THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Structural Analysis of α-Fetoprotein (AFP)-like Peptides with Anti-Breast-Cancer Properties

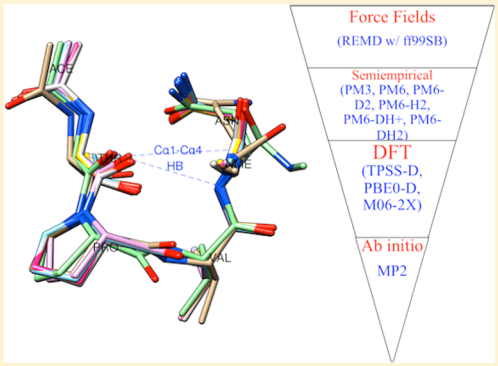
Berhane Temelso,[†,‡] Katherine A. Alser,[†] Arianne Gauthier,[‡] Amber Kay Palmer,[‡] and George C. Shields*[,†,‡]

[†]Dean's Office, College of Arts and Sciences, and Department of Chemistry, Bucknell University, Lewisburg, Pennsylvania 17837, United States

[‡]Dean's Office, College of Science and Technology, and Department of Chemistry & Physics, Armstrong Atlantic State University, 11935 Abercorn Street, Savannah, Georgia 31419, United States

Supporting Information

ABSTRACT: The abundance of α-fetoprotein (AFP), a natural protein produced by the fetal yolk sac during pregnancy, correlates with lower incidence of estrogen receptor positive (ER+) breast cancer. The pharmacophore region of AFP has been narrowed down to a four amino acid (AA) region in the third domain of the 591 AA peptide. Our computational study focuses on a 4-mer segment consisting of the amino acids threonine-proline-valine-asparagine (TPVN). We have run replica exchange molecular dynamics (REMD) simulations and used 120 configurational snapshots from the total trajectory as starting configurations for quantum chemical calculations. We optimized structures using semiempirical (PM3, PM6, PM6-D2, PM6-H2, PM6-DH+, PM6-DH2) and density functional methods (TPSS, PBE0, M06-2X). By comparing the accuracy of these methods against RI-MP2 benchmarks, we devised a protocol for calculating the lowest energy conformers of these peptides accurately and efficiently. This protocol screens out high-energy conformers using lower levels of theory and outlines a general method for predicting small peptide structures.



INTRODUCTION

Breast cancer kills more women than any other cancer except lung cancer. Activation of the estrogen receptor α (ERα) leads to the development of hormone-dependent breast cancers, and these cancers are known as estrogen receptor positive (ER+). Drugs that inhibit active folding of ERα have long been used as a primary therapeutic strategy, and aromatase inhibitors that block estradiol synthesis and antiestrogens that compete with hormone binding to the receptor are routinely prescribed.[1] Unfortunately, tumor resistance almost invariably emerges from prescribed antihormonal approaches. For instance, tamoxifen, a drug used to treat ER+ breast cancer, is effective for only two-thirds of patients, and the remaining one-third become resistant or develop lethal side affects such as uterine tumors. One promising approach is a multitarget strategy affecting key regulatory domains distinct from the ligand-binding pocket of ERα, and the synthesis of peptides that may specifically inhibit intra- or intermolecular interactions has been proposed to achieve this goal.[1] Molecular modeling shows promise in applications to anticancer drug design.[2]

α-Fetoprotein (AFP) is a protein produced by the fetal yolk sac that circulates in the maternal serum, and high concentrations of this 591 amino acid long protein and peptides derived from AFP reduce a woman's risk of breast cancer.[3−9] In addition, women who have a full term pregnancy before the age of twenty-seven reduce their risk of ER+ breast cancer, whereas those over twenty-seven have a risk profile that remains the same or increases slightly.[10] AFP binds to an unknown receptor that causes endocrine action on premalignant tissues that would have developed into cancer many years later. The cancer growth suppression of AFP was originally attributed to a 34-residue peptide (GIP-34) in the third domain covering amino acids 447−480 and then to an eight residue or cyclic 9-mer amino acid sequence.[11−13] A breakthrough resulted from our previous REMD simulations, which allowed for the prediction that peptides as small as 4-mers and 5-mers adopt a reverse β-turn conformation and should therefore retain biological activity.[14,15] These small peptides inhibit cancer growth similarly to the larger peptides.[14,16,17] One of the main aims of the present paper is to show how quantum calculations can be combined with molecular dynamics simulations to gain a high degree of confidence in the ensemble of structures that these small peptides possess in aqueous solution.

Determination of the structure and activity of proteins has been one of the holy grails of modern science.[18] In the past, techniques such as X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy were the main methods for

Received: January 1, 2014
Revised: March 30, 2014
Published: April 5, 2014

dx.doi.org/10.1021/jp500017b | J. Phys. Chem. B 2014, 118, 4514−4526

understanding structure and function, and these techniques remain invaluable today. NMR provides the three-dimensional structure of proteins in solution[19] or in the solid state,[20] whereas X-ray crystallography allows structure determination of single crystals.[21−24] In addition, mass spectrometry has become a powerful tool for analyzing peptide and protein structure.[25,26] Understanding solvation effects and how best to model them has made great progress over the past few decades.[27−44] Today, theoretical methods have made striking progress, and these improvements combined with the development of ever more powerful computers have added to our ability to understand complex biological problems.[45−48] In particular, the complex problem of protein folding has garnered a great deal of research interest and significant progress has been achieved.[49] Methods ranging from molecular dynamics to quantum mechanics are being used to gain insight into many different protein systems.[50−87] Just as protein folding complicates the determination of large protein structures, the inherent conformational flexibility of smaller peptides presents a similar challenge to understanding peptide structure because the variety of conformations that peptides adopt in solution depend on specific conditions and the presence of various ligands. Computational methods are providing alternative pathways to peptide structure determination. For instance, reversible β-peptide folding in solution at atomic detail was demonstrated by van Gunsteren's group,[88] in agreement with NMR data.[89] Another interesting example is a molecular dynamics simulation that folded the 20-residue Trp-cage peptide to a high degree of fidelity relative to the 1 Å crystal structure.[90,91] Ultimately, modeling work in this field is invaluable because it teaches us to ask and seek answers to systematic questions about sequence/structure/function relationships and about the underlying forces that stabilize biomolecular structure. Thus, the first aim of this study is to outline a protocol that is a fast, efficient, and accurate procedure for determining the structure of small peptides. The second aim is to provide convincing data to support our understanding of the ensemble of structures of the 4-mer and 5-mer peptides that have been shown to inhibit breast cancer in animal models.

METHODS

Unlike conventional molecular dynamics (MD) simulations that can be trapped in local energy wells, replica exchange MD (REMD) overcomes this problem by running multiple replicas and allowing exchanges between them at fixed intervals. The REMD simulations employed the AMBER 9 program[92] suite with AMBER's ff99SB force field[56] and were used to predict the active conformation of small peptides that possessed anti-breast-cancer activity.[14,15] Initial conjugate-gradient energy minimizations were performed on all systems studied, employing a 0.01 kcal·mol^{-1}·Å^{-1} convergence criterion in the energy gradient. Simulations were in implicit solvent, using the generalized Born (GB) model[36] implemented in AMBER with the default radii.[93] The TPVN tetrapeptide studied here was capped with an acetyl (ACE) group on the N-terminus at the threonine and an *N*-methylamine (NME) group on the C-terminus at the asparagine. The complete protocol for the REMD simulations has been explained elsewhere.[15]

A total of 120 structures were extracted from the T = 304 K replica at even intervals over the course of the simulation. These structures served as input structures for semiempirical, DFT, and ab initio calculations. Starting with these configurations from the REMD simulations, we optimized each structure using a variety of methods spanning a large range in terms of computational cost and accuracy. By comparing the accuracy of these methods with RI-MP2 (resolution-of-the-identity Møller−Plesset perturbation), we devised a protocol that would quickly deliver the most stable conformers both in gas phase and in solution. RI-MP2[94−97] is a relatively inexpensive yet accurate method for including electron correlation in quantum chemistry calculations.[36] Electron correlation is essential for capturing noncovalent interactions, such as hydrogen bonding and dispersion, which determine peptide and protein structure. MP2 captures these interactions well for water clusters where high-resolution experiments can be compared with theory.[98−100] The best approach should identify the most stable conformers accurately and efficiently, and scale to larger peptides, such as the 8-mer and 9-mer. With this in mind, we have developed a protocol using extensive benchmarking for a 4-mer peptide, TPVN.

First we calculated the optimal structures using molecular orbital theory based semiempirical (SE) methods (PM3,[101] PM6,[102,103] PM6-D2,[104,105] PM6-H2,[104,105] PM6-DH+,[104,105] and PM6-DH2[104,105]) because they are computationally cheap and they produce sufficiently accurate structures.[36,106] Since the heat of formation (ΔH_f) and relative (conformational) energies from semiempirical methods are unreliable, we next calculated absolute and conformational energies using the RI-MP2 method with Dunning's correlation consistent double-ζ basis set, cc-pVDZ (VDZ),[107] on the optimized structures obtained from the semiempirical methods. Diffuse or augmented basis functions were not included because the RI-MP2 energy and gradient calculations would have been very expensive. In addition, we optimized the geometries of all REMD structures using density functional theory (DFT) methods (TPSS,[108] PBE0,[109] and M06-2X[110]). The energy and gradient calculations with TPSS and PBE0 functionals are substantially accelerated using two approximations. First, the Coulomb matrix (J) was calculated using resolution-of-the-identity with a proper fitting basis in the RI-J approximation. Second, the exchange contribution (X) employed Neese's chain-of-spheres exchange (COSX) technique[111−113] implemented in ORCA 2.8, 2.9, and 3.0.[114] The TPSS and PBE0 functionals were also augmented with an empirical van der Waals correction that significantly improves the performance of DFT methods for capturing long-range interactions.[115−119] The development of these dispersion corrections has undergone many iterations from the original empirical D1[115] and D2[120] corrections to the atom-pairwise D3[121] and atom-pairwise with Becke−Johnson damping function (D3BJ).[122] We added the D3BJ dispersion corrections to the TPSS and PBE0 calculations. Therefore, what we describe below as TPSS-D and PBE0-D are actually RIJCOSX-TPSS-D3BJ and RIJCOSX-PBE0-D3BJ. Because conformational energies that DFT methods predict are not always reliable, we calculated the energies of the optimized DFT structures using RI-MP2/VDZ. Geometry optimizations were repeated in aqueous solution using the COSMO (conductor-like screening model) implicit solvation model.[28,29] In this polarizable continuum model, the dielectric constant of the solvent (pure water) is taken to be 74.2 at the average human body temperature of 37 °C (98.6 °F).[123] All SE calculations used MOPAC 2012,[124] while all other DFT and RI-MP2 computations used ORCA 3.0,[114] with the exception of M06-2X for which Gaussian09, rev. D.01 was used.[125]

In our initial study, 120 structures extracted from REMD simulations were optimized using the different methods above,

dx.doi.org/10.1021/jp500017b | J. Phys. Chem. B 2014, 118, 4514−4526

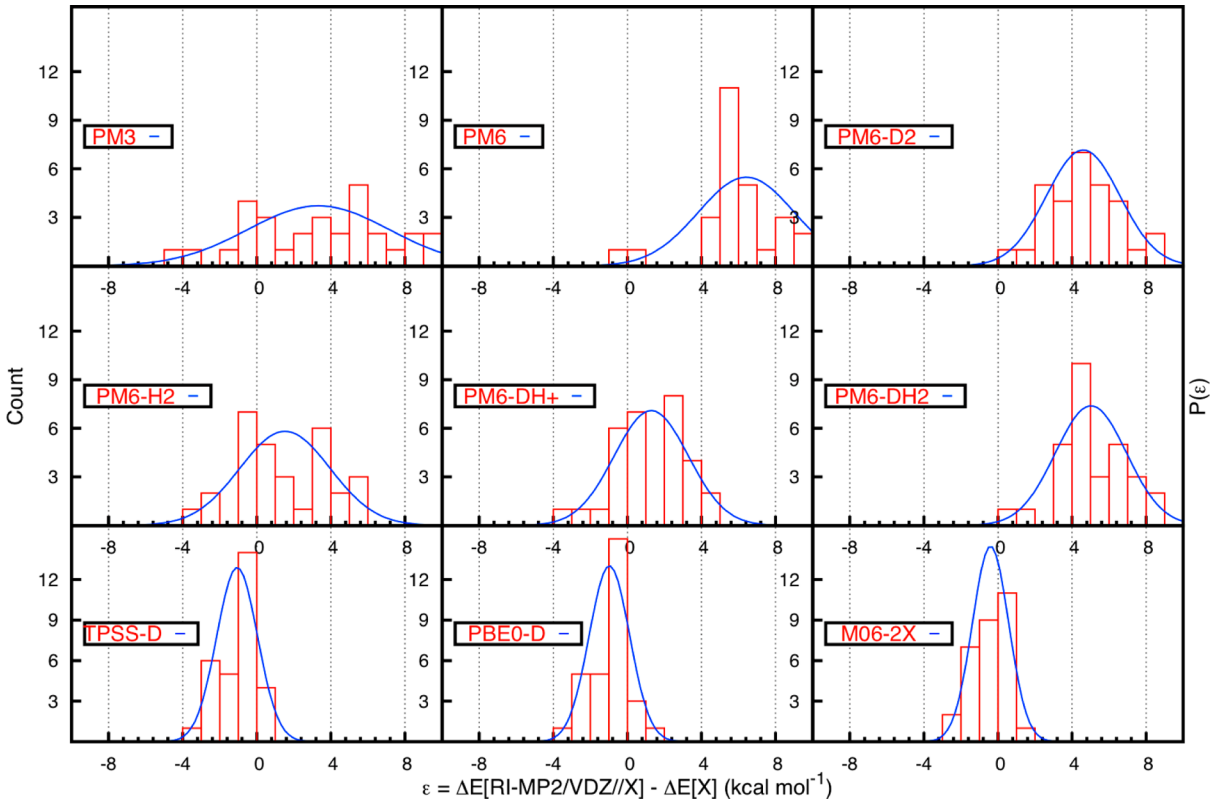


Figure 1. A histogram (red) of the error in conformational energy of the 30 lowest energy conformers in methods (X) relative to RI-MP2/VDZ//X in the gas phase. A corresponding Gaussian function based on the MSE (μ_S) and STDEV (σ_S) for N = 30 is plotted in blue. See Table 1 for the MSE (μ_S) and STDEV (σ_S) for N = 10, 20, and 30.

and their conformational energy was calculated using each method (X = PM3, PM6, PM6-D2, PM6-H2, PM6-DH+, PM6-DH2, TPSS-D, PBE0-D, M06-2X) as well as RI-MP2/VDZ single point energies performed on the X optimized geometries (RI-MP2/VDZ//X). The conformational energies calculated using each method X (ΔE^i_X) is compared with that calculated from the RI-MP2/VDZ energy at the geometry optimized using method X ($\Delta E^i_{\text{RI-MP2/VDZ//X}}$). The error in conformational energy (ε_i), defined as the difference between the conformational energy calculated using RI-MP2/VDZ//X and X, is a measure of the reliability of each method X.

$$\varepsilon_i = \Delta E^i_{\text{RI-MP2/VDZ//X}} - \Delta E^i_X \tag{1}$$

The mean signed error (MSE), mean unsigned error (MUE), and corresponding standard deviations (STDEV) for the N lowest energy conformers by their RI-MP2/VDZ//X energies are then

$$\mu_S = \text{MSE} = \frac{1}{N}\sum_{i=1}^{N}\varepsilon_i$$

$$\mu_U = \text{MUE} = \frac{1}{N}\sum_{i=1}^{N}|\varepsilon_i|$$

$$\sigma_S = \text{STDEV}_S = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\varepsilon_i - \mu_S)^2}$$

$$\sigma_U = \text{STDEV}_U = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\varepsilon_i - \mu_U)^2} \tag{2}$$

As illustrated in Figures 1 and 2 for the 30 lowest energy conformers (N = 30), the error in the conformational energies that the different methods predict relative to RI-MP2 does not follow a normal distribution, particularly for the highly parametrized semiempirical methods. Even though these errors may not necessarily follow a normal distribution especially given the small data set, plotting them as such serves an illustrative purpose. Therefore, we have used the mean errors (μ) and standard deviation (σ) to plot Gaussian distribution functions:

$$P(\varepsilon) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(\varepsilon-\mu)^2/(2\sigma^2)} \tag{3}$$

N is chosen to be the 10, 20, or 30 lowest energy conformers.

Using a multilevel protocol to obtain structures is a feasible approach for obtaining accurate quantum mechanical structures.[126−133] A protocol was developed to calculate low-energy conformers most efficiently based on these error profiles and computational cost considerations.

RESULTS

Approximately 67% of the 120,000 peptide structures from the REMD simulation were in a β-turn conformation, and 67% of the 120 starting structures for higher-level minimizations were also in a β-turn conformation. Each of the 120 structures was minimized using the four SE methods, three density functionals, and RI-MP2. On the basis of the error profile (relative to RI-MP2) that we compiled for the 10, 20, and 30 lowest energy structures for each method, we are proposing a protocol for obtaining an ensemble of RI-MP2-quality low-energy conformers. Table 1 displays the mean signed errors (MSEs)

dx.doi.org/10.1021/jp500017b | J. Phys. Chem. B 2014, 118, 4514−4526

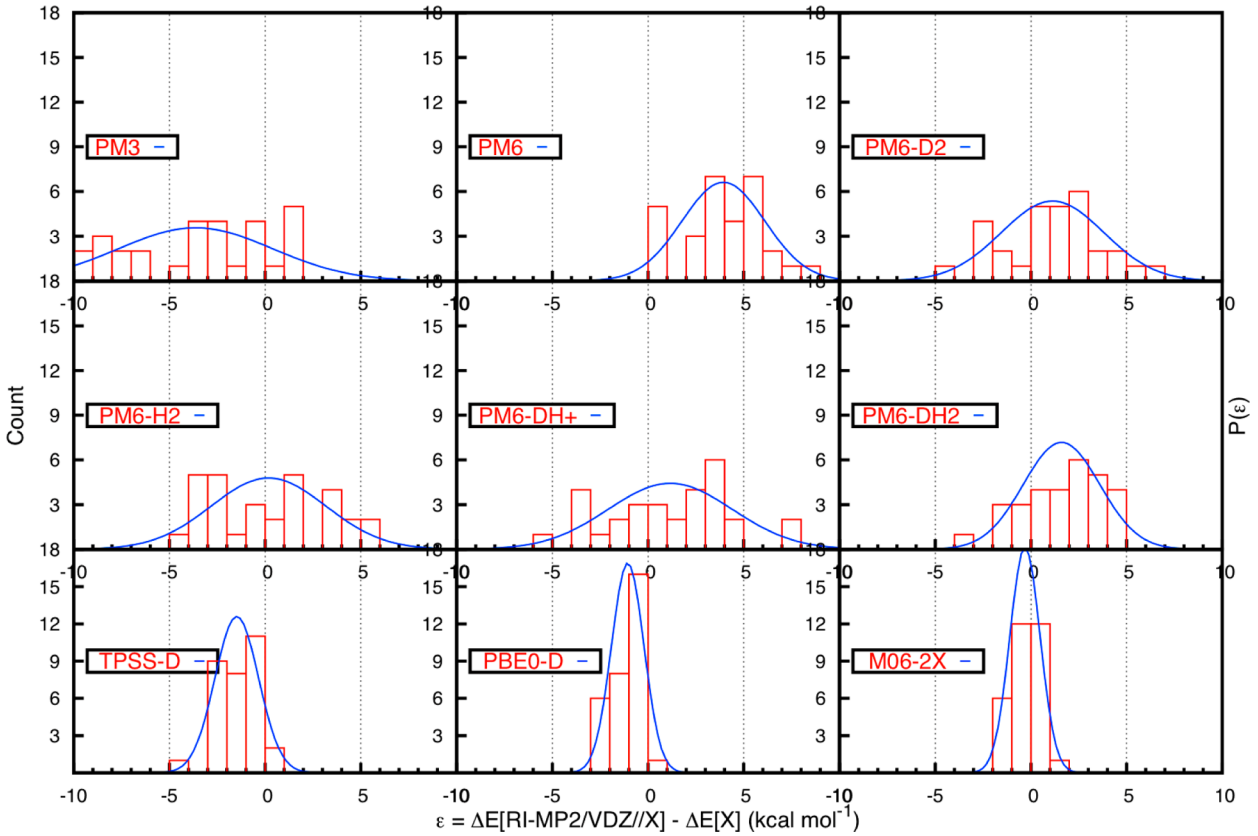


Figure 2. A histogram (red) of the error in conformational energy of the 30 lowest energy conformers in methods (X) relative to RI-MP2/VDZ//X in aqueous solution. A corresponding Gaussian function based on the MSE (μ_S) and STDEV (σ_S) for N = 30 is plotted in blue. See Table 2 for the MSE (μ_S) and STDEV (σ_S) for N = 10, 20, and 30.

Table 1. Mean Signed Error (MSE) and Standard Deviation (STDEV) of the Conformational Energies of the 10, 20, and 30 Lowest Energy TPVN Conformers in the Gas Phase Calculated Using Different Methods, X, Relative to RI-MP2/VDZ//X

	ΔE[RI-MP2/VDZ//X] − ΔE[X] (kcal mol^{-1})					
	10 lowest		20 lowest		30 lowest	
method (X)	MSE	STDEV	MSE	STDEV	MSE	STDEV
PM3	0.29	3.10	1.47	3.11	3.33	3.82
PM6	4.39	2.30	5.68	2.43	6.42	2.59
PM6-D2	3.48	1.79	3.70	1.53	4.60	1.98
PM6-H2	0.35	1.69	0.80	2.07	1.52	2.44
PM6-DH+	0.90	2.14	1.17	1.89	1.29	2.00
PM6-DH2	3.79	1.74	4.26	1.58	5.02	1.92
TPSS-D	−0.43	0.44	−0.77	0.87	−1.06	1.10
PBE0-D	−0.34	0.43	−0.68	0.87	−0.97	1.09
M06-2X	0.19	0.73	−0.34	1.11	−0.40	0.98

Table 2. MSE and STDEV of the Conformational Energies of the 10, 20, and 30 Lowest Energy TPVN Conformers in Aqueous Solution Calculated Using Methods, X, Relative to RI-MP2/VDZ//X

	ΔE[RI-MP2/VDZ//X] − ΔE[X] (kcal mol^{-1})					
	10 lowest		20 lowest		30 lowest	
method (X)	MSE	STDEV	MSE	STDEV	MSE	STDEV
PM3	−8.25	2.08	−5.41	3.51	−3.63	4.03
PM6	1.87	1.65	3.03	1.90	3.94	2.17
PM6-D2	−0.93	2.09	0.30	2.44	1.13	2.68
PM6-H2	−2.35	1.66	−1.39	2.20	0.18	3.00
PM6-DH+	−1.85	2.25	0.17	3.48	1.16	3.24
PM6-DH2	−0.38	1.50	0.69	1.78	1.60	2.00
TPSS-D	−1.54	1.65	−1.37	1.25	−1.48	1.14
PBE0-D	−0.74	0.79	−0.96	0.90	−1.06	0.85
M06-2X	−0.49	0.85	−0.33	0.89	−0.31	0.79

and standard deviations (STDEVs) for the 10, 20, and 30 lowest energy conformers of TPVN for each method X relative to the RI-MP2/VDZ//X results in the gas phase, and Table 2 contains similar results in aqueous solution. Figure 3 illustrates the error distribution based on the MSE and STDEV for the 20 lowest energy conformers found using each method in the form of a Gaussian function both in the gas phase (top) and in aqueous solution (bottom). The mean unsigned errors (MUEs) and STDEVs are given in the Supporting Information in Tables S2 and S3 and are plotted in Figure S1. Both the MSEs and STDEVs of the SE methods are large even after corrections for dispersion (D) and hydrogen bonding interactions (H). Figures 4 and 5 show the overlay of the lowest energy RI-MP2/VDZ//X conformers of TPVN in the gas phase and aqueous solution, respectively. Each figure lists the root-mean-square deviation (RMSD) of the heavy atoms of each global minimum against the RI-MP2/VDZ global minimum structure in the gas phase and in aqueous solution. While DFT structures match best with the RI-MP2 global minimum, it is clear that all the methods with the exception of PM3 give reasonably good structures. Table 3 shows the percentage of the 30 lowest energy minima of TPVN in the gas phase and aqueous solution that have a β-turn using (a) the

dx.doi.org/10.1021/jp500017b | J. Phys. Chem. B 2014, 118, 4514−4526

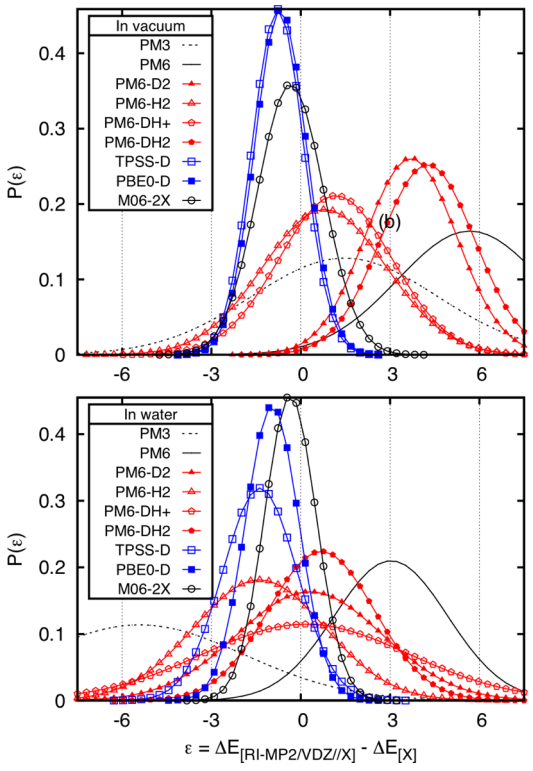


Figure 3. Gaussian distribution of errors in the conformational energies of the 20 lowest energy conformers predicted by different methods relative to RI-MP2 for TPVN in the gas phase (top) and in aqueous solution (bottom). See Tables 1 and 2 for the MSEs (μ_S) and STDEVs (σ_S) for N = 20.

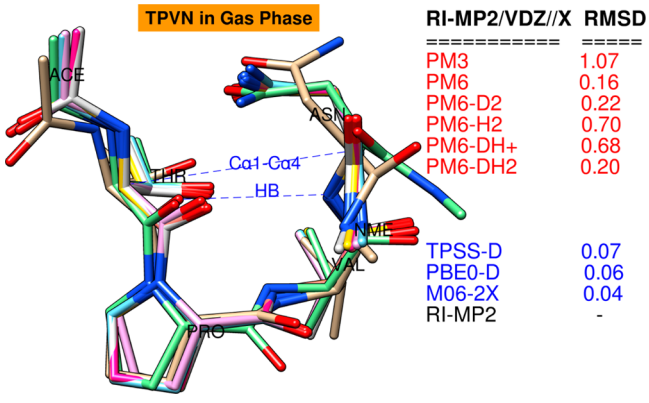


Figure 4. Overlays of the lowest energy RI-MP2/VDZ//X conformers of TPVN in the gas phase, where X is a semiempirical (PM3, PM6, PM6-H2, PM6-D2, PM6-DH+, PM6-DH2) or density functional (TPSS-D, PBE0-D, M06-2X) method. The RMSDs are determined by superimposing the heavy atoms of RI-MP2/VDZ//X global minima on the RI-MP2/VDZ global minimum.

comprehensive distance between the C_α of the first (threonine) and fourth (asparagine) residues ($C_{\alpha1}$−$C_{\alpha4}$ < 7 Å) criteria and (b) the intraturn hydrogen bond between these two residues (H-bond distance <2.5 Å and H-bond angle between 140° and 180°). Almost all the 30 lowest energy conformers have a β-turn both based on the $C_{\alpha1}$−$C_{\alpha4}$ distance criteria ($C_{\alpha1}$−$C_{\alpha4}$ < 7 Å), as well as the presence of a hydrogen bond between the threonine carbonyl group and the amide hydrogen of the asparagine, particularly in the gas phase. In aqueous solution, a majority of the conformers have a β-turn based on the $C_{\alpha1}$−$C_{\alpha4}$

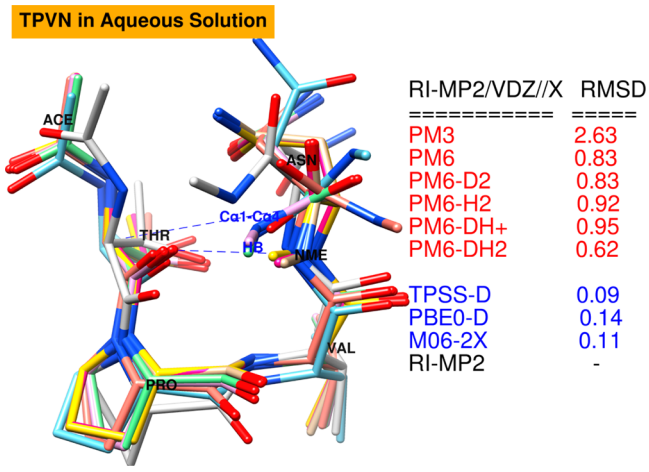


Figure 5. Overlays of the lowest energy RI-MP2/VDZ//X conformers of TPVN in water, where X is a semiempirical (PM3, PM6, PM6-H2, PM6-D2, PM6-DH+, PM6-DH2) or density functional (TPSS-D, PBE0-D, M06-2X) method. The RMSDs are determined by superimposing the heavy atoms of RI-MP2/VDZ//X global minima on the RI-MP2/VDZ global minimum.

Table 3. The Percentage of the Lowest 30 Energy Conformers Satisfying Different β-Turn Criteria for TPVN in the Gas Phase and in Water[a]

	gas phase		water	
method (X)	% β-turn $C_{\alpha1}$−$C_{\alpha4}$	% β-turn HB	% β-turn $C_{\alpha1}$−$C_{\alpha4}$	% β-turn HB
REMD/ff99SB[b]			67	50
PM3	79	14	80	27
PM6	100	45	82	36
PM6-D2	97	50	83	24
PM6-H2	100	67	79	46
PM6-DH+	100	54	87	35
PM6-DH2	100	54	100	62
TPSS-D	100	67	100	55
PBE0-D	100	67	100	67
M06-2X	90	67	93	67
RI-MP2	100	80	100	80

[a]The two criteria for a β-turn are (1) $C_{\alpha1}$−$C_{\alpha4}$ < 7 Å and (2) HB distance < 2.5 Å and 140° < HB angle < 180°. The HB here is between the threonine carbonyl group and the amide hydrogen of the asparagine. [b]Represents 120 structures extracted from 304 K replica of an REMD simulation with the ff99SB[56] force field employing the generalized Born (GB) implicit solvent model.

criteria, but fewer satisfy the hydrogen bonding criteria. Looking at the 15 lowest energy conformers reveals that virtually all of them exhibit a β-turn. Figure 6 shows an overlay of these 15 conformers at the RI-MP2/VDZ level in gas phase, and Figure 7 reveals their structures in aqueous solution. At the RI-MP2/VDZ//X level of theory, the relative energy of these 15 conformers spans a 10 kcal/mol range in the gas phase and 6 kcal/mol in aqueous solution. The peptide backbone in all these structures is very similar, and most of the differences in energy and conformation can be attributed to changes in the side chains and capping groups. Figure 8 summarizes the results of this study in one diagram that outlines the most efficient protocol for finding the lowest energy peptide structures starting from a vast ensemble of structures.

dx.doi.org/10.1021/jp500017b | J. Phys. Chem. B 2014, 118, 4514−4526

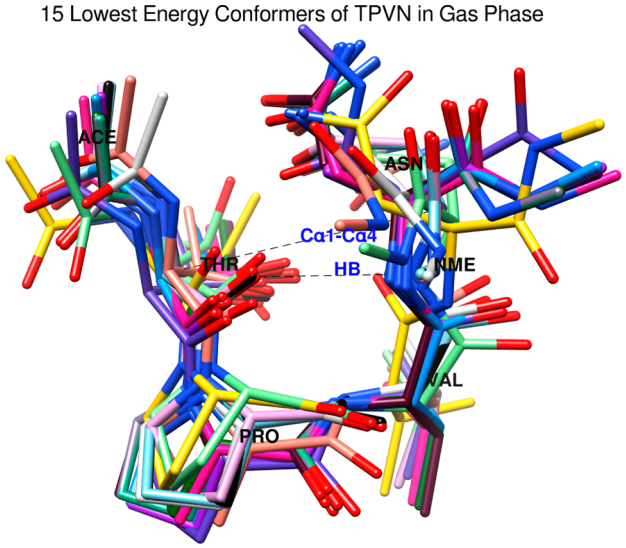


Figure 6. Overlay of the 15 lowest energy RI-MP2/VDZ conformers of TPVN in the gas phase. They all have a β-turn conformation.

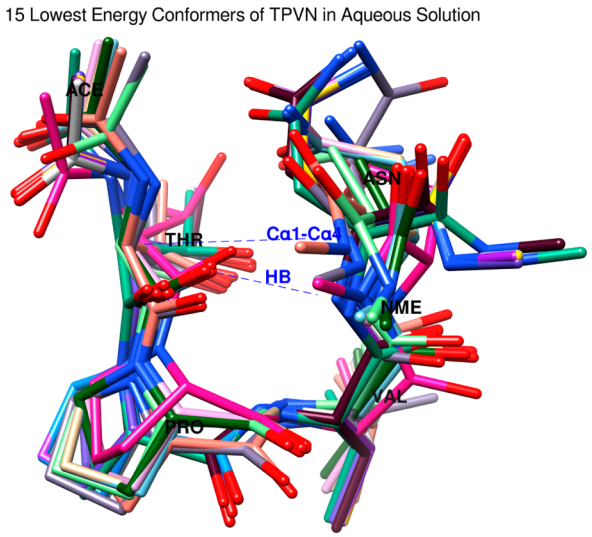


Figure 7. Overlays of the 15 lowest energy RI-MP2/VDZ conformers of TPVN in an aqueous solution. They all have a β-turn conformation.

DISCUSSION

The conformational energies predicted by methods X and RI-MP2/VDZ//X varied substantially for TPVN. While the PM3 method is well-known for its ability to model the geometries of intermolecular hydrogen bonded structures,[106,134−137] it is unable to deliver quantitative energies unless reparametrized for a specific system of interest.[138] It is no surprise that the energies obtained from the PM3 and PM6 semiempirical methods for these peptides, without correction for dispersion and hydrogen bonding, are not reliable.

For gas phase conformational energies, a comparison of the lowest 10 peptides obtained from semiempirical (SE) methods (Table 1) reveal that the RI-MP2/VDZ//PM6-H2 and RI-MP2/VDZ//PM6-DH+ methods yield the lowest MSE (0.35 and 0.90 kcal/mol) relative to RI-MP2/VDZ, but their STDEV is large (1.69 and 2.14 kcal/mol). Nevertheless, they outperform the other SE methods even for the larger grouping of 20 or 30 lowest energy conformers. In aqueous solution, RI-MP2/VDZ//PM6-D2 and RI-MP2/VDZ//PM6-DH2 yield the

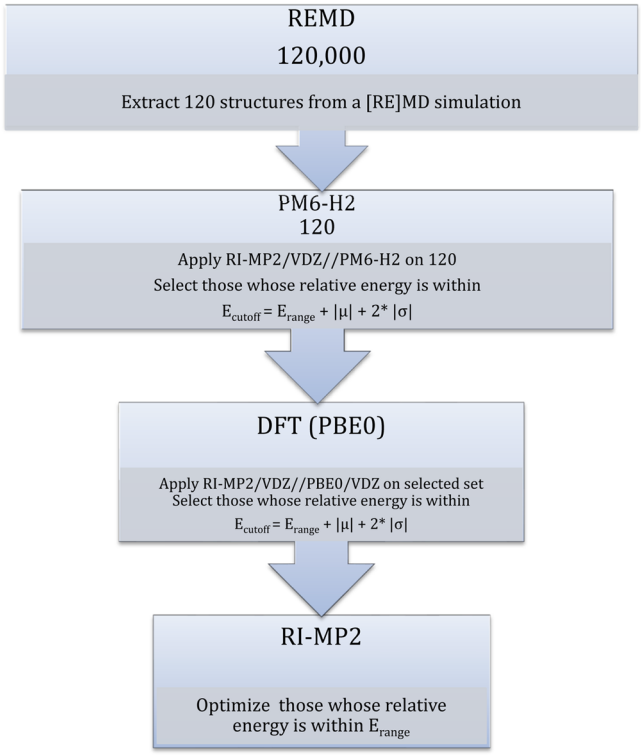


Figure 8. An efficient protocol that delivers low-energy conformers whose relative energy is within an energy range (E_{range}) of the global minimum. Optimize 120 REMD structures using PM6-H2, and select those within an energy cutoff (E_{cutoff}) using their RI-MP2/VDZ//PM6-H2 energy. Further optimize selected set using PBE0-D/VDZ, and select those within a new energy cutoff (E_{cutoff}) on the basis of their RI-MP2/VDZ//PBE0/VDZ energy. Finally, minimize the structures within the desired energy range (E_{range}) using RI-MP2/VDZ to obtain the most stable set of conformers.

smallest MSEs and STDEVs for the 10, 20, and 30 lowest energy conformers. These results indicate that no one SE method stands out among the six that have been evaluated here. In general, the dispersion or hydrogen bonding corrected methods perform better than the original PM3 and PM6 parametrizations and they should be used because their computational cost is essentially equivalent to PM6. Faver et al.[72] analyzed the performance of different methods including PM3, PM6, and PM6-DH2 in predicting the interaction energy between fragments in the protein ubiquitin (1UBQ fragment database) by comparing them against benchmark MP2/CBS and CCSD(T)/CBS values. They concluded that PM6-DH2 substantially improves on both PM3 and PM6, especially for polar interactions, but the STDEV of the errors in all three methods remains large. Faccioli and co-workers have shown that an AMBER molecular dynamics simulation of a tetraalanine peptide and a dominant reaction pathways (DRP) simulation using PM3 both fold to the elementary unit of an α-helix starting from an initial coil configuration.[139] They suggest that the similarity of the two paths suggests a two-step approach, with a quantum DRP simulation following from a classical simulation. Our results suggest that one of the modified PM6 methods would be a better semiempirical approach for quantum DRP of peptides.

For the DFT methods, it is not surprising that the hybrid PBE0-D and hybrid meta-GGA functionals M06-2X mimic the RI-MP2/VDZ conformational energy profile better than the

dx.doi.org/10.1021/jp500017b | J. Phys. Chem. B 2014, 118, 4514−4526

TPSS-D meta-GGA functional. TPSS and PBE0 functionals have performed reasonably well when augmented with the empirical dispersion correction (-D). The density functional methods with a dispersion correction performed quite well in that both energies (Table 1, MSE < 1.1 kcal/mol) and structures (Figure 4, RMSD < 0.1 Å) compare favorably with RI-MP2. The inclusion of a dispersion correction is critical. Valdes et al.[140] employed a range of wave function, density functional, and semiempirical methods to obtain benchmark structures for five dipeptides and tripeptides with aromatic side chains. They concluded that only methods that accounted in some way for London dispersion forces gave reasonably good results compared with CCSD(T). A recent review summarized the performance of different semiempirical, DFT, and ab initio methods in predicting the structures and binding energies of the S22 set of noncovalently bound molecules.[141] One of the key conclusions was that empirical dispersion corrections improve the accuracy of both semiempirical and DFT methods dramatically.

In aqueous solution, the MSEs and STDEVs computed with PM6-D2, PM6-H2, PM6-DH+, and PM6-DH2 using the COSMO solvation model are generally comparable (Table 2), while PM6 captures the first 10 structures reasonably well and then falls off for the next 10 and 20 conformers. Comparing DFT methods, M06-2X and PBE0-D perform well with MSEs largely below 1 kcal/mol while TPSS-D has larger MSEs. Figure 3 displays a Gaussian (normal) distribution of the errors in the different methods found by plugging in the MSEs (μ_S) and STDEVs (σ_s) from Tables 1 and 2 to eq 3 for the 20 lowest energy conformers (N = 20). The top of the figure illustrates the error distribution in the gas phase, and the bottom of the figure corresponds to that in aqueous solution. All three density functionals perform well, but PBE0-D and M06-2X give more consistent error profiles in both the gas phase and aqueous solution.

A comparison of the lowest energy conformers using the different methods (RI-MP2/VDZ//X) showed reasonably good agreement with the RI-MP2/VDZ optimized geometries (Figures 4 and 5). An overlay of the gas phase structures in Figure 4 reveals that the RI-MP2/VDZ//X global minima compare well with the RI-MP2/VDZ analogue. The agreement between the DFT methods and RI-MP2 is particularly good, with the heavy atom RMSDs being less than 0.1 Å. In aqueous solution, both the semiempirical and DFT results deteriorate somewhat. The disagreement is particularly noticeable for the semiempirical methods. A closer examination of the semiempirical RI-MP2/VDZ//X global minima reveals that most of their deviation from the RI-MP2 global minimum is in the capping groups (ACE and NME) rather than the main peptide residues. In general, including solvation effects using the implicit COSMO model appears to have changed structures very little, as evidenced by the similarity between the global minima in the gas phase (Figure 4) and aqueous solution (Figure 5). The gas phase structures clearly capture the central feature of the TPVN peptide, so geometry optimization in solution may not even be necessary. Comparing the COSMO model to other solvation models is beyond the scope of this work. However, it should be noted that the use of implicit solvation models in REMD simulations overstabilizes ion pairs and affects the secondary structure of proteins.[142,143] Since our peptide has no ionic groups or a salt bridge, implicit solvation should work reasonably well.

A key secondary structural motif in the results is that of a type I reverse turn, or β-turn, where the carbonyl oxygen of the first amino acid is hydrogen-bonded to the amino hydrogen of fourth amino acid. There are two other diagnostics for determining whether a given peptide structure adopts a type I reverse β-turn conformation. For a given sequence of four amino acids, this reverse turn can be defined by the phi (ϕ) and psi (ψ) angles of the second and third amino acids of the sequence. An ideal type I β-turn has ϕ/ψ values of −60° and −30° for the second amino acid and ϕ/ψ values of −90° and 0° for the third amino acid. In the REMD simulation of the TPVN peptide, the average ϕ/ψ values for the proline were −68°/−19°, and for the valine, they were −88°/−11°.[14] Since only about 75% of β-turns have the intraturn hydrogen bond, a more definitive structural definition is that the distance between the C_α atoms of residues one and four are less than 7 Å.[144]

Based on the $C_{\alpha1}-C_{\alpha4}$ interresidue distance criterion ($R[C_{\alpha1}-C_{\alpha4}]$ < 7 Å), over 79% of the ensemble of 30 lowest energy TPVN structures in the gas phase in each method has a β-turn. Semiempirical methods suggest that only 14−67% of these structures have a hydrogen bond between the threonine carbonyl group and amide hydrogen of the asparagine, but their high β-turn ratios based on the $C_{\alpha1}-C_{\alpha4}$ interresidue distance criterion stems from the presence of other hydrogen bonds holding these structures in this bent conformation. It is a feature of β-turns that this reverse turn exists even when an actual intraturn hydrogen bond between residue one and four is not present, approximately 25% of the time, and this feature is captured by the DFT methods. Some of these other hydrogen bonds are between the capping groups (ACE and NME) and the other standard residues (Thr, Pro, Val, Asn), as well as between the standard residues themselves. In fact, common hydrogen bonds include those between the carbonyl oxygen of any residue and an amide hydrogen of another residue. Also, the threonine's hydroxyl side chain often serves both as a hydrogen bond donor and as an acceptor. Therefore, the hydrogen bond between the threonine carbonyl group and amide hydrogen of the asparagine is not the only interaction keeping the peptide in a β-turn conformation. It is not unusual to observe between three and five hydrogen bonds keeping these structures in a β-turn conformation. Density functionals and RI-MP2 methods lead to the prediction that a majority (>67%) of the 30 lowest energy conformers have the canonical hydrogen bond that is at least partly responsible for the 90−100% β-turn ratio in gas phase.

Placing the peptide in an aqueous solution does not yield significantly different structures. It lowers the percentage of structures having a β-turn very slightly for the semiempirical methods (Table 3). The difference between the % β-turn in vacuum and that in solution is more pronounced for the semiempirical methods than it is for DFT and RI-MP2, as can be seen from the lower number of structures satisfying both % β-turn criteria for the SE methods in solution compared with the gas phase (Table 3).

The 15 lowest energy RI-MP2/VDZ conformers in the gas phase are shown in Figure 6, and those in aqueous solution are displayed in Figure 7. In both cases, all (100%) the conformers adapt a β-turn conformation and 13 of 15 (87%) have a canonical hydrogen bond between the threonine carbonyl group and amide hydrogen of the asparagine. The 15 structures in the gas phase and in aqueous solution span a 10 and 6 kcal/mol range, respectively. In each case, there are two conformers that lie within 1 kcal/mol of the respective global minimum,

dx.doi.org/10.1021/jp500017b | J. Phys. Chem. B 2014, 118, 4514−4526

excluding the global minimum itself. All semiempirical methods, density functionals, and the RI-MP2 method reveal that the lowest energy conformers have a β-turn in the case of TPVN, which is the main pharmacophore region for AFP and related peptides that show anti-breast-cancer activity.[14,16] We strongly confirm the evidence for the presence of a β-turn in the AFPeps that was predicted by REMD simulations using the AMBER force field.[14,15] To the extent that the β-turn percentages correlate with inhibition rates, these peptides should have inhibiting properties.

We now propose a protocol (Figure 8) to predict *ab initio* structures of larger peptides in the gas phase and in solution based on the error profile of the methods explored here (See Supporting Information, Table S1, for computational cost of each method). To obtain the lowest energy conformers, begin with a large number of configurations (120 initial structures for the 4-mers studied in this paper) from an REMD trajectory or another technique that assures sufficient sampling. This is essential for producing a range of structures that encompass as much of the eventual quantum potential energy surface as possible. Once the structures have been optimized using one of the dispersion and/or hydrogen bonding corrected SE methods (PM6-D2, PM6-H2, PM6-DH+, or PM6-DH2), select the lowest set of conformers based on their RI-MP2/VDZ//X conformational energies, ΔE[RI-MP2/VDZ//X]. We illustrated Figure 8 with PM6-H2, but the other three corrected semiempirical methods generally perform comparably. Selecting the conformers with energies above a certain energy cutoff (E_{cutoff}) using SE methods reduces the required number of more expensive DFT and ab initio calculations to a small set of structures. The energy cutoff (E_{cutoff}) to use is the sum of the desired energy range (E_{range}), the mean signed errors (μ), and twice the standard deviation (σ):

$$E_{cutoff} = E_{range} + |\mu| + 2|\sigma| \quad (4)$$

Assuming that the error distribution is normal, there is a 95% chance of capturing all structures within the desired E_{range} within this E_{cutoff}. Based on the RI-MP2/VDZ//PM6-H2 values for the 30 lowest energy structures (μ_{30} and σ_{30}), the E_{cutoff} to capture isomers within 2 kcal/mol range would be 2.00 + 1.52 + 2 × 2.44 = 8.40 kcal/mol in gas phase and 2.00 + 0.18 + 2 × 3.00 = 8.18 kcal/mol in aqueous solution.

The next step is a DFT optimization. Clearly the DFT method must be able to include dispersion, and the best method to use will depend on the size of the peptide and the accuracy and speed of the latest density functionals.[145−147] For TPVN, PBE0-D and M06-2X perform well, with their μ and σ values being largely less than 1 kcal/mol. To determine the energy cutoff (E_{cutoff}) necessary to capture all conformers within a range (E_{range}) of 2 kcal/mol, we can again apply eq 4. With RI-MP2/VDZ//PBE0-D/VDZ, the cutoff (E_{cutoff}) would be 2.00 + 0.97 + 2 × 1.09 = 5.15 kcal/mol in the gas phase and 2.00 + 1.06 + 2 × 0.85 = 4.76 kcal/mol in aqueous solution. Using the energy cutoff first for RI-MP2/VDZ//PM6-H2 energies and then for RI-MP2/VDZ//PBE0-D/VDZ on a much smaller set of configurations should yield an ensemble of low-energy structures within the desired energy range. To finalize the procedure, we can optimize the structures within the desired energy range (E_{range}) using RI-MP2/VDZ to obtain a final ensemble of the lowest energy conformers.

The protocol is intended for AFPeps, particularly those we have studied for their anticancer properties previously.[14,15,17] The current semiempirical, density functional, and ab initio work confirms that the ensemble of low-energy conformers of these peptides do indeed have β-turn structures as predicted from our previous REMD simulations.[14,15] The protocol should presumably work for other small peptides regardless of their conformation or composition; however, testing and validating it for a large set of peptides is beyond the scope of this work. The most challenging aspect of obtaining the structure of peptides is capturing the long-range interactions that determine secondary structure. Our protocol includes dispersion and hydrogen bonding corrections at the semiempircal (PM6-H2) and density functional (PBE0-D) levels, while MP2 inherently includes these interactions. Even though none of the amino acid residues of TPVN are charged at the physiological pH of interest for drug design, the protocol should presumably work for peptides containing amino acids with different side chain properties. In our Supporting Information, we have complemented the TPVN results reported in the main manuscript with similar data for TOVN and PVNP.

CONCLUSION

We have performed a thorough computational investigation of the TPVN four amino acid peptide, which encompasses the pharmacophore region for AFP and AFP-like peptides, starting from REMD simulations. These inexpensive simulations using AMBER's ff99SB force field have shown that the anticancer activity of the small AFPeps correlates with their ability to form a β-turn motif. We extracted 120 structures from REMD simulations, which were subsequently minimized using more reliable semiempirical, density functional, and wave function methods. The more rigorous computational methods conclusively determined that the most stable conformers adopt a β-turn motif both in the gas phase and in solution. The agreement between simple and rigorous methods for determining the structures of TPVN suggests that the correlation between the presence of β-turn conformation and anti-breast-cancer activity holds true for larger peptides. Even though the mechanism of action is not known, our findings do provide a clear structure−activity relationship that may lead to the design of a drug based on this pharmacophore.

We used semiempirical, DFT, and ab initio methods to produce an ensemble of low-energy RI-MP2/VDZ structures. By using continuum calculations to simulate solvent, peptide structures in both the gas and aqueous phases were deduced. By comparing the accuracy of these methods against RI-MP2 benchmarks, we devised a protocol for calculating the lowest energy conformers of larger peptides accurately and efficiently. This protocol screens out high-energy conformers using lower levels of theory and predicts accurate peptide structures for the most stable conformers. This multilevel methodology could be a useful tool to complement experimental studies using X-ray crystallography, NMR, and mass spectrometry.

ASSOCIATED CONTENT

Supporting Information

The RI-MP2/VDZ//X (X = PM3, PM6, PM6-D2, PM6-H2, PM6-DH+, PM6-DH2, TPSS-D, PBE0-D, or M06-2X) global minimum coordinates in PDB format, a comparison of timing to perform semiempirical, DFT, and ab initio gradient calculations, mean unsigned errors (MUEs), and standard deviations (STDEVs). This material is available free of charge via the Internet at http://pubs.acs.org/.

dx.doi.org/10.1021/jp500017b | J. Phys. Chem. B 2014, 118, 4514−4526

AUTHOR INFORMATION

Corresponding Author

*E-mail: george.shields@bucknell.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Acknowledgment is made to the NSF and Bucknell University for their support of this work. This project was supported in part by NSF Grants CHE-0848827 and CHE-1213521 and by NSF Grants CHE-0116435, CHE-0521063, CHE-0849677, and CHE-1229354 as part of the MERCURY high-performance computer consortium (http://www.mercuryconsortium.org). This research used the National Science Foundation TeraGrid resources provided by the Texas Advanced Computing Center (TACC) under Grant Numbers TG-CHE090095 and TG-CHE120025. This research also used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

REFERENCES

(1) Leclercq, G.; Gallo, D.; Cossy, J.; Laios, I.; Larsimont, D.; Laurent, G.; Jacquot, Y. Peptides Targeting Estrogen Receptor Alpha-Potential Applications for Breast Cancer Treatment. *Curr. Pharm. Des.* **2011**, *17*, 2632−2653.

(2) Rosales-Hernandez, M. C.; Bermudez-Lugo, J.; Garcia, J.; Trujillo-Ferrara, J.; Correa-Basurto, J. Molecular Modeling Applied to Anti-Cancer Drug Development. *Anticancer Agents Med. Chem.* **2009**, *9*, 230−238.

(3) Jacobson, H.; Bennett, J.; Mizejewski, G. Inhibition of Estrogen-Dependent Breast-Cancer Growth by a Reaction-Product of Alpha-Fetoprotein and Estradiol. *Cancer Res.* **1990**, *50*, 415−420.

(4) Mizejewski, G. J.; Dias, J. A.; Hauer, C. R.; Henrikson, K. P.; Gierthy, J. Alpha-fetoprotein derived synthetic peptides: Assay of an estrogen-modifying regulatory segment. *Mol. Cell. Endocrinol.* **1996**, *118*, 15−23.

(5) Eisele, L. E.; Mesfin, F. B.; Bennett, J. A.; Andersen, T. T.; Jacobson, H. I.; Soldwedel, H.; MacColl, R.; Mizejewski, G. J. Studies on a growth-inhibitory peptide derived from alpha-fetoprotein and some analogs. *J. Pept. Res.* **2001**, *57*, 29−38.

(6) Mizejewski, G. J.; Butterstein, G. Survey of functional activities of alpha-fetoprotein derived growth inhibitory peptides: Review and prospects. *Curr. Protein Pept. Sci.* **2006**, *7*, 73−100.

(7) Mizejewski, G. J.; Eisele, L.; Maccoll, R. Anticancer versus antigrowth activities of three analogs of the growth-inhibitory peptide: Relevance to physicochemical properties. *Anticancer Res.* **2006**, *26*, 3071−3076.

(8) Sierralta, W. D.; Epunan, M. J.; Reyes, J. M.; Valladares, L. E.; Andersen, T. T.; Bennett, J. A.; Jacobson, H. I.; Pino, A. M. A peptide derived from alpha-fetoprotein inhibits the proliferation induced by estradiol in mammary tumor cells in culture. *Oncol. Rep.* **2008**, *19*, 229−235.

(9) Mizejewski, G. Mechanism of Cancer Growth Suppression of Alpha-Fetoprotein Derived Growth Inhibitory Peptides (GIP): Comparison of GIP-34 versus GIP-8 (AFPep). Updates and Prospects. *Cancers* **2011**, *3*, 2709−2733.

(10) Richardson, B. E.; Hulka, B. S.; Peck, J. L. D.; Hughes, C. L.; van den Berg, B. J.; Christianson, R. E.; Calvin, J. A. Levels of maternal serum alpha-fetoprotein (AFP) in pregnant women and subsequent breast cancer risk. *Am. J. Epidemiol.* **1998**, *148*, 719−727.

(11) Mesfin, F. B.; Bennett, J. A.; Jacobson, H. I.; Zhu, S. J.; Andersen, T. T. Alpha-fetoprotein-derived antiestrotrophic octapeptide. *Biochim. Biophys. Acta* **2000**, *1501*, 33−43.

(12) Bennett, J. A.; Mesfin, F. B.; Andersen, T. T.; Gierthy, J. F.; Jacobson, H. I. A peptide derived from alpha-fetoprotein prevents the growth of estrogen-dependent human breast cancers sensitive and resistant to tamoxifen. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2211−2215.

(13) DeFreest, L. A.; Mesfin, F. B.; Joseph, L.; McLeod, D. J.; Stallmer, A.; Reddy, S.; Balulad, S. S.; Jacobson, H. I.; Andersen, T. T.; Bennett, J. A. Synthetic peptide derived from alpha-fetoprotein inhibits growth of human breast cancer: investigation of the pharmacophore and synthesis optimization. *J. Pept. Res.* **2004**, *63*, 409−419.

(14) Kirschner, K. N.; Lexa, K. W.; Salisburg, A. M.; Alser, K. A.; Joseph, L.; Andersen, T. T.; Bennett, J. A.; Jacobson, H. I.; Shields, G. C. Computational design and experimental discovery of an antiestrogenic peptide derived from alpha-fetoprotein. *J. Am. Chem. Soc.* **2007**, *129*, 6263−6268.

(15) Lexa, K. W.; Alser, K. A.; Salisburg, A. M.; Ellens, D. J.; Hernandez, L.; Bono, S. J.; Michael, H. C.; Derby, J. R.; Skiba, J. G.; Feldgus, S.; et al. The search for low energy conformational families of small peptides: Searching for active conformations of small peptides in the absence of a known receptor. *Int. J. Quantum Chem.* **2007**, *107*, 3001−3012.

(16) Joseph, L. C.; Bennett, J. A.; Kirschner, K. N.; Shields, G. C.; Hughes, J.; Lostritto, N.; Jacobson, H. I.; Andersen, T. T. Antiestrogenic and anticancer activities of peptides derived from the active site of alpha-fetoprotein. *J. Pept. Sci.* **2009**, *15*, 319−325.

(17) Shields, G. C. Computational approaches for the design of peptides with anti-breast cancer properties. *Future Med. Chem.* **2009**, *1*, 201−212.

(18) Schlick, T.; Collepardo-Guevara, R.; Halvorsen, L. A.; Jung, S.; Xiao, X. Biomolecular modeling and simulation: a field coming of age. *Q. Rev. Biophys.* **2011**, *44*, 191−228.

(19) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons, Ltd: New York, 2001.

(20) Rienstra, C. M.; Tucker-Kellogg, L.; Jaroniec, C. P.; Hohwy, M.; Reif, B.; McMahon, M. T.; Tidor, B.; Lozano-Perez, T.; Griffin, R. G. De novo determination of peptide structure with solid-state magic-angle spinning NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10260−10265.

(21) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **1958**, *181*, 662−666.

(22) Schultz, S. C.; Shields, G. C.; Steitz, T. A. Crystallization of *Escherichia coli* Catabolite Gene Activator Protein with its DNA Binding Site: The Use of Modular DNA. *J. Mol. Biol.* **1990**, *213*, 159−166.

(23) Schultz, S. C.; Shields, G. C.; Steitz, T. A. Crystal Structure of a CAP-DNA Complex: The DNA Is Bent by 90 Degrees. *Science* **1991**, *253*, 1001−1007.

(24) Schroder, G. F.; Levitt, M.; Brunger, A. T. Super-resolution biomolecular crystallography with low-resolution data. *Nature* **2010**, *464*, 1218−1222.

(25) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989**, *246*, 64−71.

(26) Yates, J. R. A century of mass spectrometry: From atoms to proteomes. *Nat. Methods* **2011**, *8*, 633−637.

(27) Cramer, C. J.; Truhlar, D. G. AM1-SM2 AND PM3-SM3 parametrized SCF solvation models for free energies in aqueous solution. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 629−666.

(28) Klamt, A.; Schuurmann, G. COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799−805.

(29) Andzelm, J.; Kolmel, C.; Klamt, A. Incorporation of solvent effects into density functional calculations of molecular energies and geometries. *J. Chem. Phys.* **1995**, *103*, 9312−9320.

(30) Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. Ab initio study of solvated molecules: A new implementation of the polarizable continuum model. *Chem. Phys. Lett.* **1996**, *255*, 327−335.

(31) Cances, E.; Mennucci, B.; Tomasi, J. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys.* **1997**, *107*, 3032−3041.
(32) Barone, V.; Cossi, M.; Tomasi, J. A new definition of cavities for the computation of solvation free energies by the polarizable continuum model. *J. Chem. Phys.* **1997**, *107*, 3210−3221.
(33) Barone, V.; Cossi, M.; Tomasi, J. Geometry optimization of molecular structures in solution by the polarizable continuum model. *J. Comput. Chem.* **1998**, *19*, 404−417.
(34) Cramer, C. J.; Truhlar, D. G. Implicit solvation models: Equilibria, structure, spectra, and dynamics. *Chem. Rev.* **1999**, *99*, 2161−2200.
(35) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. New universal solvation model and comparison of the accuracy of the SM5.42R, SM5.43R, C-PCM, D-PCM, and IEF-PCM continuum solvation models for aqueous and organic solvation free energies and for vapor pressures. *J. Phys. Chem. A* **2004**, *108*, 6532−6542.
(36) Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd ed.; Wiley: Chichester, U.K., 2004.
(37) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. SM6: A density functional theory continuum solvation model for calculating aqueous solvation free energies of neutrals, ions, and solute-water clusters. *J. Chem. Theory Comput.* **2005**, *1*, 1133−1152.
(38) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999−3093.
(39) Caricato, M.; Ingrosso, F.; Mennucci, B.; Tomasi, J. A time-dependent polarizable continuum model: Theory and application. *J. Chem. Phys.* **2005**, *122*, No. 154501.
(40) Curutchet, C.; Orozco, M.; Luque, F.; Mennucci, B.; Tomasi, J. Dispersion and repulsion contributions to the solvation free energy: Comparison of quantum mechanical and classical approaches in the polarizable continuum model. *J. Comput. Chem.* **2006**, *27*, 1769−1780.
(41) Iozzi, M.; Cossi, M.; Improta, R.; Rega, N.; Barone, V. A polarizable continuum approach for the study of heterogeneous dielectric environments. *J. Chem. Phys.* **2006**, *124*, No. 184103.
(42) Cramer, C. J.; Truhlar, D. G. A universal approach to solvation modeling. *Acc. Chem. Res.* **2008**, *41*, 760−768.
(43) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on the Generalized Born Approximation with Asymmetric Descreening. *J. Chem. Theory Comput.* **2009**, *5*, 2447−2464.
(44) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378−6396.
(45) Sherer, E. C.; Yang, G.; Turner, G. M.; Shields, G. C.; Landry, D. W. Comparison of experimental and theoretical structures of a transition state analogue used for the induction of anti- cocaine catalytic antibodies. *J. Phys. Chem. A* **1997**, *101*, 8526−8529.
(46) Shields, G. C.; Laughton, C. A.; Orozco, M. Molecular dynamics simulations of the d(T·A·T) triple helix. *J. Am. Chem. Soc.* **1997**, *119*, 7463−7469.
(47) Shields, G. C.; Laughton, C. A.; Orozco, M. Molecular dynamics simulation of a PNA·DNA·PNA triple helix in aqueous solution. *J. Am. Chem. Soc.* **1998**, *120*, 5895−5904.
(48) Sherer, E. C.; Bono, S. J.; Shields, G. C. Further Quantum Mechanical Evidence that Difluorotoluene does not Hydrogen Bond. *J. Phys. Chem. B* **2001**, *105*, 8445−8451.
(49) Shea, J. E.; Brooks, C. L., III from Folding Theories to Folding Proteins: A Review and Assessment of Simulation Studies of Protein Folding and Unfolding. *Annu. Rev. Phys. Chem.* **2001**, *52*, 499−535.
(50) Gogonea, V.; Westerhoff, L. M.; Merz, K. M. Quantum mechanical/quantum mechanical methods. I. A divide and conquer strategy for solving the Schrodinger equation for large molecular systems using a composite density functional-semiempirical Hamiltonian. *J. Chem. Phys.* **2000**, *113*, 5604−5613.
(51) Van der Vaart, A.; Gogonea, V.; Dixon, S. L.; Merz, K. M. Linear scaling molecular orbital calculations of biological systems using the semiempirical divide and conquer method. *J. Comput. Chem.* **2000**, *21*, 1494−1504.
(52) van der Vaart, A.; Suarez, D.; Merz, K. M. Critical assessment of the performance of the semiempirical divide and conquer method for single point calculations and geometry optimizations of large chemical systems. *J. Chem. Phys.* **2000**, *113*, 10512−10523.
(53) Simmerling, C.; Strockbine, B.; Roitberg, A. E. All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc.* **2002**, *124*, 11258−11259.
(54) Estiu, G.; Merz, K. M. Enzymatic catalysis of urea decomposition: Elimination or hydrolysis? *J. Am. Chem. Soc.* **2004**, *126*, 11832−11842.
(55) Crespo, A.; Marti, M. A.; Estrin, D. A.; Roitberg, A. E. Multiple-steering QM-MM calculation of the free energy profile in chorismate mutase. *J. Am. Chem. Soc.* **2005**, *127*, 6940−6941.
(56) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712−725.
(57) Hornak, V.; Okur, A.; Rizzo, R. C.; Simmerling, C. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 915−920.
(58) Estiu, G.; Suarez, D.; Merz, K. M., Jr. Quantum mechanical and molecular dynamics simulations of ureases and Zn beta-lactamases. *J. Comput. Chem.* **2006**, *27*, 1240−1262.
(59) Wollacott, A. M.; Merz, K. M., Jr. Assessment of semiempirical quantum mechanical methods for the evaluation of protein structures. *J. Chem. Theory Comput.* **2007**, *3*, 1609−1619.
(60) Wickstrom, L.; Bi, Y.; Hornak, V.; Raleigh, D. P.; Simmerling, C. Reconciling the solution and X-ray structures of the villin headpiece helical subdomain: Molecular dynamics simulations and double mutant cycles reveal a stabilizing cation-pi interaction. *Biochemistry* **2007**, *46*, 3624−3634.
(61) Ding, F.; Layten, M.; Simmerling, C. Solution structure of HIV-1 protease flaps probed by comparison of molecular dynamics simulation ensembles and EPR experiments. *J. Am. Chem. Soc.* **2008**, *130*, 7184−7185.
(62) Okur, A.; Wickstrom, L.; Simmerling, C. Evaluation of salt bridge structure and energetics in peptides using explicit, implicit, and hybrid solvation models. *J. Chem. Theory Comput.* **2008**, *4*, 488−498.
(63) Mart, M. A.; Lebrero, M. C. G.; Roitberg, A. E.; Estrin, D. A. Bond or cage effect: How nitrophorins transport and release nitric oxide. *J. Am. Chem. Soc.* **2008**, *130*, 1611−1618.
(64) Demir, O.; Roitberg, A. E. Modulation of Catalytic Function by Differential Plasticity of the Active Site: Case Study of Trypanosoma cruzi trans-Sialidase and Trypanosoma rangeli Sialidase. *Biochemistry* **2009**, *48*, 3398−3406.
(65) Seabra, G. d. M.; Walker, R. C.; Roitberg, A. E. Are Current Semiempirical Methods Better Than Force Fields? A Study from the Thermodynamics Perspective. *J. Phys. Chem. A* **2009**, *113*, 11938−11948.
(66) Wang, X. S.; Roitberg, A. E.; Richards, N. G. J. Computational Studies of Ammonia Channel Function in Glutamine 5′-Phosphoribosylpyrophosphate Amidotransferase. *Biochemistry* **2009**, *48*, 12272−12282.
(67) He, X.; Fusti-Molnar, L.; Cui, G.; Merz, K. M., Jr. Importance of Dispersion and Electron Correlation in ab Initio Protein Folding. *J. Phys. Chem. B* **2009**, *113*, 5290−5300.
(68) Galiano, L.; Ding, F.; Veloro, A. M.; Blackburn, M. E.; Simmerling, C.; Fanucci, G. E. Drug Pressure Selected Mutations in HIV-1 Protease Alter Flap Conformations. *J. Am. Chem. Soc.* **2009**, *131*, 430−431.
(69) Sun, L.; Simmerling, C.; Ojima, I. Recent Advances in the Study of the Bioactive Conformation of Taxol. *ChemMedChem.* **2009**, *4*, 719−731.
(70) Sindhikara, D. J.; Roitberg, A. E.; Merz, K. M., Jr. Apo and Nickel-Bound Forms of the *Pyrococcus horikoshii* Species of the

Metalloregulatory Protein: NikR Characterized by Molecular Dynamics Simulations. *Biochemistry* **2009**, *48*, 12024−12033.
(71) He, X.; Merz, K. M., Jr. Divide and Conquer Hartree-Fock Calculations on Proteins. *J. Chem. Theory Comput.* **2010**, *6*, 405−411.
(72) Faver, J. C.; Benson, M. L.; He, X.; Roberts, B. P.; Wang, B.; Marshall, M. S.; Sherrill, C. D.; Merz, K. M. The Energy Computation Paradox and ab initio Protein Folding. *PLoS One* **2011**, *6*, No. e18868.
(73) Rodriguez Limardo, R. G.; Ferreiro, D. N.; Roitberg, A. E.; Marti, M. A.; Turjanski, A. G. p38 gamma Activation Triggers Dynamical Changes in Allosteric Docking Sites. *Biochemistry* **2011**, *50*, 1384−1395.
(74) Bergonzo, C.; Campbell, A. J.; de los Santos, C.; Grollman, A. P.; Simmerling, C. Energetic Preference of 8-oxoG Eversion Pathways in a DNA Glycosylase. *J. Am. Chem. Soc.* **2011**, *133*, 14504−14506.
(75) Pierdominici-Sottile, G.; Horenstein, N. A.; Roitberg, A. E. Free Energy Study of the Catalytic Mechanism of Trypanosoma cruzi trans-Sialidase. From the Michaelis Complex to the Covalent Intermediate. *Biochemistry* **2011**, *50*, 10150−10158.
(76) Chakravorty, D. K.; Wang, B.; Lee, C. W.; Giedroc, D. P.; Merz, K. M., Jr. Simulations of Allosteric Motions in the Zinc Sensor CzrA. *J. Am. Chem. Soc.* **2012**, *134*, 3367−3376.
(77) Shang, Y.; Simmerling, C. Molecular Dynamics Applied in Drug Discovery: The Case of HIV-1 Protease. In *Computational Drug Discovery and Design*; Baron, R., Ed.; Methods in Molecular Biology; Humana Press Inc: Totowa, NJ, 2012; Vol. *819*; pp 527−549.
(78) Fu, Z.; Li, X.; Merz, K. M. Conformational Analysis of Free and Bound Retinoic Acid. *J. Chem. Theory Comput.* **2012**, *8*, 1436−1448.
(79) Roberts, B. P.; Miller, B. R.; Roitberg, A. E.; Merz, K. M. Wide-Open Flaps Are Key to Urease Activity. *J. Am. Chem. Soc.* **2012**, *134*, 9934−9937.
(80) Boechi, L.; Arrar, M.; Marti, M. A.; Olson, J. S.; Roitberg, A. E.; Estrin, D. A. Hydrophobic Effect Drives Oxygen Uptake in Myoglobin via Histidine E7. *J. Biol. Chem.* **2013**, *288*, 6754−6762.
(81) Swails, J. M.; Roitberg, A. E. Enhancing Conformation and Protonation State Sampling of Hen Egg White Lysozyme Using pH Replica Exchange Molecular Dynamics. *J. Chem. Theory Comput.* **2012**, *8*, 4393−4404.
(82) Mulholland, A. J.; Roitberg, A. E.; Tuñón, I. Enzyme dynamics and catalysis in the mechanism of DNA polymerase. *Theor. Chem. Acc.* **2012**, *131*, 1−4.
(83) Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314−3321.
(84) Di Russo, N. V.; Estrin, D. A.; Marti, M. A.; Roitberg, A. E. pH-Dependent Conformational Changes in Proteins and Their Effect on Experimental pK_as: The Case of Nitrophorin 4. *PLoS Comput. Biol.* **2012**, *8*, No. e1002761.
(85) Dashti, D. S.; Meng, Y. L.; Roitberg, A. E. pH-Replica Exchange Molecular Dynamics in Proteins Using a Discrete Protonation Method. *J. Phys. Chem. B* **2012**, *116*, 8805−8811.
(86) Chakravorty, D. K.; Parker, T. M.; Guerra, A. J.; Sherrill, C. D.; Giedroc, D. P.; Merz, K. M. Energetics of Zinc-Mediated Interactions in the Allosteric Pathways of Metal Sensor Proteins. *J. Am. Chem. Soc.* **2013**, *135*, 30−33.
(87) Fu, Z.; Li, X.; Miao, Y. P.; Merz, K. M. Conformational Analysis and Parallel QM/MM X-ray Refinement of Protein Bound Anti-Alzheimer Drug Donepezil. *J. Chem. Theory Comput.* **2013**, *9*, 1686−1693.
(88) Daura, X.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Reversible peptide folding in solution by molecular dynamics simulation. *J. Mol. Biol.* **1998**, *280*, 925−932.
(89) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide Folding: When Simulation Meets Experiment. *Angew. Chem., Int. Ed.* **1999**, *38*, 236−240.
(90) Simmerling, C.; Strockbine, B.; Roitberg, A. E. All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc.* **2002**, *124*, 11258−11259.
(91) Qiu, L. L.; Pabit, S. A.; Roitberg, A. E.; Hagen, S. J. Smaller and faster: The 20-residue Trp-cage protein folds in 4 mu s. *J. Am. Chem. Soc.* **2002**, *124*, 12952−12953.
(92) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(93) Onufriev, A.; Bashford, D.; Case, D. A. Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem. B* **2000**, *104*, 3712−3720.
(94) Bernholdt, D. E.; Harrison, R. J. Large-scale correlated electronic structure calculations: The RI-MP2 method on parallel computers. *Chem. Phys. Lett.* **1996**, *250*, 477−484.
(95) Weigend, F.; Haser, M. RI-MP2: First derivatives and global consistency. *Theor. Chem. Acc.* **1997**, *97*, 331−340.
(96) Werner, H. J.; Manby, F. R.; Knowles, P. J. Fast linear scaling second-order Møller-Plesset perturbation theory (MP2) using local and density fitting approximations. *J. Chem. Phys.* **2003**, *118*, 8149−8160.
(97) Marshall, M. S.; Sears, J. S.; Burns, L. A.; Bredas, J. L.; Sherrill, C. D. An Error and Efficiency Analysis of Approximations to Møller-Plesset Perturbation Theory. *J. Chem. Theory Comput.* **2010**, *6*, 3681−3687.
(98) Temelso, B.; Archer, K. A.; Shields, G. C. Benchmark Structures and Binding Energies of Small Water Clusters with Anharmonicity Corrections. *J. Phys. Chem. A* **2011**, *115*, 12034−12046.
(99) Pérez, C.; Muckle, M. T.; Zaleski, D. P.; Seifert, N. A.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. Structures of Cage, Prism, and Book Isomers of Water Hexamer from Broadband Rotational Spectroscopy. *Science* **2012**, *336*, 897−901.
(100) Pérez, C.; Lobsiger, S.; Seifert, N. A.; Zaleski, D. P.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. Broadband Fourier transform rotational spectroscopy for structure determination: The water heptamer. *Chem. Phys. Lett.* **2013**, *571*, 1−15.
(101) Stewart, J. J. P. Optimization of parameters for semiempirical methods. I. Method. *J. Comput. Chem.* **1989**, *10*, 209−220; Optimization of parameters for semiempirical methods. II. Applications. *J. Comput. Chem.* **1989**, *10*, 221−264.
(102) Stewart, J. J. P. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* **2007**, *13*, 1173−1213.
(103) Stewart, J. J. P. Application of the PM6 method to modeling proteins. *J. Mol. Model.* **2009**, *15*, 765−805.
(104) Rezac, J.; Fanfrlik, J.; Salahub, D.; Hobza, P. Semiempirical Quantum Chemical PM6Method Augmented by Dispersion and H-Bonding Correction Terms Reliably Describes Various Types of Noncovalent Complexes. *J. Chem. Theory Comput.* **2009**, *5*, 1749−1760.
(105) Korth, M.; Pitonak, M.; Rezac, J.; Hobza, P. A Transferable H-Bonding Correction for Semiempirical Quantum-Chemical Methods. *J. Chem. Theory Comput.* **2010**, *6*, 344−352.
(106) Jurema, M. W.; Shields, G. C. Ability of the PM3 Quantum Mechanical Method to Model *Inter*molecular Hydrogen Bonding between Neutral Molecules. *J. Comput. Chem.* **1993**, *14*, 89−104.
(107) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(108) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91*, No. 146401.
(109) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158−6170.
(110) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.

(111) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree-Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree-Fock exchange. *Chem. Phys.* **2009**, *356*, 98−109.
(112) Kossmann, S.; Neese, F. Comparison of two efficient approximate Hartee-Fock approaches. *Chem. Phys. Lett.* **2009**, *481*, 240−243.
(113) Kossmann, S.; Neese, F. Efficient Structure Optimization with Second-Order Many-Body Perturbation Theory: The RIJCOSX-MP2Method. *J. Chem. Theory Comput.* **2010**, *6*, 2325−2338.
(114) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73−78.
(115) Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25*, 1463−1473.
(116) Jurečka, P.; Černý, J.; Hobza, P.; Salahub, D. R. Density functional theory augmented with an empirical dispersion term. Interaction energies and geometries of 80 noncovalent complexes compared with ab initio quantum mechanics calculations. *J. Comput. Chem.* **2007**, *28*, 555−569.
(117) Sherrill, C. D. Frontiers in electronic structure theory. *J. Chem. Phys.* **2010**, *132*, 7.
(118) Marom, N.; Tkatchenko, A.; Rossi, M.; Gobre, V. V.; Hod, O.; Scheffler, M.; Kronik, L. Dispersion Interactions with Density-Functional Theory: Benchmarking Semiempirical and Interatomic Pairwise Corrected Density Functionals. *J. Chem. Theory Comput.* **2011**, *7*, 3944−3951.
(119) Hujo, W.; Grimme, S. Performance of the van der Waals Density Functional VV10 and (hybrid)GGA Variants for Thermochemistry and Noncovalent Interactions. *J. Chem. Theory Comput.* **2011**, *7*, 3866−3871.
(120) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(121) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, No. 154104.
(122) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456−1465.
(123) Meissner, T.; Wentz, F. J. The complex dielectric constant of pure and sea water from microwave satellite observations. *IEEE Geosci. Remote Sens. Lett.* **2004**, *42*, 1836−1849.
(124) Stewart, J. J. P. MOPAC2012 Computational Chemistry, http://OpenMOPAC.net (2012). In *MOPAC*, 2009 ed.; Colorado Springs, CO, USA, 2008.
(125) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision D.01; Gaussian, Inc.: Wallingford, CT, 2009.
(126) Day, M. B.; Kirschner, K. N.; Shields, G. C. Global search for minimum energy (H2O)21H+ clusters, n=3−5. *J. Phys. Chem. A* **2005**, *109*, 6773−6778.
(127) Kirschner, K. N.; Hartt, G. M.; Evans, T. M.; Shields, G. C. In search of $CS_2(H_2O)$(n=1−4) clusters. *J. Chem. Phys.* **2007**, *126*, No. 154320.
(128) Hartt, G. M.; Shields, G. C.; Kirschner, K. N. Hydration of OCS with one to four water molecules in atmospheric and laboratory conditions. *J. Phys. Chem. A* **2008**, *112*, 4490−4495.
(129) Morrell, T. E.; Shields, G. C. Atmospheric Implications for Formation of Clusters of Ammonium and 1−10 Water Molecules. *J. Phys. Chem. A* **2010**, *114*, 4266−4271.
(130) Shields, R. M.; Temelso, B.; Archer, K. A.; Morrell, T. E.; Shields, G. C. Accurate Predictions of Water Cluster Formation, $(H_2O)_{n=2-10}$. *J. Phys. Chem. A* **2010**, *114*, 11725−11737.
(131) Husar, D. E.; Temelso, B.; Ashworth, A. L.; Shields, G. C. Hydration of the Bisulfate Ion: Atmospheric Implications. *J. Phys. Chem. A* **2012**, *116*, 5151−5163.
(132) Temelso, B.; Morrell, T. E.; Shields, R. M.; Allodi, M. A.; Wood, E. K.; Kirschner, K. N.; Castonguay, T. C.; Archer, K. A.; Shields, G. C. Quantum Mechanical Study of Sulfuric Acid Hydration: Atmospheric Implications. *J. Phys. Chem. A* **2012**, *116*, 2209−2224.
(133) Temelso, B.; Phan, T. N.; Shields, G. C. Computational Study of the Hydration of Sulfuric Acid Dimers: Implications for Acid Dissociation and Aerosol Formation. *J. Phys. Chem. A* **2012**, *116*, 9745−9758.
(134) Zheng, Y.; Merz, K. M., Jr. Study of hydrogen bonding interactions relevant to biomolecular structure and function. *J. Comput. Chem.* **1992**, *13*, 1151−1169.
(135) Jurema, M. W.; Kirschner, K. N.; Shields, G. C. Modeling of Magic Water Clusters $(H_2O)_{20}$ and $(H_2O)_{21}H^+$ with the PM3 Quantum-Mechanical Method. *J. Comput. Chem.* **1993**, *14*, 1326−1332.
(136) Kirschner, K. N.; Shields, G. C. Quantum-Mechanical Investigation of Large Water Clusters. *Int. J. Quantum Chem.* **1994**, 349−360.
(137) Lively, T. N.; Jurema, M. W.; Shields, G. C. Hydrogen Bonding of Nucleotide Base Pairs: Application of the PM3Method. *Int. J. Quantum Chem.* **1994**, 95−107.
(138) Sherer, E. C.; York, D. M.; Cramer, C. J. Fast approximate methods for calculating nucleic acid base pair interaction energies. *J. Comput. Chem.* **2003**, *24*, 57−67.
(139) Beccara, S. A.; Faccioli, P.; Sega, M.; Pederiva, F.; Garberoglio, G.; Orland, H. Dominant folding pathways of a peptide chain from ab initio quantum-mechanical simulations. *J. Chem. Phys.* **2011**, *134*, No. 024501.
(140) Valdes, H.; Pluháčková, K.; Pitonák, M.; Rezáč, J.; Hobza, P. Benchmark database on isolated small peptides containing an aromatic side chain: Comparison between wave function and density functional theory methods and empirical force field. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2747−2757.
(141) Riley, K. E.; Pitonak, M.; Jurecka, P.; Hobza, P. Stabilization and Structure Calculations for Noncovalent Interactions in Extended Molecular Systems Based on Wave Function and Density Functional Theories. *Chem. Rev.* **2010**, *110*, 5023−5063.
(142) Zhou, R.; Berne, B. J.; Germain, R. The free energy landscape for β hairpin folding in explicit water. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14931.
(143) Geney, R.; Layten, M.; Gomperts, R.; Hornak, V.; Simmerling, C. Investigation of salt bridge stability in a generalized born solvent model. *J. Chem. Theory Comput* **2006**, *2*, 115−127.
(144) Chou, K. Prediction of Tight Turns and Their Types in Proteins. *Anal. Biochem.* **2000**, *286*, 1−16.
(145) Burns, L. A.; Mayagoitia, A. V.; Sumpter, B. G.; Sherrill, C. D. Density-functional approaches to noncovalent interactions: A comparison of dispersion corrections (DFT-D), exchange-hole dipole moment (XDM) theory, and specialized functionals. *J. Chem. Phys.* **2011**, *134*, No. 084107.
(146) Hao, P.; Sun, J.; Xiao, B.; Ruzsinszky, A.; Csonka, G. I.; Tao, J.; Glindmeyer, S.; Perdew, J. P. Performance of meta-GGA Functionals on General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2012**, *9*, 355−363.
(147) Zhao, Y.; Truhlar, D. G. Density Functional Theory for Reaction Energies: Test of Meta and Hybrid Meta Functionals, Range-

Separated Functionals, and Other High-Performance Functionals. *J. Chem. Theory Comput.* **2011**, *7*, 669−676.