystallogr., Sect. D* **2004**, *60*, 2184–2195.
(35) Sheldrick, G. M. *Acta Crystallogr., Sect. A* **2008**, *64*, 112–122.
(36) Nissink, J. W. M.; Murray, C.; Hartshorn, M.; Verdonk, M. L.; Cole, J. C.; Taylor, R. *Proteins−Struct., Funct. Genet.* **2002**, *49*, 457–471.
(37) Hauptmann, J.; Sturzebecher, J. *Thromb. Res.* **1999**, *93*, 203–241.
(38) Grzesiak, A.; Helland, R.; Smalas, A. O.; Krowarsch, D.; Dadlez, M.; Otlewski, J. *J. Mol. Biol.* **2000**, *301*, 205–217.
(39) Dullweber, F.; Stubbs, M. T.; Musil, D.; Sturzebecher, J.; Klebe, G. *J. Mol. Biol.* **2001**, *313*, 593–614.
(40) Guvench, O.; Price, D. J.; Brooks, C. L. *Proteins−Struct., Funct. Bioinf.* **2005**, *58*, 407–417.
(41) Ota, N.; Stroupe, C.; Ferreira-da-Silva, J. M. S.; Shah, S. A.; Mares-Guia, M.; Brunger, A. T. *Proteins−Struct., Funct. Genet.* **1999**, *37*, 641–653.
(42) Radmer, R. J.; Kollman, P. A. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 215–227.
(43) Tominey, A. F.; Docherty, P. H.; Rosair, G. M.; Quenardelle, R.; Kraft, A. *Org. Lett.* **2006**, *8*, 1279–1282.
(44) Alig, L.; Edenhofer, A.; Hadvary, P.; Hurzeler, M.; Knopp, D.; Muller, M.; Steiner, B.; Trzeciak, A.; Weller, T. *J. Med. Chem.* **1992**, *35*, 4393–4407.
(45) Adler, M.; Davey, D. D.; Phillips, G. B.; Kim, S. H.; Jancarik, J.; Rumennik, G.; Light, D. R.; Whitlow, M. *Biochemistry* **2000**, *39*, 12534–12542.

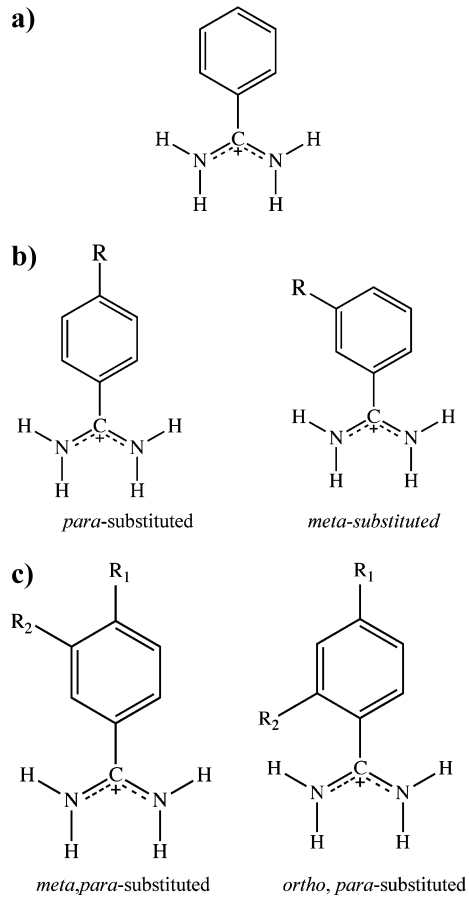

Figure 1. Structures of benzamidinium (a) and benzamidinium derivatives with one (b) and two (c) substituent groups found in protein−ligand complexes.

MM refinements.[46,47] The 1RKW and 1RPW QacR (quaternary ammonium compound repressor) dimer structures had resolutions of 2.62 and 2.90 Å, respectively. In our calculations only one subunit bound with a ligand was used in the refinement, while the other subunit was kept fixed. The sulfate molecules were also kept fixed during the refinement process, while crystal waters were included in the refinement.

The 1Y3X trypsin-ligand structure has a higher resolution of 1.7 Å. There are multiple conformations given for several of the trypsin side chains; however, these side chains are remote from the ligand binding, so in the QM/MM refinements only the main side-chain conformer was considered.

Due to system size limitations of *ab initio* methods, the QM region included the ligand and polar residues and/or water molecule within 3 Å of the ligand itself. The Hartree−Fock (HF) level of theory with the 6-31G* basis set was the *ab initio* method used for the QM region in all cases. All the reflections downloaded from the PDB were considered, of which 5% of the reflections were randomly extracted and used as an external validation set. The QM/MM energy refinements were done using AMBER 10[48] combined with our in-house *ab initio* code QUICK.[49] The X-ray target function was evaluated by CNS (Crystallography and NMR System).[33] The energy target function used in the QM/MMM refinements is given by eq 1,

$$E = E_{\mathrm{QM/MM}} + w_{\mathrm{a}}E_{\mathrm{X\text{-}ray}} \quad (1)$$

where w_a is the weighting factor for the X-ray signal (we generally use a range of 0.1−1.0, but higher values can be used), $E_{QM/MM}$ is the QM/MM energy and gradients of the system, and $E_{X\text{-}ray}$ is the pseudo energy function obtained from CNS along with the gradients. Thus, the gradients from the QM/MM region were merged with the X-ray gradients in order to update the coordinates in the minimization process.

In order to determine the dependence of the ligand conformation on the energy function scheme used in the refinement, the traditional MM-based refinements using Engh−Huber parameters[50] were also carried out. The topology and parameter files used for the ligands within the CNS (Crystallography and NMR System)[33] program were obtained from the Hetero-compound Information Centre-Uppsala[51,52] (Hic-Up server, http://alpha2.bmc.uu.se/hicup/) as was done previously in the reported QacR refinements. [53] Each refinement took about 600−700 refinement steps, and each step involved ∼60−80 atoms and required 40−50 min per cycle on a single 2.4 GHz AMD CPU. These are long runs, but we are parallelizing the code as well as trying different minimizers (we are using AMBER's conjugate gradient optimizer) to reduce the number of refinement cycles in order to significantly decrease the refinement run times.

3. Benzamidinium Torsion Profile Calculations. The torsional profiles were calculated using Gaussian 03.[54] The torsion profiles for free benzamidinium and hydroxyl-substituted benzamidinium were calculated using the 6-31G* and aug-cc-pvtz basis sets at the HF and MP2 levels of theory. In order to obtain the torsion energy profiles, torsional scans were carried out by varying the N−C−C−C (the last two carbon atoms come from the phenyl ring, while the first two atoms are from the amidinium group) torsion angle between 0° and 180°, in 10° intervals. For each structure along the torsion profile, a constrained energy minimization was carried out at fixed N−C−C−C torsion angles while relaxing all other geometric variables.

In order to obtain insights into the conformational preferences of the bound benzamidinium derivatives and to further verify our QM/MM refinement results, we carried out active-site cluster calculations that examined the N−C−C−C torsion angle. The model structures were built from crystal structures taken from the PDB (1RKW, 1RPW, and 1Y3X). Hydrogen atoms were added and then optimized initially using the SANDER module from the AMBER suite of programs. In order to preserve protein−ligand interactions, protein atoms with van der Waals contacts with the ligand were included in the calculations. The resultant cluster models are shown in Figure 2. Starting with these geometries, only the atoms involved in the N−C−C−C torsion angle and the attached hydrogen atoms were allowed to freely rotate, while the remaining atoms were kept frozen at their original positions. These gas-phase cluster calculations were done at the HF/6-31G* level of theory in order to match the method and basis set used in the QM/MM refinement studies themselves. These calculations only gave qualitative insights into the conformational preferences around the initial structure. This is because of the fixed geometries that were used. To examine the role geometric relaxation plays, molecular dynamics (MD) simulations were carried out.

4. Molecular Dynamics Simulations. The QacR-pentadiamidine and QacR-hexadiamidine X-ray structures (PDB codes 1RKW and

(46) Murray, D. S.; Schumacher, M. A.; Brennan, R. G. *J. Biol. Chem.* **2004**, *279*, 14365–14371.
(47) Fokkens, J.; Klebe, G. *Angew. Chem., Int. Ed.* **2006**, *45*, 985–989.
(48) Case, D. A.; et al. *AMBER 10*; University of California: San Francisco, 2008.
(49) He, X.; Ayers, K. B.; Brothers, E. N.; Merz, K. M. *QUICK*; University of Florida: Gainesville, FL, 2008.
(50) Engh, R. A.; Huber, R. *Acta Crystallogr., Sect. A* **1991**, *47*, 392–400.
(51) Kleywegt, G. J.; Jones, T. A. *Acta Crystallogr., Sect. D* **1998**, *54*, 1119–1131.
(52) Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M. *Acta Crystallogr., Sect. A* **1991**, *47*, 110–119.
(53) Brooks, B. E.; Piro, K. M.; Brennan, R. G. *J. Am. Chem. Soc.* **2007**, *129*, 8389–8395.
(54) Frisch, M. J.; et al. *Gaussian 03*, Rev. D.01; Gaussian, Inc.: Wallingford, CT, 2004.

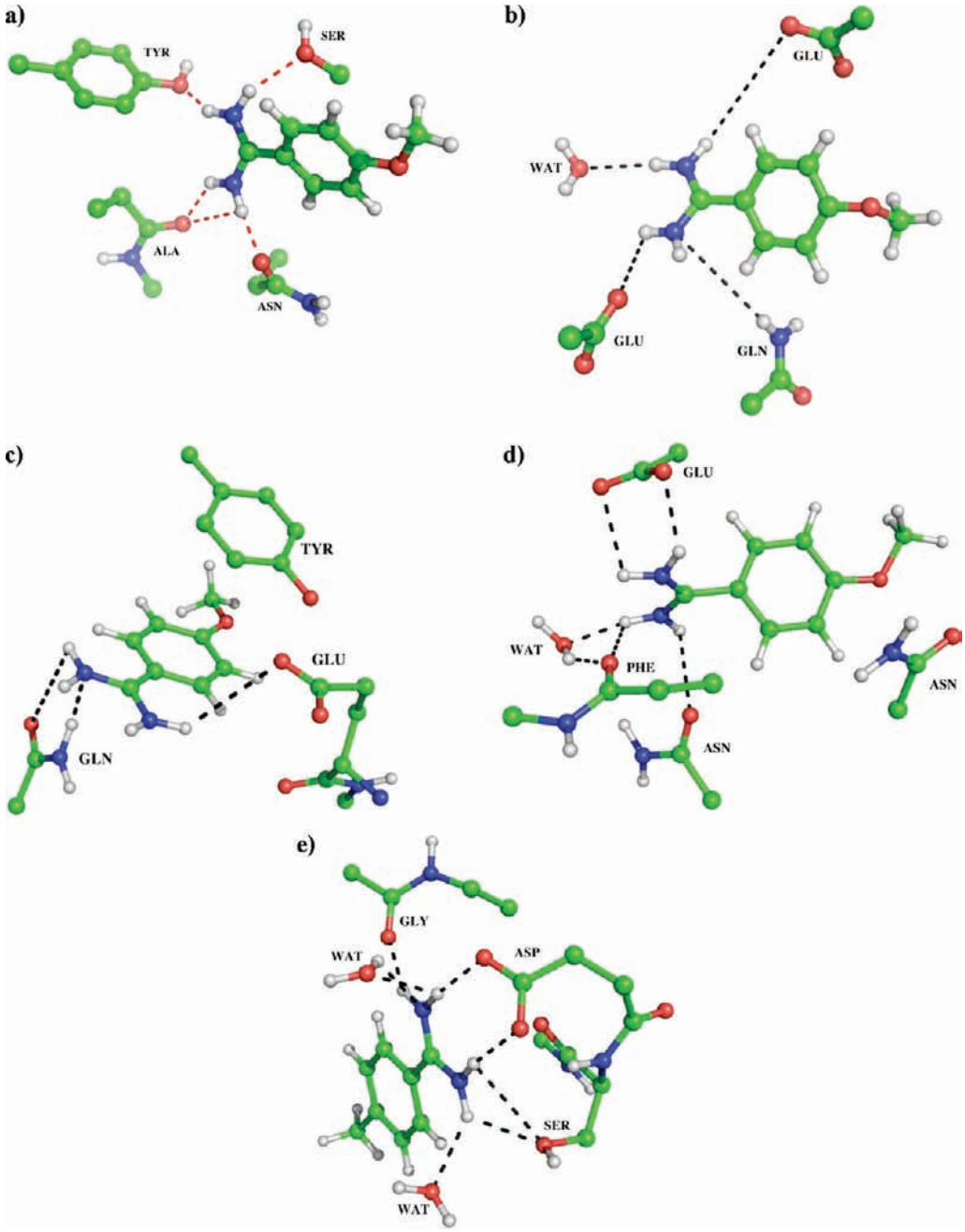

Figure 2. Model cluster structures used in the torsion profile calculations (nonpolar hydrogen atoms are omitted for clarity): (a) 1RKW, PNT binding site 1; (b) 1RKW, PNT binding site 2; (c) 1RPW, DID binding site R6G; (d) 1RPW, DID binding site Et; (e) 1Y3X, UIB binding site.

1RPW, respectively) were used as the starting point for MD simulations using the AMBER force field ff99sb[55] for the protein atoms and gaff[56] for the ligands. In order to investigate the benzamidinium torsional angle preference in the protein−ligand complex, we had to re-parametrize the force field for the dihedral angle between the amidine group and the phenyl ring (N−C−C−C torsion) using the energy profile obtained from high-level quantum mechanical calculations. The ligand partial charges were derived from a two-stage restrained electrostatic potential (RESP) fitting procedure.[57] The entire protein−ligand complex, including all crystallographic waters, was solvated in an octahedral TIP3P[58] water box with each side at least 8 Å from the nearest solute atom. Applying a uniform neutralizing plasma neutralized the net charge of the entire system. The SHAKE algorithm was employed to constrain bonds X−H bonds to their equilibrium values.[59] The systems were minimized and then gradually heated from 0 to 300 K, slowly decreasing the weak restraints on the heavy atoms of the complex. During the last step of equilibration, the restraints were removed entirely, and the production simulations were carried out at 300 K for 8 ns with a 2 fs time step. Constant temperature was maintained using a Berendsen temperature bath with a coupling

(55) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712–725.
(56) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
(57) Wang, W.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049–1074.
(58) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(59) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.

strength of 1.0 ps.[60] Snapshots for subsequent analysis were taken every 2 ps. All simulations were carried out using the PMEMD module from the AMBER package.[48]

Results and Discussion

1. Observed Conformations of Benzamidinium and Benzamidinium Derivatives in Protein−Ligand Complexes. The benzamidinium or benzamidinium derivatives included in the PDB can be divided into three classes on the basis of the substitution of the parent benzamidinium: (a) Unsubstituted amidine, benzamidinium. In this class, there were 87 crystal structures with a total of 153 benzamidinium moieties present in these structures. It should be noted that the same inhibitor may occur multiple times in one crystal structure at different positions. However, the inhibitors with disordered conformations were excluded from the study. (b) One substitution at the *para* or *meta* position of benzamidinium. The *ortho*-substituted benzamidinium was not observed in the crystal structures from the PDB databank. A total of 46 ligands were substituted at the *para* position, including several bivalent dibenzamidinium derivatives. The identities of the four bivalent dibenzamidinium derivatives are abbreviated as BBA,[61,62] PNT,[46] DID,[46] BRN (PBD 2GBY), respectively. In these structures, all substitutions are at the *para* positions. There are 14 crystal structures that contain one substituent in the *meta* position. In the structures 1Y5A and 1Y5U, both *para* and *meta* substitutions were observed in one inhibitor. (c) Two substituents on the parent benzamidinium. The two substituent motifs can have three possible combinations: *m*,*p*-, *o*,*p*-, or *m*,*o*- derivatives. No inhibitor with both *o*- and *m*- substituents was found in the PDB. Only one inhibitor with *p*- and *o*- substitution was observed (PDB ID 1YGC).[63] The intramolecular interaction (2.53 Å heavy atom−heavy atom distance) between the oxygen atom of the *ortho*-hydroxy group and the amidinium nitrogen atom in the inhibitor, G17905, introduces structural and energetic constraints to the rotation of the amidinium group. Hence, G17905 will be excluded from our statistical analysis. Six protein−ligand structures were found in which the inhibitors have substituents at the *para* and *meta* positions of benzamidinium.

In structure-based drug design, important intramolecular geometric parameters are the torsion angles around rotatable bonds, since they influence the overall shape of a molecule far more than do the bond lengths and valence angles. The main conformation difference for the benzamidinium-containing compounds is the torsion angle formed by the N−C−C−C atoms (the latter two C atoms come from the phenyl ring, while the first two atoms come from the amidinium moiety). The conformational histograms of this torsion angle are depicted in Figure 3.

The experimentally found conformational distribution for benzamidinium from the PDB indicates two major preferences for the dihedral angles. The most frequently encountered one is the planar conformation, in which the torsion angles fall into the range of −10° to 10°. The second one is the twisted conformation, which lies outside the previous range (>10° and <−10°). No relationship between the resolution of the structures and the number of occurrences of a specific torsional angle was observed. Both for unsubstitituted and substituted benzamidiniums, similar torsional preferences were observed (see Figure 3a,b).

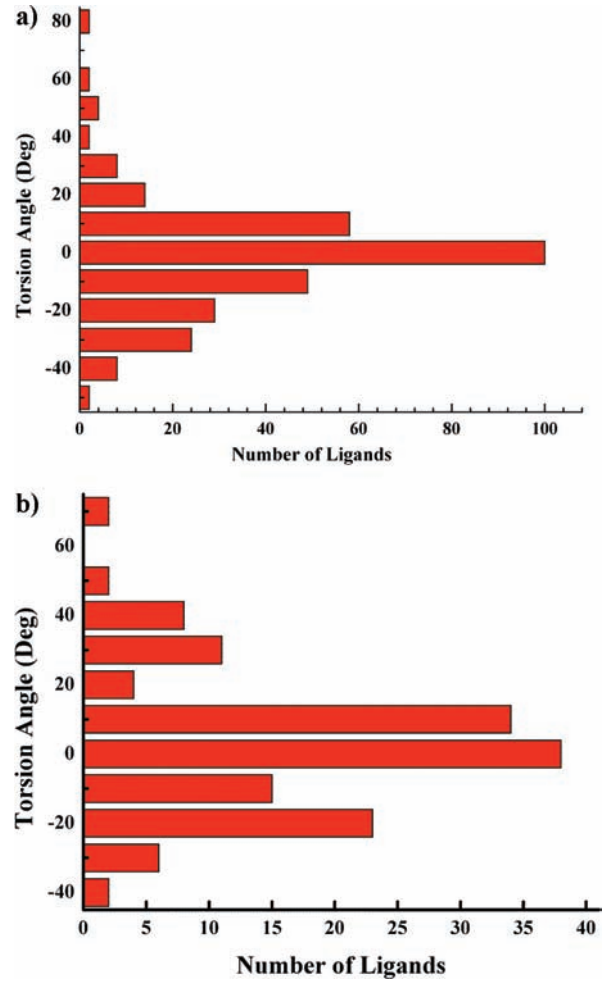

Figure 3. Conformational preference distributions observed in benzamidinium-containing PDB structures. Values of the torsion angle N−C−C−C found in benzamidinium−protein complexes (a) and benzamidinium derivative−protein complexes (b).

The preferred conformation of putative ligands obtained from small-molecule crystal structure data can also be used to validate the observed conformations found in protein−ligand complexes.[8,10,15] Sometimes the torsion distributions of molecular fragments in the CSD compared to the corresponding fragments in protein-bound ligands are similar, which suggests geometry distributions in small-molecule crystal structures generally serve as useful guides to geometric preferences in protein−ligand complexes. However, the structures of benzamidinium and benzamidinium derivatives found in the CSD are different from what we observed in protein−ligand complexes.[64] Only the nonplanar conformation was observed. The average value is 33°, with the lowest value being 15° and the highest 55°. As we mentioned before, due to limitations of the X-ray experiment itself, the structure factors, and the refinement procedure, the preferred or observed conformation of protein-bound ligands may contain some uncertainties. Importantly, structure-based drug design relies heavily on experimental protein structures, taken from the PDB, so understanding the accuracy of the

(60) Berendsen, H. J. C.; Potsma, J. P. M.; van Gunsteren, W. F.; DiNola, A. D.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
(61) Renatus, M.; Bode, W.; Huber, R.; Sturzebecher, J.; Prasa, D.; Fischer, S.; Kohnert, U.; Stubbs, M. T. *J. Biol. Chem.* **1997**, *272*, 21713–21719.
(62) Rauh, D.; Klebe, G.; Stubbs, M. T. *J. Mol. Biol.* **2004**, *335*, 1325–1341.
(63) Olivero, A. G.; et al. *J. Biol. Chem.* **2005**, *280*, 9160–9169.
(64) Zheng, S. private communication.

representation of a given protein−ligand complex is critical for success. Hence, it is worth taking a closer look at predicted protein−benzamidinium geometries in order to determine whether the wide range of torsional preferences is due to the nature of the complexes or due to shortcomings of the refinement protocols.

2. QM/MM Refinement on Protein−Ligand Complexes. The QM/MM approach has been extensively used to study a number of chemical and biological problems.[65−67] This method is typically applied when QM methods are too expensive to use on the entire system while at the same time MM methods themselves are unable to provide an accurate representation.[68−71] Generally, this involves the study of reactive process, but not always. For example, QM/MM methods have been used in X-ray refinement studies in order to improve local geometries or to determine the protonation status of metal-bound ligands.[72−82] In this study, we also used the QM/MM refinement method to enhance our understanding of the preferred ligand conformation in selected protein−ligand complexes.

The first two structures we examined were QacR−ligand complexes where two benzamidinium moieties were connected by either a pentyl or hexyl linker (shown in Figure 4a,b).[46,53,83,84] Each ligand in 1RKW and 1RPW has two positively charged benzamidinium groups bound to QacR in different ways. It is interesting that each torsion angle shows different preferences upon binding to different residues from QacR (from 0° to 70°). The drug-binding pocket of QacR is quite extensive and has been described as having two independent but overlapping drug-binding sites that are designated the rhodamine 6G (R6G) and ethidium (Et) binding sites. A key feature of the binding pocket is the presence of several glutamates and a large number of aromatic residues. Combinations of these residues are used to define the specific small-molecule-binding sites and provide maximal shape, chemical, and electrostatic complementarity.

The structure of the QacR−pentadiamidine complex was refined at 2.6 Å resolution. In the original crystal structure, the pentadiamidine molecule was significantly twisted about its central linker. In the binding site 1 for pentadiamidine (PNT, shown in Figure 5a), no Glu or Asp residues lie within van der Waals/hydrogen contact of the positively charged benzamidinium group. However, a number of hydrogen bonds are formed with neutral residues via oxygen atoms. The nearest contacts were made with the backbone carbonyl oxygen of Ala153 and the side-chain hydroxyl oxygen of Tyr127, both of which engaged in hydrogen-bond interactions with the amidinium group. The side-chain carbonyl oxygen of Asn157 and the hydroxyl oxygen of Ser86 were also proximal to the benzamidinium moiety and contributed to the overall negative electrostatic character of site 1. At the other end of the pentadiamidine-binding site (designated as site 2; see Figure 5b), the benzamidinium moiety forms a Pi-stack with Tyr93 (Tyr93 not shown for clarity) along with several hydrogen-bonding interactions. Glu63 and Glu57 both form hydrogen bonds with the positively charged benzamidinium moiety. One solvent molecule was also observed in the site 2 pentadiamidine-binding pocket.

The preferred value for the dihedral angle of the benzamidinium moiety in each binding site is planar. However, QM/MM refinement shows different preferences in the QacR−PNT complexes. In the PNT−QacR complex (see Figure 6, selected refined structures at $w_a = 0.8$), the torsion angle becomes 34° in binding site 1 and 31° in binding site 2, whereas the torsion angles from the MM refinements are both around 0°. The end-

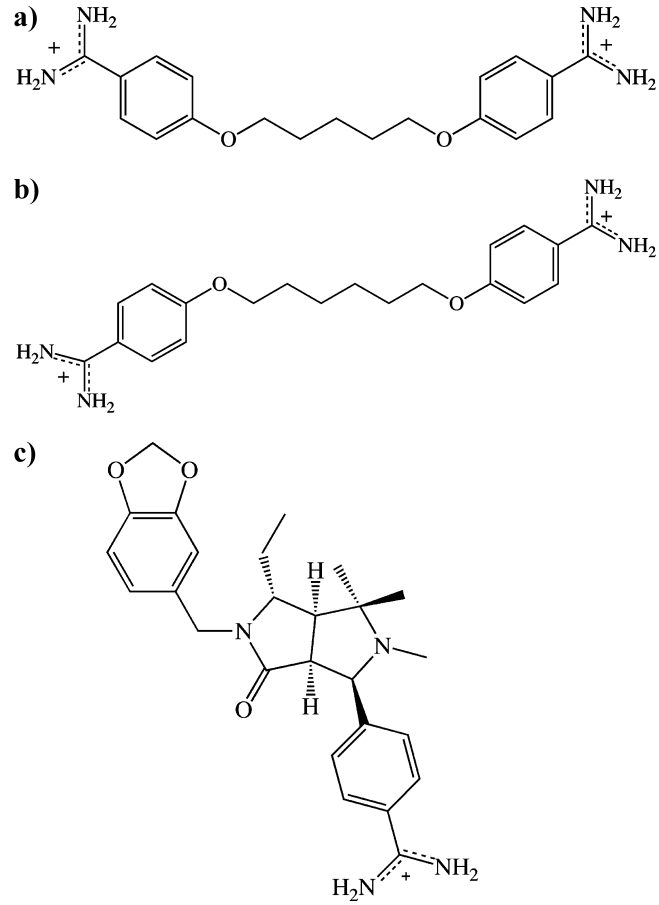

Figure 4. Structures selected for detailed *ab initio* QM/MM refinement: (a) 1,3-bis(4-amidinophenoxy)pentane (pentadiamidine, PNT); (b) 4,4′-[1,6-hexanediylbis(oxy)]bisbenzenecarboximidamide (hexadiamidine, DID); (c) (1*R*,3A*S*,4*R*,8A*S*,8B*R*)-4-{5-(phenyl[1,3]dioxol-5-ylmethyl)-4-ethyl-2,3,3-trimethyl-6-oxo-octahydro-pyrrolo[3,4-*c*]pyrrol-1-yl}benzamidinium (UIB).

(65) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569–585.
(66) Monard, G.; Merz, K. M., Jr. *Acc. Chem. Res.* **1999**, *32*, 904–911.
(67) Ma, S.; Devi-Kesavan, L. S.; Gao, J. *J. Am. Chem. Soc.* **2007**, *129*, 13633–13645.
(68) Gao, J. L.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467–505.
(69) Ryde, U. *Curr. Opin. Chem. Biol.* **2003**, *7*, 136–142.
(70) Talbi, H.; Monard, G.; Loos, M.; Billaud, D. *Synth. Met.* **1999**, *101*, 115–116.
(71) Maseras, F.; Lledos, A.; Clot, E.; Eisenstein, O. *Chem. Rev.* **2000**, *100*, 601–636.
(72) Ryde, U. *Dalton Trans.* **2007**, 607–625.
(73) Ryde, U.; Hsiao, Y. W.; Rulisek, L.; Solomon, E. I. *J. Am. Chem. Soc.* **2007**, *129*, 726–727.
(74) Soderhjelm, P.; Ryde, U. *J. Mol. Struct.−Theochem* **2006**, *770*, 199–219.
(75) Rulisek, L.; Ryde, U. *J. Phys. Chem. B* **2006**, *110*, 11511–11518.
(76) Nilsson, K.; Hersleth, H. P.; Rod, T. H.; Andersson, K. K.; Ryde, U. *Biophys. J.* **2004**, *87*, 3437–3447.
(77) Nilsson, K.; Ryde, U. *J. Inorg. Biochem.* **2004**, *98*, 1539–1546.
(78) Ryde, U.; Nilsson, K. *J. Am. Chem. Soc.* **2003**, *125*, 14232–14233.
(79) Raha, K.; Peters, M. B.; Wang, B.; Yu, N.; WollaCott, A. M.; Westerhoff, L. M.; Merz, K. M. *Drug Discovery Today* **2007**, *12*, 725–731.
(80) Cui, G. L.; Li, X.; Merz, K. M. *Biochemistry* **2007**, *46*, 1303–1311.
(81) Yu, N.; Li, X.; Cui, G. L.; Hayik, S. A.; Merz, K. M. *Protein Sci.* **2006**, *15*, 2773–2784.
(82) Yu, N.; Hayik, S. A.; Wang, B.; Liao, N.; Reynolds, C. H.; Merz, K. M. *J. Chem. Theory Comput.* **2006**, *2*, 1057–1069.
(83) Schumacher, M. A.; Miller, M. C.; Grkovic, S.; Brown, M. H.; Skurray, R. A.; Brennan, R. G. *Science* **2001**, *294*, 2158–2163.
(84) Grkovic, S.; Brown, M. H.; Schumacher, M. A.; Brennan, R. G.; Skurray, R. A. *J. Bacteriol.* **2001**, *183*, 7102–7109.

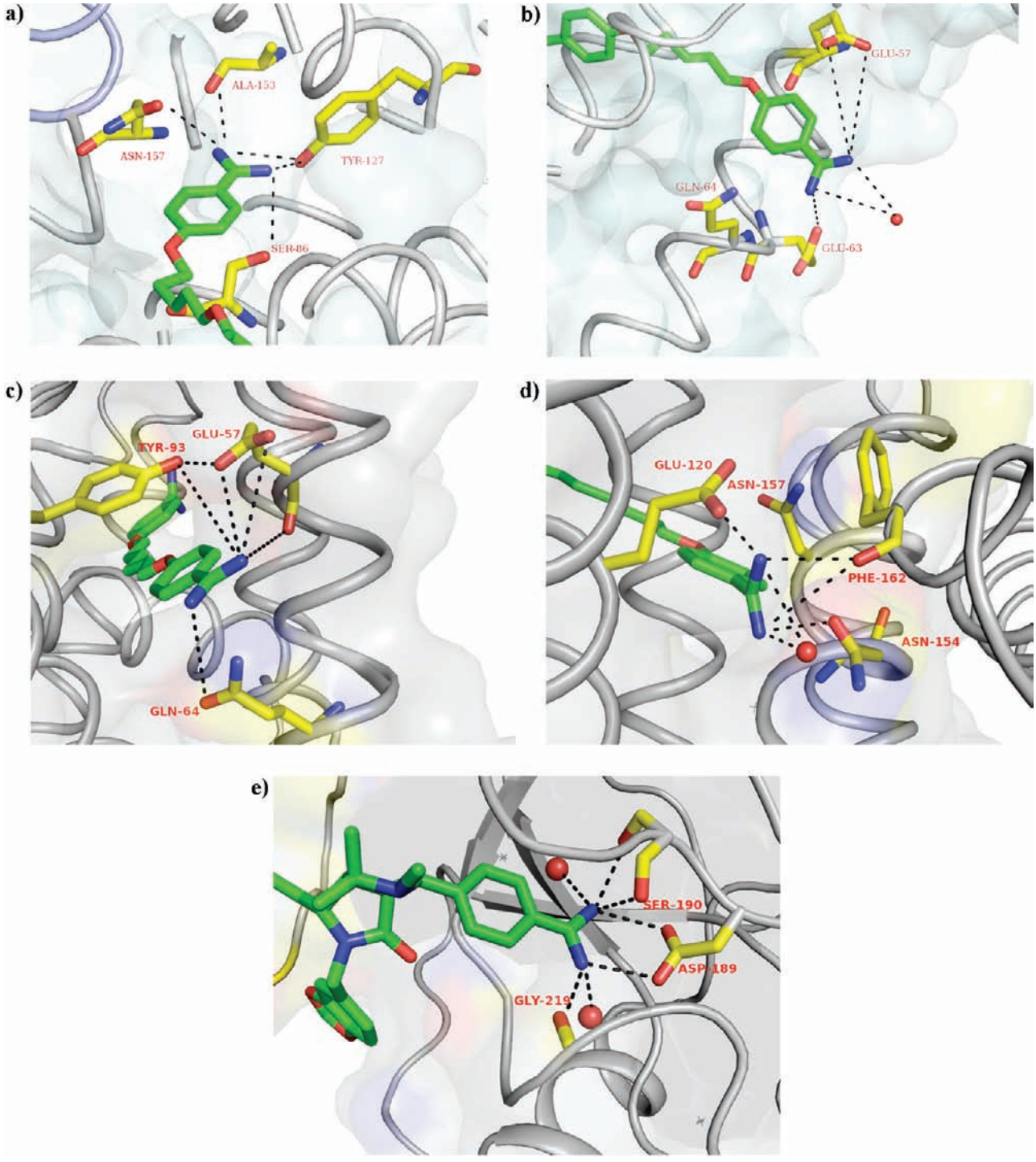

Figure 5. Select protein−ligand contacts: (a) PNT binding site 1, (b) PNT binding site 2, (c) DID binding site R6G, (d) DID binding site Et, and (e) UIB binding site. The residues in contact with the ligand in the active site are shown as yellow sticks, while the ligands are shown as green sticks. The oxygen atoms in water molecules are given as red spheres. All other residues are gray. Black dashed lines indicate hydrogen bond interactions. O and N atoms are colored red and blue, respectively. All structures are taken from the PDB.

to-end length of PNT is 16.03 Å, which is shorter than the value obtained after MM refinement (17.55 Å). The distances from N1 of PNT and Ser86/OG to N2 of PNT and Asn157/OD1 are shorter than those obtained from the MM refined structure. The nitrogen atom of Asn157 moved away from the N2 atom of PNT by 0.9 Å, thereby reducing repulsive interactions between the two nitrogen atoms after QM/MM refinement. In binding site 2, the two nitrogen atoms from the amidinium group of PNT after QM/MM refinement are much closer to the negatively charged carboxylate oxygen atoms from Glu63 and Glu57, which increases favorable electrostatic interactions between the ligand and protein. The water molecule stays at nearly the same location in both refinements.

The QacR−hexadiamidine complex (DID, shown in Figure 5c,d) was refined at 2.9 Å resolution and shows a binding mode very different from that of PNT. The residues from the R6G (Figure 5c) pocket include Glu57, Gln64, and Tyr93. Phe162′ (Figure 5d) and the carboxylate of Glu120 interact with the

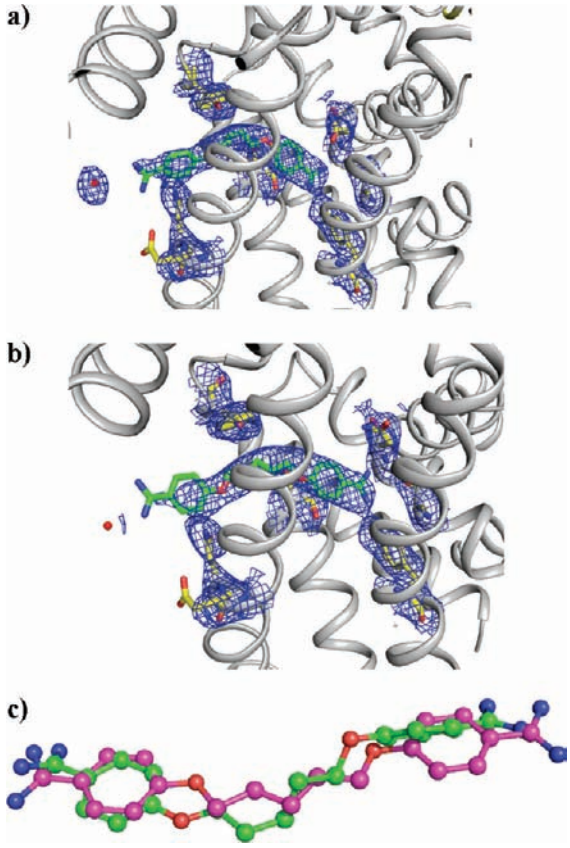

Figure 6. View of the $2F_o - F_c$ difference electron density of PNT contoured at 1.0σ at $w_a = 0.8$. The difference density is shown as a blue mesh, and the underlying structure is shown as a stick model: (a) QM/MM refined structure, (b) MM refined structure with Engh−Huber parameters, and (c) overlay of the QM/MM refined ligand (green) and the CNS model (magenta).

amidinium group within the Et binding site. In the Et pocket Asn154 contributes polar contacts between QacR and DID. Asn157 also has nonpolar contacts with the phenyl ring. A water molecule was also found in the Et binding pocket that was involved in water-mediated protein−small molecule interactions. The observed torsion angle for the benzamidinium fragment is −66°, indicating a very different conformational preference for this torsion angle in the Et binding site, while the same torsion angle is 0° in the R6G binding site (see Figure 7). These divergent conformational preferences for the two torsion angles were still observed in the structure after QM/MM refinement starting from the original crystal structure. For the selected structure at $w_a = 0.8$ after QM/MM refinement, the torsion angle preference is about −4° at the R6G binding site and −80° at the Et binding site. Compared to the structures obtained from the PDB and MM refinement, the R6G binding-site amidinium group is closer to the active-site residues, indicating stronger interaction between the ligand and protein in the binding site. In contrast, in the Et binding site, the distances between atoms from the ligand and protein residues are in between the distance observed from the PDB and re-refined MM structure. The planar conformation obtained from the MM re-refinement weakens the interactions between the ligand and protein active site.

The 1Y3X crystal structure was originally refined to 1.7 Å resolution. Each amidinium nitrogen atom from the UIB ligand (Figure 5e) makes hydrogen bonds to carboxylate oxygens of Asp189. One of the nitrogen atoms forms a hydrogen bond with the backbone carbonyl of Gly219, while the second nitrogen

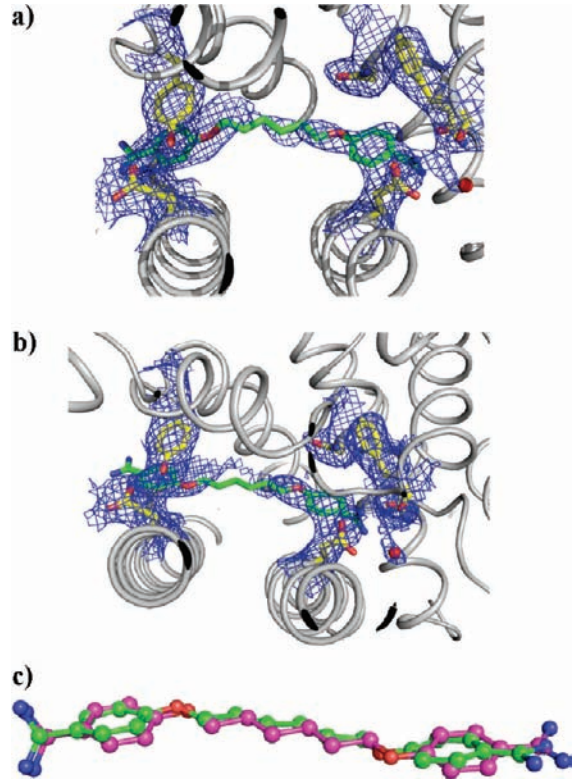

Figure 7. View of the $2F_o - F_c$ difference electron density of DID contoured at 0.5σ at $w_a = 0.8$. The difference density is shown as a blue mesh, and the underlying structure is shown as a stick model: (a) QM/MM refined structure; (b) MM refined structure with Engh−Huber parameters; and (c) overlay of the QM/MM refined ligand (green) and the CNS model (magenta).

forms a hydrogen bond with the hydroxyl oxygen and backbone carbonyl oxygen atom from Ser190. A water molecule also makes a hydrogen bond to this nitrogen atom and completes the hydrogen bond network between the benzamidinium fragment and the protein. After QM/MM refinement (see Figure 8a), the torsion angle is around 25°, which is different from the structures refined (Figure 8b) with MM (torsion angle is about 0°). No significant changes were observed for the corresponding distances between the benzamidinium fragment and protein residues after both refinements.

The QM/MM refinements on the three selected protein−benzamidinium derivative complexes show distinctly different preferences for the benzamidinium torsion relative to the deposited structures as well as re-refined MM structures. When comparing the structure quality of the refinements carried out using different methods, such as the electron density maps for the ligand (shown in Figures 6−8) and the R and R_{free} values (listed in Table 1), it is clear that the ligand structure after QM/MM refinement better fits the observed electron density than the one refined using a MM potential. Even though the R and R_{free} factors are indicators of global structure quality, we still observe that the QM/MM refined structures have better R and R_{free} values. The local structure for the small-molecule ligands has also been greatly improved by QM/MM-based refinement methods. The real-space R-values for the small-molecule ligands (listed in Table 1) are lower than those obtained after MM refinement. Taken together, this suggests that QM/MM-based refinement is a very useful tool for improving the local structure quality for small-molecule ligands.

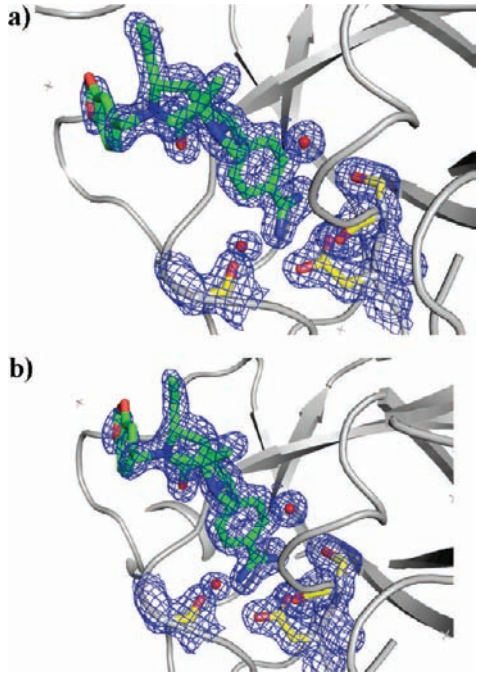

Figure 8. View of the $2F_o - F_c$ difference electron density of UIB contoured at 1.0σ at $w_a = 0.8$. The difference density is shown as a blue mesh, and the underlying structure is shown as a stick model: (a) QM/MM refined structure; (b) MM refined structure with Engh−Huber parameters.

Table 1. X-ray Refinement Parameters for the MM- and QM/MM-Based Refinement Studies

ligand	PDB ID	w_a	method	R	R_{free}	real-space R
PNT	1RKW	0.8	MM	0.237	0.243	0.252
			QM/MM	0.233	0.237	0.151
DID	1RPW	0.8	MM	0.254	0.286	0.345
			QM/MM	0.246	0.259	0.302
UIB	1Y3X	0.4	MM	0.177	0.190	0.111
			QM/MM	0.184	0.203	0.085

Table 2. Selected Torsion Angles (°) Obtained from Different Refinement Procedures

ligand	methods	site 1		site 2	
PNT		N1−C9−C4−C3	N2−C9−C4−C5	N1*−C9*−C4*−C3*	N2*−C9*−C4*−C5*
	PDB	1.60	1.45	0.46	−0.02
	QM/MM	34.06	33.19	25.55	31.09
	CNS	0.23	0.11	0.11	−0.02
DID		Et		R6G	
	PDB	−66.16	−66.14	−0.30	0.15
	QM/MM	−79.79	−74.93	−4.34	−0.87
	CNS	−71.34	−69.46	−0.02	0.00
UIB	PDB	−2.11	−1.46		
	QM/MM	−25.65	−24.67		
	CNS	−0.38	−0.34		

3. Computed Torsion Profiles. The refinement studies summarized above yielded contradictory results for the preferred conformation of benzamidinium derivatives when bound to a receptor. Table 2 summarizes the observed torsion values from the PDB for CNS MM and QM/MM refinement. A quick review of this table indicates that, depending on the refinement protocol, different results can be obtained. For example, for the PNT ligand the PDB and CNS MM refined structure gives nearly planar torsion angles, while the QM/MM model prefers rotated conformations of ∼30°. From small-molecule X-ray studies on benzamidinium and related derivatives (noted above), the value for this torsion was found to be on average ∼33° in the crystalline phase. In order to better understand the conformational preferences, we first calculated the torsion profile for the benzamidinium and *p*-hydroxybenzamidinium ions using *ab initio* QM calculations. In order to include the active-site context, we also carried out a series of QM calculations on the ligand with residues that are in close contact using *ab initio* methods as well.

The results from our *ab initio* gas-phase calculations, using different levels of theory, are shown in Figure 9. All the potential energy profiles for rotation of this dihedral angle showed a minimum at +40° and identically at 140° (or −40°) due to symmetry, which agrees with other theoretical studies.[42] The energy difference between our best level of theory (MP2/Aug-ccpvtz) and HF/6-31G* at 90° is ∼0.5 kcal/mol, and at the planar conformation (0 or 180°) the difference is ∼1.5 kcal/mol. Increasing the basis set size stabilizes the planar conformation, but even so the barrier to rotation is ∼2.5 kcal/mol, while that for the orthogonal structure is slightly higher at ∼3.3 kcal/mol. Adding a hydroxyl substituent increases the orthogonal barrier height by a small amount (∼0.2 kcal/mol for the *meta* derivative) to 1 kcal/mol for the *para* derivative. The effect on the planar conformations is generally small except for the *ortho* derivative, where intramolecular hydrogen-bonding interactions alter the torsion profile significantly (see Figure 9b). Nonetheless, for the *para* derivatives refined herein, it is clear that the gas-phase energy minimum is ∼30°−40°.

From the gas-phase torsion profiles it is clear that a twisted conformation is preferred for the parent compound, but benzamidinium-based ligands found in the PDB show various conformational preferences when bound to a protein receptor (see Figure 3). In the majority of cases, the flexible ligand binds at a higher energy conformation than found from our gas-phase computations or in small-molecule crystal structures. However, due to limited X-ray resolution and force field parameter issues for unusual ligands encountered in protein crystallography, it is difficult to know whether the predicted ligand conformation obtained from X-ray refinement is correct or is biased. In order to delve into this further, we have carried out *in situ* conformational analyses based on the fragment regions shown in Figure 2. In these calculations, we only rotated the torsion and then computed the energy. Hence, the resultant energies are high because we did not energy-minimize the system, but even these qualitative calculations provided useful insights.

From our cluster QM calculations we find that the QM/MM refinement structures always correspond to the global energy minimum for the ligand in its bound state (see Figure 10). From Figure 10a,b it is clear that, for the PNT ligand, both sites 1 and 2 prefer the twisted conformation. The situation for the DID ligand is the most interesting case in that both terminal benzamidinium groups adopt high-energy conformations. The R6G sites are nearly planar (see Figure 10c and Table 2), which from our gas-phase computations is 2−3 kcal/mol higher in energy that the slightly twisted conformation. Interestingly, the Et binding site of the molecule twists to ∼80°, which is ∼3−4 kcal/mol higher than the preferred twisted conformation (see Figure 9b) seen in our *ab initio* calculations. Hence, in contrast to the PNT ligand, we conclude that the DID ligand binds to the active site in a significantly strained conformation.

For the Et binding site, hydrogen-bonding interactions between Glu120, the backbone carbonyl oxygen atom from Phe162′, and the side-chain carbonyl oxygen of Asn154 are sufficient to distort the preferred conformation of the benzami-

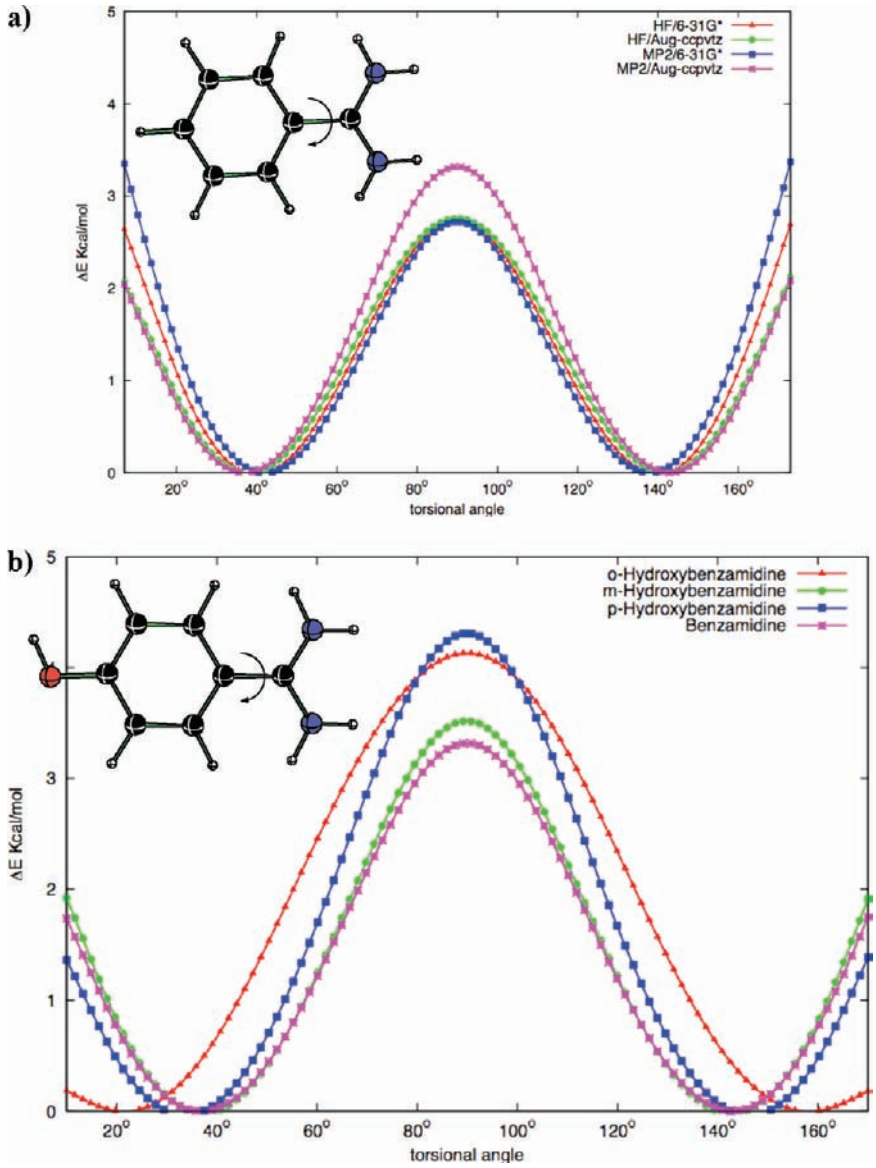

Figure 9. Torsion profiles obtained from different levels of theory for the benzamidinium (a) and hydroxybenzamidinium (b) ions.

dinium group. In the R6G binding site of DID, the amidinium group has several close contacts with the active-site residues. Unfavorable interactions with the side-chain amide group of Gln 64, coupled with favorable interactions with the carboxylate ion of Glu57, force the benzamidinium moiety to adopt the planar conformation as the global energy minimum.

For UIB we observe the expected behavior in that the benzamidinium group adopts the twisted conformation in the QM/MM-based refinement. Like the PNT case, the CNS MM and PDB structures adopt the incorrect planar conformation.

5. Torsion Angle Distributions Obtained from MD Simulations. Using our *ab initio* computed torsion profile, we built an AMBER force field model in order to carry out MD simulations on the PNT and DID complexes to get a sense of how flexible the binding sites might be and to further corroborate our QM/MM refinement results. Previous MD simulations of bovine trypsin complexed with benzamidinium, using CHARMM22 force field[85] and TIP3P water model,[58] indicated that the preferred torsion value is ±25° and that the planar conformation obtained from crystallographic refinement is an average of two symmetric, nonplanar conformations.[40]

Figures 11 and 12 show the benzamidinium dihedral angle distributions for both binding sites in the QacR−pentadiamidine (PNT) and QacR−hexadiamidine (DID) complexes, respectively, during the MD simulations. The N−C−C−C torsion angle in binding site 1 of PNT is maintained at around 30°, whereas the N−C−C−C torsion in site 2 is maintained at 40° throughout the entire 8 ns production simulation. These conformations are significantly different from the planar structures reported from the PDB X-ray analyses, but they are similar to our QM/MM refined structure. In our MD simulation on DID, the torsion angle value in the R6G site is on average ∼10°, indicating the benzamidinium moiety is nearly planar. The benzamidinium in the Et site is twisted with a dihedral angle of −40°. This observation matches the X-ray structure and our QM/MM refined structures in the sense that the twisted conformation is preferred, but the MD simulation has allowed the active site to relax for the much more twisted conformation

(85) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

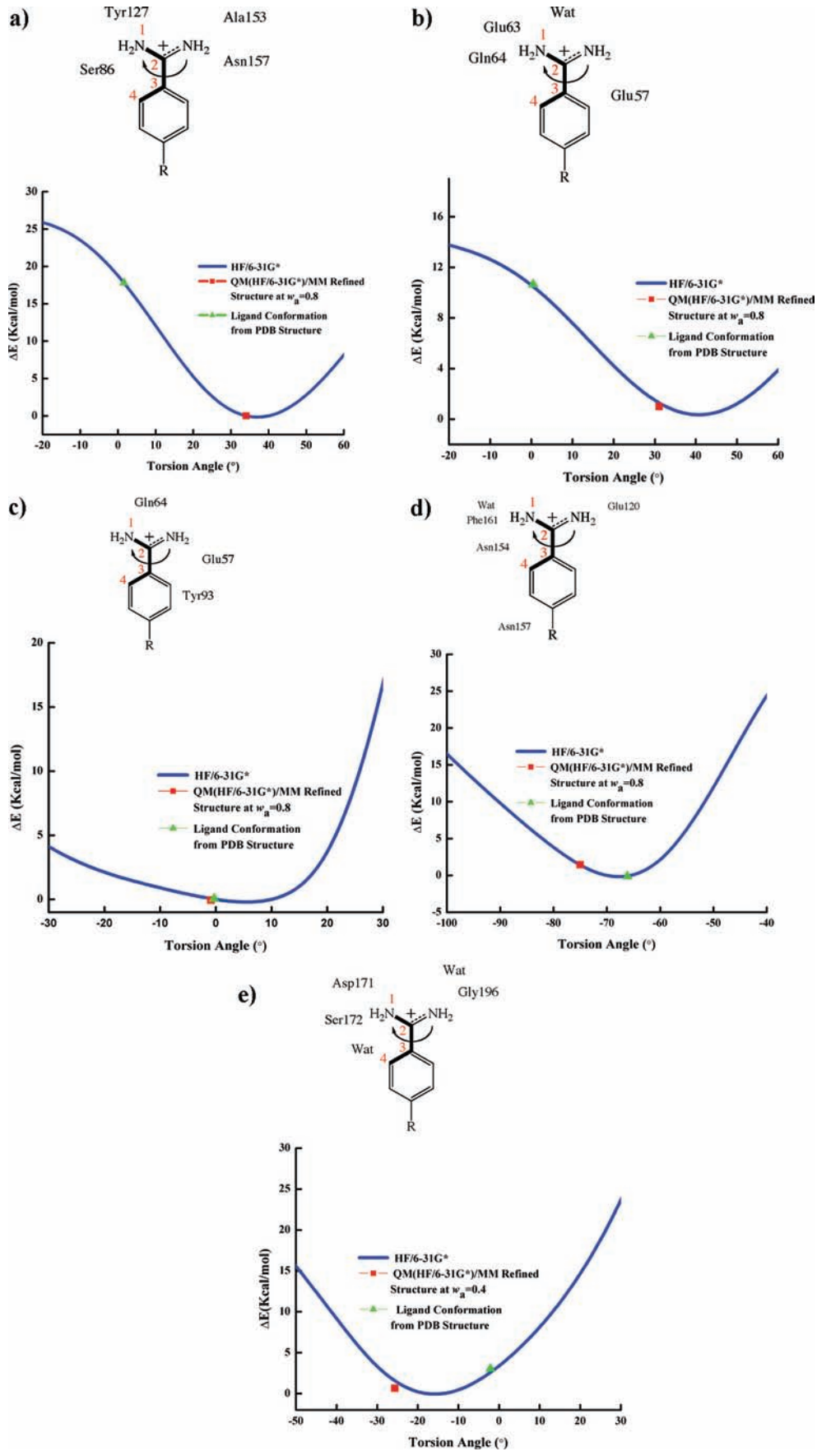

Figure 10. Cluster torsion energy profiles for the bound ligands: (a) PNT binding site 1, (b) PNT binding site 2, (c) DID binding site R6G, (d) DID binding site Et, and (e) UIB binding site.

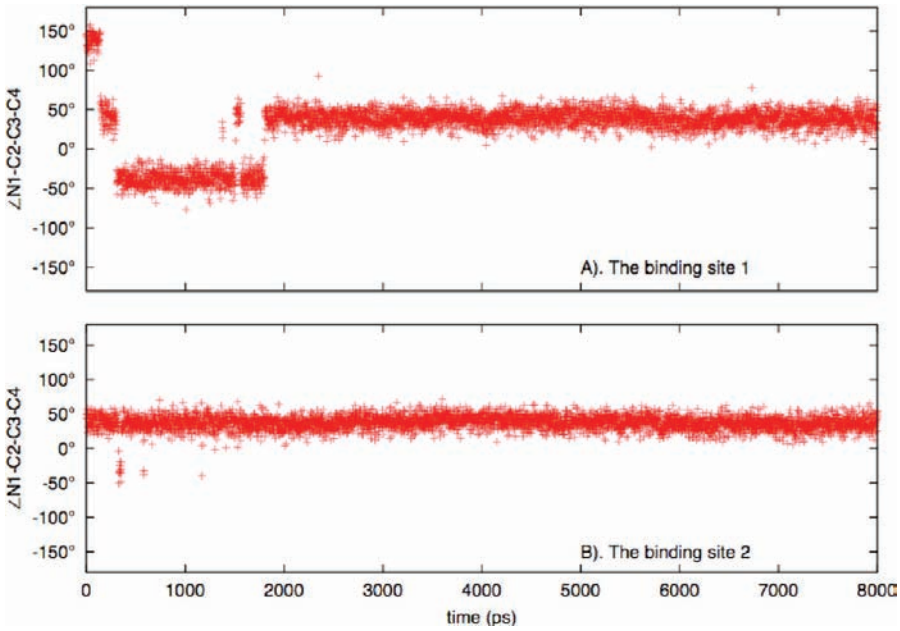

Figure 11. Dihedral angle distribution of the benzamidinium fragment in the QacR−pentadiamidine complex taken from an 8 ns MD simulation.

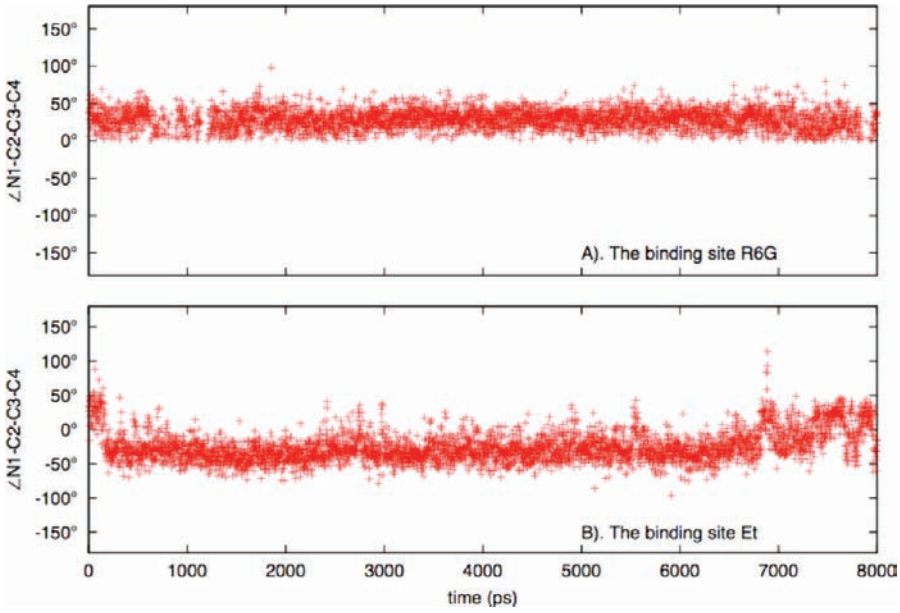

Figure 12. Dihedral angle distribution of the benzamidinium fragment in the QacR−hexadiamidine complex taken from an 8 ns MD simulation.

predicted from the QM/MM refinement study. Nonetheless, the MD results confirm the results from the QM/MM refinement studies. Moreover, they do not suggest that the observed planar structures arise due to the sampling of two symmetrically related structures as was observed in earlier MD simulations.

Conclusions

In this article we have shown that QM/MM-based X-ray refinement can improve the analysis of protein−ligand complexes both by improving R and R_{free} values and by providing a more accurate representation of the intra- and intermolecular interactions of the ligand with itself and its environment. The reason this approach succeeds depends on the resolution of the primary X-ray data and the ability of standard MM potentials within refinement programs to handle a specific pharmacophore. At high resolution defects in a given MM potential can be overwhelmed by the X-ray signal, while in lower resolution cases defects in a MM model can cause problems.

For benzamidinium and benzamidinium derivatives examined in detail herein, the protein−ligand complexes obtained from the PDB were found to have a strong preference for the planar conformation (see Figure 3). However, our QM/MM refinement studies of benzamidinium derivatives suggest that the ligand conformation presented in the PDB may not correspond to the "preferred" energy minimum and that this is at least partially due to the choice of MM potential used during the course of refinement. We find that QM/MM refinement of these complexes locally improved the ligand structure and tended to yield refined structures with the preferred (low-energy) nonplanar conformation. Importantly, in cases where the active-site environment imposes constraints on the "preferred" conformation, the QM/MM approach is able to adapt to these perturba-

tions. Using our gas-phase QM calculations, we built improved MM potentials to carry out MD simulation studies of QacR complexed with the two benzamidinium derivatives studied herein. The MD simulations also confirmed the preferred conformations for these inhibitors obtained by QM/MM-based X-ray refinement.

Our observations on benzamidinium are important on another, more qualitative level. In SBDD studies, refined X-ray structures are relied on to suggest ways in which to modify or build new inhibitor structures. This effort depends critically on accurate protein−ligand structures in order to guide the design process. The PNT case (QacR−pentadiamidine complex) illustrates clearly what can go awry. The MM refined and PDB structures (see Figure 5a,b) both predict that the two benzamidiniums are planar, while the QM/MM refined structures (see Figures 6) are both twisted out of plane (sites 1 and 2). Indeed, these small differences in the computed structures might lead to different conclusions about what molecules to make next. Improved force fields can allay this issue, but given the creativity of modern organic chemistry with respect to the molecular scaffolds one can create, this is an ongoing and challenging problem from which QM-based methods simply do not particularly suffer. Thus, it is safe to conclude that the QM/MM approach is significantly more general than any MM approach.

In this work we have focused on one class of small molecules (benzamidinium derivatives), but there are numerous other examples of classes of drug-like molecules that can be examined. Given the large number of crystal structures of pharmaceutically relevant protein−ligand complexes that are available, a comprehensive study of systems with available structure factors using the QM/MM refinement approach could yield interesting insights into conformational preferences. We are in the process of investigating several other classes of molecules along the lines described herein.

Acknowledgment. The authors thank Dr. Shaoliang Zheng at SUNY-Buffalo for valuable discussions on the torsion motif preferences of small crystal structures from CSD. We thank the NIH for supporting this research via grants GM044974 and GM066859.

Supporting Information Available: Complete refs 23, 48, 54, 63, and 85. This material is available free of charge via the Internet at http://pubs.acs.org.

JA9010833