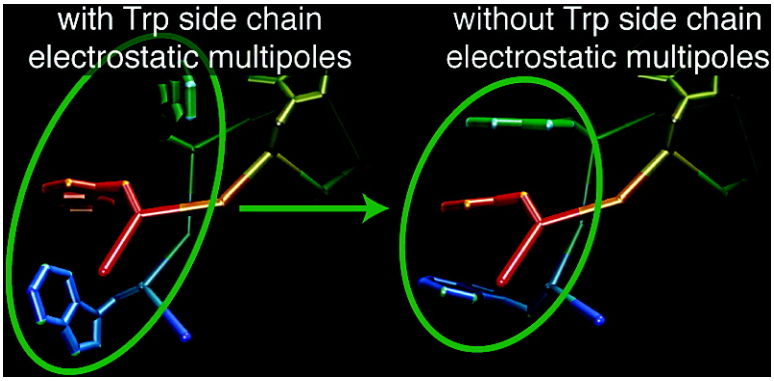
Article

Tryptophan Side Chain Electrostatic Interactions Determine Edge-to-Face vs Parallel-Displaced Tryptophan Side Chain Geometries in the Designed β-Hairpin "trpzip2"

Olgun Guvench, and Charles L. Brooks

J. Am. Chem. Soc., **2005**, 127 (13), 4668-4674• DOI: 10.1021/ja043492e • Publication Date (Web): 12 March 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 10 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 03/12/2005

Tryptophan Side Chain Electrostatic Interactions Determine Edge-to-Face vs Parallel-Displaced Tryptophan Side Chain Geometries in the Designed β-Hairpin "trpzip2"

Olgun Guvench and Charles L. Brooks III*

Contribution from the Department of Molecular Biology (TPC-6), The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received October 27, 2004; E-mail: brooks@scripps.edu

Abstract: The interaction geometries of the four tryptophan (Trp) side chains in the 12-residue designed β-hairpin trpzip2 are investigated using all-atom explicit-solvent molecular dynamics simulations. The experimentally observed edge-to-face (EtF) pairwise interaction geometries are stable on a time scale of 10 ns. However, removing the electrostatic multipoles of the Trp side chains while retaining the dipoles of the side chains' NH moieties induces a conformational change to a geometry in which three of the four side chains interact in a parallel-displaced (PD) manner. Free energy simulations of the Etf to PD conformational change reveal that, with the side chain multipole moments intact (+MP), the EtF conformation is preferred by 5.79 kcal/mol. Conversely, with only the dipole moments of the side chain NH moieties intact (−MP), the PD conformation's free energy is more favorable by 1.71 kcal/mol. In contrast to energetic similarities for Trp side chain−water electrostatic and Trp side chain−Trp side chain and Trp side chain−water van der Waals, +MP Trp side chain−Trp side chain electrostatic interactions are more favorable by 4.21 kcal/mol in the EtF conformation, while in the −MP case the EtF and PD conformations' Trp side chain−Trp side chain electrostatic energies are nearly identical. The results highlight the importance of electrostatic multipole moments in determining aromatic−aromatic interaction geometries in aqueous biomolecular systems and argue for the inclusion of this physics in simplified models used for protein−ligand docking and protein structure prediction, possibly through a truncated Coulomb term between aromatic moieties.

Introduction

The interaction between a pair of aromatic moieties is widely recognized to be important in biomolecular systems.[1,2] This interaction, though weak relative to hydrogen bonding,[3,4] has been found to be quite prevalent in structural studies of proteins[5] and protein−ligand complexes.[2] Extensive experimental investigations of model systems have shown that the aromatic−aromatic interaction is both favorable and specific, not only in organic solvents but also in water, and help to provide a rationale for the observations in structural studies.[6]

The benzene dimer has served as the canonical model system for studying aromatic−aromatic interactions. Early work employing molecular beam spectroscopy implied the favored gas-phase conformation of benzene to be an edge-to-face dimer (Figure 1a).[7] Subsequent quantum chemical calculations revealed that, in addition to the edge-to-face geometry, the parallel-displaced geometry (Figure 1b) is also a favorable conformation.[8] Recent high level ab initio quantum chemical calculations on this system provide accurate energetic insight into the competition between edge-to-face and parallel-displaced interaction geometries: in gas phase, the two conformations are essentially isoenergetic ($\Delta E = 0.02$ kcal/mol)[4] and with a nearly flat energy profile for interconversion (barrier = 0.2 kcal/mol).[9]

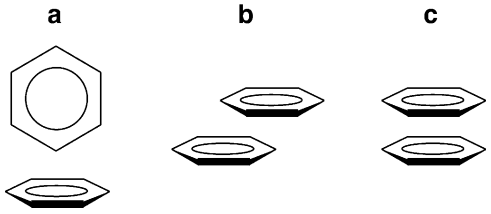


Figure 1. Benzene dimer interaction geometries: (a) edge-to-face, (b) parallel-displaced, (c) face-to-face.

A survey of pairs of aromatic side chains in high-resolution X-ray crystal structures found that there is a full continuum of values for the interplanar angles formed by the aromatic pairs that ranges from coplanar to perpendicular.[10] Additionally, this survey finds only a 0.5−0.75 kcal/mol preference for parallel-

(1) Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. *J. Chem. Soc., Perkin Trans. 2* **2001**, 651−669.
(2) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed.* **2003**, *42*, 1210−1250.
(3) Mas, E. M.; Szalewicz, K. *J. Chem. Phys.* **1996**, *104*, 7606−7614.
(4) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2002**, *124*, 104−112.
(5) Burley, S. K.; Petsko, G. A. *Science* **1985**, *229*, 23−28.
(6) Waters, M. L. *Curr. Opin. Chem. Biol.* **2002**, *6*, 736−741.
(7) Janda, K. C.; Hemminger, J. C.; Winn, J. S.; Novick, S. E.; Harris, S. J.; Klemperer, W. *J. Chem. Phys.* **1975**, *63*, 1419−1421.
(8) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Chem. Phys.* **1990**, *93*, 5893−5897.
(9) Tsuzuki, S.; Uchimaru, T.; Sugawara, K.; Mikami, M. *J. Chem. Phys.* **2002**, *117*, 11216−11221.

10.1021/ja043492e CCC: $30.25

displaced Phe−Phe side chain pair geometries using 300 K as the temperature value for converting database statistics to free energies; this energy range would be downward revised by another 67% using a temperature of 100 K, a common temperature used in high-resolution X-ray crystal structure determination. There thus appears to be a good correlation between benzene dimer gas phase conformational energetics and the observed geometries of Phe−Phe side chain pairs.

The favorable interaction geometries for the benzene dimer represent a compromise between maximizing dispersion and electrostatic interactions.[4] If the dimerization were governed purely by dispersion interactions, the conformation with a maximum of intermolecular contact would be an energetic minimum. However, the excess of negative charge on the electron-rich carbon atoms and the corresponding positive charge character of the electron-deficient hydrogens preclude such a direct face-to-face stacked conformation (Figure 1c) because of the intermolecular electrostatic repulsion. In contrast, both the parallel-displaced and the edge-to-face conformations allow favorable interaction between the benzene quadrupole moments. This physics of compromise between dispersion and electrostatic interactions extends to more complex aromatic systems, for example in the stacking of bases in DNA and RNA.[11]

Though exceedingly accurate, gas-phase quantum chemical studies of aromatic−aromatic interactions lack solvent effects, whereas aromatic−aromatic interactions in biomolecules occur in either an aqueous environment if solvent-exposed or an apolar solventlike environment if, for example, buried in the hydrophobic core or binding site of a protein. Early solvated molecular mechanics studies employing Coulomb and Lennard−Jones interaction sites on the carbon and the hydrogen positions showed that, in aqueous solvent, benzene dimer interaction energies are attenuated relative to gas phase.[12−14] More recent solvated molecular mechanics studies have gone beyond the benzene dimer to include amino acid β and α atoms as well as the hydroxyl moiety characteristic of tyrosine and have found that parallel-displaced conformations are free energy minima for Phe−Phe, Tyr−Tyr, and Phe−Tyr side chain pairs and are ∼1 kcal/mol lower in free energy than the edge-to-face conformation, as opposed to the preferred edge-to-face geometry of the solvated benzene dimer.[15,16] Interestingly, a pair of solvent-exposed cross-strand Phe residues in a designed β-hairpin peptide were found to assume an edge-to-face geometry in aqueous solvent,[17] in contrast to what might be inferred from these latter simulation studies.

Despite the aforementioned computational studies and the widely accepted use of all-atom models in explicit-solvent molecular mechanics simulations of biomolecular systems,[18] there remains some debate about the importance of electrostatic interactions relative to dispersion and hydrophobic effects in determining aromatic−aromatic interaction geometries,[6] as well as a lack of simulation studies on actual biomolecular systems of known structure exploring the importance of the multipole moments of the aromatic amino acid side chains. Perhaps reflecting this, many simplified models used in protein−ligand docking and protein structure prediction, while accounting for bonded interactions, volume exclusion, solvent effects, and hydrogen bonding, lack terms which can be expected to capture the specific, geometric nature of aromatic−aromatic interactions.

To extend the simulation studies of aromatic−aromatic interactions to Trp side chain pairs and also to do so in the context of a real, experimentally characterized system, we have undertaken explicit-water molecular dynamics simulations of the designed 12-residue β-hairpin "trpzip2."[19,20] This so-called "tryptophan zipper" has two pairs of cross-strand Trp side chain−Trp side chain interactions, both of which assume edge-to-face geometries, and is 1.3 kcal/mol stable in water at 298 K. Owing to its small size and unusual stability, trpzip2 is an attractive system for computational studies, and prior simulations using implicit solvent models have been used to characterize its folding thermodynamics and kinetics and also for validation of peptide backbone force field parameters.[21−23]

In the present work, simulations with and without multipoles on the Trp side chains were run so as to separate electrostatic multipole effects from dispersion and hydrophobic forces in determining the Trp side chain interaction geometries. In both cases the large electrostatic dipoles on the Trp side chain NH moieties were maintained so as not to perturb canonical hydrogen bonding involving the side chains. Lack of the electrostatic multipoles, which normally arise from the net neutral collection of partial charges on the Trp side chain aromatic carbon and hydrogen atoms, induced a conformational change from the experimental edge-to-face geometry to parallel-displaced. Free energy calculations were used to quantitate the relative preferences for the two conformations with and without multipoles. Energetic analysis revealed that electrostatic multipole interactions between the Trp side chains are a major contributor to the preferred edge-to-face side chain geometries despite the solvent-exposed nature of all four Trp side chains, suggesting that a term based on Coulomb's law may be sufficient to account for aromatic−aromatic interactions in simplified models employed in protein−ligand docking and protein structure prediction.

Results and Discussion

The seventh of twenty structures in the trpzip2 NMR ensemble (PDB code 1le1) was randomly selected to serve as the starting point for the all-atom explicit-water molecular dynamics (MD) simulations. The peptide was solvated in a truncated octahedron of water molecules such that a minimum distance of 14 Å separated the peptide from the edges of the truncated octahedron, resulting in a system with a minimum dimension of 50 Å and composed of the peptide plus 3183 water molecules.

(10) McGaughey, G. B.; Gagne, M.; Rappe, A. K. *J. Biol. Chem.* **1998**, *273*, 15458−15463.
(11) Sponer, J.; Leszczynski, J.; Hobza, P. *Biopolymers* **2001**, *61*, 3−31.
(12) Jorgensen, W. L.; Severance, D. L. *J. Am. Chem. Soc.* **1990**, *112*, 4768−4774.
(13) Linse, P. *J. Am. Chem. Soc.* **1992**, *114*, 4366−4373.
(14) Linse, P. *J. Am. Chem. Soc.* **1993**, *115*, 8793−8797.
(15) Chipot, C.; Jaffe, R.; Maigret, B.; Pearlman, D. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 11217−11224.
(16) Chelli, R.; Gervasio, F. L.; Procacci, P.; Schettino, V. *J. Am. Chem. Soc.* **2002**, *124*, 6133−6143.
(17) Tatko, C. D.; Waters, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 9372−9373.
(18) Ponder, J. W.; Case, D. A. *Protein Simulations*; Academic Press: San Diego, 2003; Vol. 66, pp 27−85.
(19) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5578−5583.
(20) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9081.
(21) Yang, W. Y.; Pitera, J. W.; Swope, W. C.; Gruebele, M. *J. Mol. Biol.* **2004**, *336*, 241−251.
(22) Snow, C. D.; Qiu, L. L.; Du, D. G.; Gai, F.; Hagen, S. J.; Pande, V. S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 4077−4082.
(23) Okur, A.; Strockbine, B.; Hornak, V.; Simmerling, C. *J. Comput. Chem.* **2003**, *24*, 21−31.

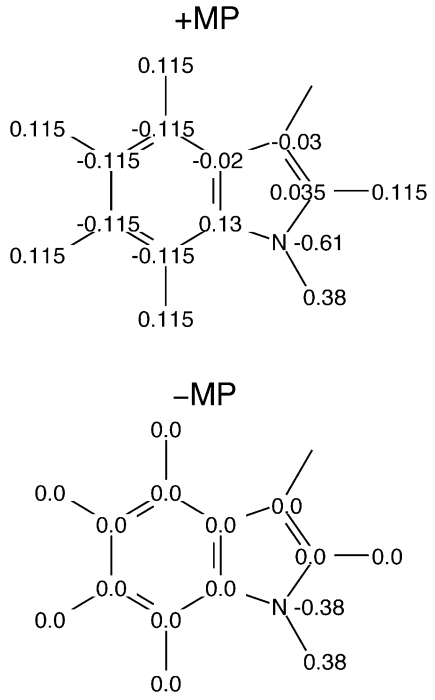


Figure 2. Trp side chain partial charges. +MP: with multipole intact. −MP: without multipole.

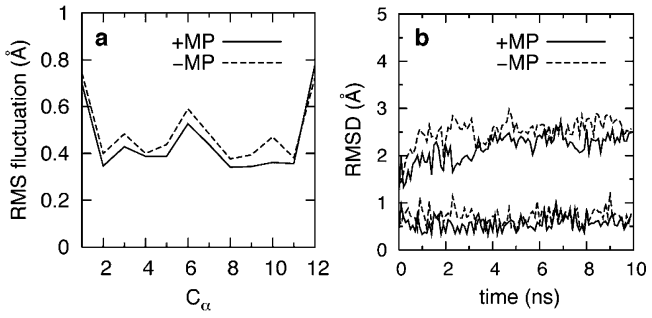


Figure 3. (a) C_α atoms' RMS fluctuations about average positions. (b) RMSD with respect to starting structure vs time. The two time series with larger values are all-atom RMSDs, and the two with smaller values are C_α RMSDs. All-atom RMSDs are calculated after least-squares alignment of all atoms; C_α RMSDs are calculated after least-squares alignment of C_α atoms.

To evaluate the effect of the aromatic Trp side chains' multipoles on trpzip2 structure and dynamics, MD simulations were performed with the Trp side chains' aromatic electrostatic multipoles intact (+MP) and without these multipoles (−MP). Removal of the multipoles was accomplished by selectively setting to 0 atom-centered partial charges used in calculating the molecular mechanics electrostatic energy and forces (Figure 2). The Trp side chains' NH dipoles were maintained in the −MP simulation, and thus the side chains' abilities to form canonical hydrogen bonds were unaltered.

C_α atom fluctuations about average coordinates for the 10-ns +MP and −MP MD simulations are nearly identical (Figure 3a). Unremarkably, the largest fluctuations are seen in Asn6 and Gly7, which form the hairpin turn, and in the terminal residues. From the time courses of the C_α RMSDs relative to the NMR starting structure, the backbone geometry for both the +MP and −MP simulations remains constant and very close to the initial geometry for the duration (Figure 3b). On the 10-ns time scale, the presence or absence of Trp side chain multipoles has no effect on the secondary structure of this 12-residue β-hairpin, and in both cases the backbone geometry is very stable and rigid. The backbone stability and rigidity in the +MP simulation is expected, as there is little variation in the positions of the backbone atoms in the 20 models in the NMR ensemble. The similar result for the −MP simulation likely reflects the fact that the Trp side chains, despite their lack of multipoles, still protect cross-strand backbone hydrogen bonds from exposure to solvent simply by virtue of their steric bulk.

Table 1. Tryptophan Side Chains' Average χ_1/χ_2 Values[a]

	NMR	+MP	–MP
Trp2	−76/102	−63(9)/99(13)	−63(9)/90(15)
Trp4	178 /−99	−178(8) /−96(14)	−174(9) /−92(15)
Trp9	−67/104	−66(8)/105(11)	**−128(43) /50(50)**
Trp11	−175/−95	−170(11)/**0(88)**	**−110(54)/36(79)**

[a] RMS fluctuations about mean values are in parentheses. Values with RMS fluctuations > 20° are in bold.

The time courses of the +MP and −MP all-atom RMSDs show significantly greater deviation from the starting conformation than the C_α RMSDs (2.5 Å compared to 0.5 Å for C_α only) (Figure 3b), implying motion in the amino acid side chains, which are solvent exposed. This is in accord with the NMR ensemble whose 20 structures' non-Trp side chains show considerable variability in their coordinates.

Unlike the side chains of the other residues, the four Trp side chains in the 20 NMR models show a constancy of atomic coordinates, with conserved edge-to-face interaction geometries for both the Trp4−Trp9 and Trp2−Trp11 side chain pairs. Extending the analysis of MD fluctuations about average structures from C_α atoms to the Trp side chains reveals a marked difference between the +MP and −MP cases, in contrast to the similar simulation backbone dynamics and stability. The −MP simulation shows large fluctuations about the average values in both the χ_1 ($N-C_\alpha-C_\beta-C_\gamma$) and χ_2 ($C_\alpha-C_\beta-C_\gamma-C_{\delta^1}$) side chain dihedrals for Trp residues 9 and 11, while +MP has large fluctuations only in Trp11 χ_2. In both cases, the Trp dihedrals with small fluctuations also have average χ values that are close to the values from the NMR structure that serves as the initial condition for the MD simulations. Also in both cases, when a dihedral fluctuation is large, the deviation of the average value of that dihedral relative to the starting NMR structure is large (Table 1). This suggests that the calculated fluctuations arise from conformational changes and not simply from floppiness in the Trp side chains.

Tracking the relative geometries of the four Trp side chains as a function of time confirms the presence of Trp side chain conformational changes in the −MP simulation. Figure 4 illustrates the progression of the interplanar angles between the side chains of Trp4 and Trp9, Trp9 and Trp2, and Trp2 and Trp11, using the $C_{\epsilon^2}-N_{\epsilon^1}$ and $C_{\epsilon^2}-C_{\delta^2}$ bond vectors to define the side chain planes. The +MP data show that the Trp side chains retain edge-to-face interaction geometries for the full 10 ns. Conversely, 2 ns into the −MP simulation, the Trp4−Trp9 and Trp9−Trp2 interplanar angles change so that Trp9 and Trp2 become coplanar and Trp4 and Trp9 become nearly coplanar for a 0.5-ns interval. These residues revert to their starting relative geometries for a subsequent 1.5 ns, after which they once again assume coplanar geometries for the remaining 6 ns of the 10-ns simulation. The lack of Trp side chain multipoles in the −MP simulation also induces a conformational change

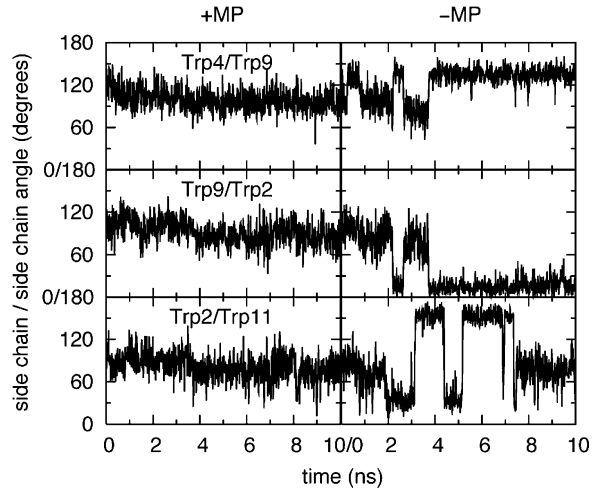


Figure 4. Trp side chain−Trp side chain interplanar angles vs time.

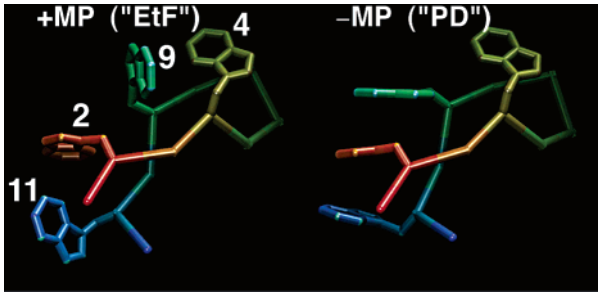


Figure 5. Average structures from +MP and −MP trajectories.

in the Trp2−Trp11 side chain interaction. During the 3-ns to 4-ns and the 5-ns to 7.5-ns intervals, this pair also assumes a coplanar geometry.

Figure 5 contrasts the coplanar parallel-displaced Trp side chain geometries induced by a lack of multipoles in the −MP simulation with the edge-to-face geometries in the +MP simulation. The +MP structure (denoted "EtF" for the edge-to-face Trp side chain interaction geometries) is an average of snapshots from the 0.5-ns to 3.5-ns interval, and the Trp side chain geometries of this average are equivalent to those of the NMR starting structure. By 0.5 ns into the +MP simulation, the all-atom RMSD has reached a plateau, while at 3.5 ns Trp11 undergoes a change in χ_2 such that it becomes more solvent exposed, which accounts for the large fluctuation in this dihedral (Table 1). Trp11 resumes its initial geometry for the 7-ns to 8-ns interval and assumes its alternate solvent-exposed geometry in the 8-ns to 10-ns interval. As apparent from Figure 4, the alternate solvent-exposed geometry retains the near-perpendicular angle between the Trp2 and Trp11 side chain planes, and the several interconversions experienced by Trp11 suggest it is in rapid equilibrium between the published NMR structure and a more solvent-exposed conformation. The mobility of the Trp11 χ_2 dihedral does not compromise the near-perpendicular Trp side chain interplanar angles, which are maintained for the entire 10 ns.

In the −MP simulation, though only the Trp9 and Trp11 χ_1 and χ_2 dihedrals have average values deviating from the NMR starting structure and large fluctuations about these average values, the conformational changes in Trp9 and Trp11 that are the basis for these average values and fluctuations have a large effect on the overall interaction geometries of all four Trp side chains. The −MP structure in Figure 5 (denoted "PD" for the parallel-displaced Trp side chain interaction geometries), an average of snapshots from the 5.5-ns to 6.5-ns interval, illustrates how changes in Trp9 and Trp11 side chain conformations lead to a set of Trp side chain interactions strikingly different from the NMR starting structure. The change in conformation of Trp9 leads to a loss of the edge-to-face Trp4−Trp9 interaction. This conformational change has the additional effect of creating a stacking interaction between Trp9 and Trp2. The Trp11 side chain motion leads to conversion of the Trp2−Trp11 edge-to-face interaction geometry to a parallel-displaced geometry. Thus, the net effect of the removal of the Trp side chain multipoles is a loss of two pairwise edge-to-face aromatic interactions and the creation of a stacked geometry involving three of the four Trp side chains.

The secondary structure is constant and the same in both the +MP and −MP simulations, and hence possible Trp side chain geometries are similarly constrained for both because of the fixed relative positions of the C_α and C_β atoms. Positional constraint from the backbone is therefore not the determining factor for Trp geometries in this system, unlike in a study of pairwise interactions of Phe and Tyr side chain analogues.[16] Rather, simple alteration of the χ_1 and χ_2 dihedrals for two of the four residues is possible and is sufficient to transform the experimentally observed edge-to-face interactions to a stacked geometry but is precluded by the Trp side chain multipoles. In contrast to the importance of the Trp side chains' multipole moments in preserving the edge-to-face geometry, the side chains' NH dipoles cannot by themselves, either through peptide−peptide or peptide−solvent electrostatic interactions, preserve the correct geometry. In both structures in Figure 5, solvent exposure of the Trp side chains' NH moieties is sufficient to satisfy their hydrogen bonding potential.

Further simulations were undertaken to estimate the free energy change associated with the EtF → PD conformational change for both the +MP and the −MP systems. To effect the conformational transformation, all Trp side chain atoms (including C_β and H_β atoms) were duplicated, one set of duplicates assigned to a set "X," the other assigned to a set "Y," and the rest of the system (nonduplicated atoms) assigned to a set "Z." The energies for X−Z and Y−Z interactions were scaled by 0.5, the energies for X−Y interactions were set to 0, and the energies for X−X and Y−Y interactions were unaltered. The NMR structure coordinates were assigned to nonduplicated peptide atoms, and after least-squares alignment of C_α atoms, set X coordinates were assigned from average structure EtF and set Y coordinates from average structure PD. The hybrid peptide was then energy minimized with and without Trp side chain multipoles, and the resultant X and Y coordinates defined the end points for the EtF → PD conformational changes for the +MP and −MP cases, respectively.

Intermediate hybrid conformations were constructed to span the endpoints of the EtF → PD conformational change. $\delta\chi$ values were calculated for each Trp's two χ dihedrals such that $\delta\chi = (\chi^Y - \chi^X)/20$. The X coordinates for hybrid conformation 1 were those of the EtF endpoint, and the Y coordinates were determined by constraining X atoms' as well as Trp C_α and N atoms' coordinates, restraining the Y set's χ values to $\chi^{EtF} + \delta\chi$ and minimizing the peptide. The X coordinates for conformation 2 were those of the Y coordinates for conformation 1,

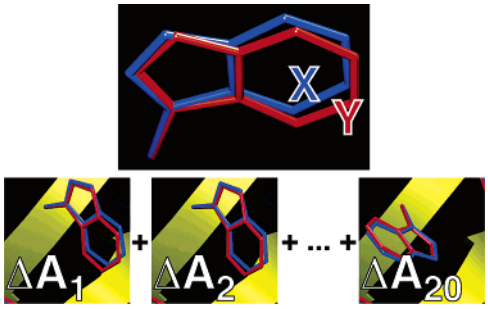


Figure 6. Hybrid Trp residue used in EtF → PD conformational change free energy simulations.

and the Y coordinates for conformation 2 were determined by restraining the Y set's χ values to $\chi^{EtF} + 2\delta\chi$, constraining the X set atoms as well as Trp C_α and N atoms, and minimizing the peptide. Intermediates 3 through 19 were similarly constructed, and the Y coordinates for intermediate conformation 20 were the PD endpoint coordinates. In the +MP case, because of a large distance between the X and Y coordinates for one Trp residue that resulted during minimization from an unfavorable electrostatic interaction between two proximal hydrogens, two conformations, 7a and 7b, were constructed to split the conformational space originally spanned by conformation 7.

For each of these intermediate hybrid conformations, an MD simulation was run with constraints on the X and Y sets' coordinates using the scaled interaction energy scheme involving sets X, Y, and Z described above. Subsequently, the free energy change ΔA_i associated with the X → Y conformational change for each hybrid conformation i was calculated using the thermodynamic integration formalism.[24] Because free energy is a state function, the total free energy change ΔA associated with transforming the EtF endpoint to the PD endpoint is then simply calculated as

$$\Delta A = \sum_{i=1}^{n} \Delta A_i$$

Figure 6 illustrates the hybrid Trp9 side chain from conformation 1 and the conformational change undergone by this side chain in going from conformation 1 to conformation 20. The very small difference in coordinates between the X and Y copies of the duplicated atoms in each conformation, along with the scaling of X−Z and Y−Z interactions, allows for a good sampling of atoms in set Z, especially those belonging to water molecules interacting with the Trp side chains. A possible alternate approach to calculating the free energy change of the EtF → PD conformational change would be to use umbrella sampling with restraining potentials applied to the Trp χ dihedrals. The drawback to this approach is that, with 8 biases (4 Trp residues × 2 restrained dihedral angles/residue), unbiasing would require eight-dimensional bins,[25] an impractical proposition given limits on computer memory and CPU time. We note that, with the present approach, the partition function of the system has been altered relative to the unconstrained peptide due to the use of positional constraints on the Trp side chain atoms; however, the peptide's rotational and translational contributions to the partition functions of the solvated EtF and PD conformations are expected to be nearly identical and thus cancel in calculating the free-energy change for the EtF → PD transition.

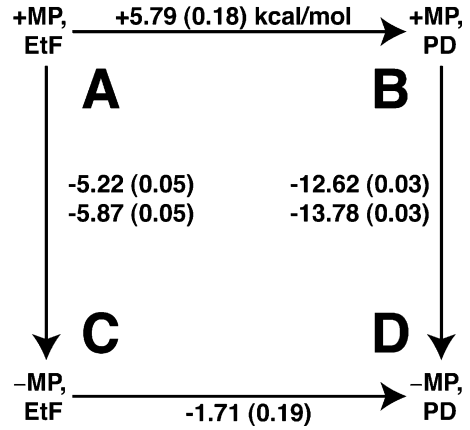


Figure 7. Thermodynamic cycle for EtF → PD conformational changes and +MP → −MP removal of Trp residues multipoles.

The EtF → PD conformational change for the −MP system has an associated ΔA of −1.71 (0.19) kcal/mol (the number in parentheses is the standard error of the mean), illustrating the slight preference for stacking of the Trp side chains over the NMR edge-to-face geometry in the absence of Trp side chain multipoles. Though favoring PD, the free energy difference between EtF and PD is not very large. This fact is reflected in Figure 4 wherein the Trp11 side chain appears to be in dynamic equilibrium between conformations parallel to and perpendicular to the Trp2 side chain. In contrast, the EtF → PD transition for the +MP system carries a +5.79 (0.18) kcal/mol cost, showing that the PD conformation is negligibly populated under laboratory conditions due to relatively unfavorable interactions involving Trp side chain electrostatic multipoles.

To validate the ΔA results for the +MP and −MP EtF → PD conformational changes, a complete thermodynamic cycle was constructed by calculating charging free energies for the EtF and the PD conformations and confirming that free energy changes from one corner of the cycle to the other were path-independent. Conversion of the +MP charges to the −MP charges was accomplished with linear charge scaling over 10 steps (λ = 0.95, 0.85, ..., 0.05) combined with thermodynamic integration. Figure 7 illustrates the thermodynamic cycle, with the horizontal legs showing the EtF → PD free energies and the vertical legs showing the charging free energies. Because the EtF and PD conformations for the +MP and −MP systems were subtly different in their Trp geometries due to the inclusion/exclusion of Trp side chain multipoles in the initial minimization stages of preparing these endpoints, charging free energies were calculated using both the +MP and −MP EtF and PD conformations; each of the vertical legs in Figure 7 shows both values respectively. From the thermodynamic cycle, the "hysteresis," defined as ΔA(A → B → D) − ΔA(A → C → D), is very small, less than the thermal energy RT (0.6 kcal/mol at 298 K). Specifically, the hysteresis using the +MP endpoints for calculating the EtF and PD charging free energies is +0.10 (0.27) kcal/mol and using the −MP endpoints is −0.41 (0.27) kcal/mol. The small hysteresis gives confidence that the EtF →

(24) Simonson, T. In *Computational biochemistry and biophysics*; Becker, O. M., MacKerell, A. D., Roux, B., Watanabe, M., Eds.; Marcel Dekker: New York, 2001; pp 169−197.
(25) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

Table 2. Energy Difference Decomposition for the EtF → PD Conformational Change, $\Delta E \equiv \langle E \rangle_{PD} - \langle E \rangle_{EtF}$

		+MP	–MP
vdW	ΔE(Trp side chain−water)	2.27(0.07)	2.91(0.08)
	ΔE(Trp side chain−Trp side chain)	−5.59(0)	−6.72(0)
elec	ΔE(Trp side chain−water)	2.43(0.15)	2.43(0.13)
	ΔE(Trp side chain−Trp side chain)	4.21(0)	0.70(0)

PD ΔA values are well converged and do not suffer from systematic error.

Table 2 lists the differences between the average energies (ΔE's) of the Trp side chain−water and Trp side chain−Trp side chain interactions in the EtF and PD conformations for both the +MP and −MP systems. The Trp side chain−Trp side chain van der Waals (vdW) interactions, as measured by the Lennard−Jones molecular mechanics energy, favor the PD conformation to the EtF because of the greater contact surface in the former, whereas Trp side chain−water vdW interactions favor the more solvent-exposed Trp side chain conformations of the EtF structure. Likewise, the energy decomposition shows Trp side chain−water electrostatic interactions to favor the more solvent-exposed Trp side chain conformations of the EtF structure. The near-identity in the electrostatic Trp side chain−water ΔE's for the +MP and −MP systems suggests that it is primarily the side chain NH dipole that is better solvated in the EtF conformation, with a near-zero contribution to conformational preference due to electrostatic interactions between water molecules and aromatic carbon and hydrogen atoms. In contrast to the similarity in the ΔE values for the +MP and −MP systems for the previous three types of interactions, there is a large difference between the +MP and −MP cases with respect to electrostatic components of ΔE for Trp side chain−Trp side chain interactions. In the absence of Trp side chain multipoles, Trp side chain−Trp side chain electrostatic interactions play a negligible part in determining the EtF vs PD conformational preference, despite the presence of the large dipoles of the side chain NH moieties. With the multipoles intact, the Trp side chain−Trp side chain electrostatic interactions in the EtF conformation are more favorable by 4.21 (0) kcal/mol compared to the PD conformation. Nearly 75% of the 5.79 kcal/mol free energy change (ΔA) for the +MP EtF → PD conformational change is accounted for by Trp side chain−Trp side chain electrostatic multipole interactions, despite the solvent-exposed nature of the four Trp side chains. Additionally, repulsion between side chain electrostatic multipoles in the +MP PD conformation accounts for the 1.13 kcal/mol $\Delta\Delta E$ in the Trp side chain−Trp side chain vdW interactions by preventing the same degree of favorable close-contact dispersion interactions as in the −MP PD conformation. Finally, we note that the values in Table 2 do not sum to the values in Figure 7. This is for two reasons: Table 2 does not show values for water−water interactions (large fluctuations in the energy of this interaction type preclude the collection of meaningful statistics on the time scale of the simulations), and the Table 2 values are average energies and do not account for entropic contributions, in contrast to the Figure 7 values, which are free energies, and therefore include entropic contributions and are equivalent to the reversible work at constant temperature and volume.

Conclusion

The present study demonstrates the importance of electrostatic multipole moments in determining the edge-to-face interaction geometries of the two pairs of solvent-exposed Trp side chains in the designed 12-residue β-hairpin trpzip2. The presence of multipoles leads to a very large free energy preference for the experimentally observed edge-to-face pairwise interaction geometries, in large part because the electrostatic multipole interaction energies between these side chains favor the edge-to-face conformation. The absence of these interactions leads to a modest preference for a parallel-displaced conformation, one that optimizes Trp side chain−Trp side chain van der Waals interactions and decreases the peptide's solvent-accessible surface area.

The importance of the inclusion of electrostatic multipoles in obtaining correct aromatic−aromatic geometries in solvated molecular mechanics studies of peptides has heretofore been assumed based on gas phase and solvated calculations on model compounds. The current results validate the use of multipoles originating from partial-charge centers located on aromatic hydrogen and carbon atoms in explicit-water molecular dynamics simulations by showing that their absence leads to loss of experimentally determined side chain conformations in a well-structured peptide. It is also seen that direct electrostatic interactions between the solvent-exposed Trp side chains strongly contribute to the observed geometry, contrary to possible expectations that solvent effects would minimize the role of direct side chain−side chain electrostatic interactions.

Due to the resource-intensive nature of explicit-water molecular dynamics simulations, simpler models are attractive in situations where extensive sampling is required, such as in docking small molecule inhibitors to macromolecular targets and in protein structure prediction. The prevalence of aromatic groups in druglike molecules and of aromatic−aromatic interactions in protein tertiary structure contacts argues for the inclusion of a term to account for aromatic−aromatic interactions in simple models. Prior work suggests the functional form for such a term: of various molecular mechanics approaches, a 12-point model with Coulomb and Lennard−Jones centers on all atoms was found to be the most promising for a force field approach to a benzene dimer.[26] The total interaction energy between two aromatic moieties decays much more rapidly than the underlying pairwise Coulomb interaction between partial charges, which independently decay as $1/r$, because it is an interaction between multipoles of an order greater than 1, not point charges. As an example, the quadrupole−quadrupole interaction characteristic of a benzene dimer decays as $1/(r^5)$. This rapid decay suggests Coulomb's law with truncation at ∼7 Å as a reasonable approximation, with possibly some empirical scaling of partial charges to account for environmental effects in a mean-field fashion. Inclusion of such an energy term may therefore be useful in situations wherein aromatic−aromatic interactions are expected to be important and where rigor must be balanced with speed of computation.

(26) Smith, G. D.; Jaffe, R. L. *J. Phys. Chem.* **1996**, *100*, 9624−9630.

Methods

The CHARMM22 all-atom force field[27] with a grid-based term for improved peptide backbone energetics[28,29]was used in conjunction with a modified version of the TIP3P water model.[30,31] Periodic boundary conditions were employed, atom-based force shifting[32] with a 12-Å cutoff was applied to electrostatic interactions, an atom-based switching function[33] was applied in the interval between 8 and 12 Å to smoothly truncate Lennard−Jones interactions, and a 14-Å atom-based pair list was maintained and updated heuristically. SHAKE[34] was used to constrain all bonds to hydrogen atoms at their equilibrium lengths and to maintain rigid water geometries. All simulations were performed with the CHARMM software.[33] Trajectories were visualized, and three-dimensional figures prepared using the VMD software.[35]

Molecular Dynamics. Following 1000 steps of steepest descent minimization,[36] the system was equilibrated for 20 ps with harmonic restraints on peptide heavy atom positions, followed by another 80 ps of unrestrained equilibration. Data for analysis were collected every 500 integration steps (1 ps) from a subsequent 10 ns of production molecular dynamics. Pressure was maintained at 1 atm using a Langevin piston[37] with a piston mass of 100 amu, a collision frequency of 10 ps^{-1}, and a piston bath temperature of 298 K. The simulation temperature was maintained at 298 K using velocity reassignment from a Gaussian distribution[38] every 500 dynamics steps during the 100-ps equilibration phase, and subsequently during the production phase using the Nosé−Hoover method[39] with a thermal-piston mass of 1000 kcal mol^{-1} ps^2. The equations of motion were integrated using the "leapfrog" alogorithm[40] with a 2-fs time step.

Free Energy Simulations. The protocol for the free energy simulations was the same as the above molecular dynamics protocol with the following exceptions: simulations were done at constant volume using average volumes calculated from the 10-ns +MP and −MP MD and were 96 731.99 $Å^3$ and 96 732.15 $Å^3$, respectively; the "velocity Verlet" integrator[41] was used; Trp side chain atom coordinates were constrained; and data were collected every 20 integration steps (40 fs). Prior to initial minimization, any water molecules overlapping peptide atoms were randomly displaced into the bulk solvent, and all simulations had 3183 water molecules. For the conformational change free energy calculations, each window had a 1-ns production phase, and for the charging free energy calculations, each window had a 0.5-ns production phase. Average energies were calculated from 6-ns trajectories of the +MP and −MP EtF and PD conformational endpoints. Standard errors of the mean were calculated using data from 100-ps blocks and confirmed with estimates of the statistical inefficiency of the simulations.[42] Bonded energy terms were excluded in the thermodynamic integration for the +MP and −MP EtF → PD conformational changes; average bonded energies from the unrestrained MD simulations showed that the EtF and PD conformations had identical bonded energies to within 0.39 kcal/mol.

Acknowledgment. We are grateful to Prof. Alexander MacKerell for providing parameters for the grid-based backbone energy term and to Dr. Daniel Price and Dr. Ilja Khavrutskii for helpful discussions. O.G. thanks the ARCS Foundation of San Diego for fellowship support. Financial support from the NIH (GM37554/GM48805) is appreciated.

JA043492E

(27) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(28) MacKerell, A. D.; Feig, M.; Brooks, C. L. *J. Am. Chem. Soc.* **2004**, *126*, 698−699.
(29) MacKerell, A. D.; Feig, M.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(30) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(31) Durell, S. R.; Brooks, B. R.; Ben-Naim, A. *J. Phys. Chem.* **1994**, *98*, 2198−2202.
(32) Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, *15*, 667−683.
(33) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(34) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(35) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.
(36) Levitt, M.; Lifson, S. *J. Mol. Biol.* **1969**, *46*, 269−279.
(37) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613−4621.
(38) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384−2393.
(39) Hoover, W. G. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695−1697.
(40) Hockney, R. W. In *Methods in computational physics*; Academic Press: New York, 1970; Vol. 9, pp 136−211.
(41) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1982**, *76*, 637−649.
(42) Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids*; Oxford University Press: Oxford, 1987.