Comparison of some dispersion-corrected and traditional functionals as applied to peptides and conformations of cyclohexane derivatives

Mateusz Marianski,¹ Amparo Asensio,² and J. J. Dannenberg^{1,a)}¹Department of Chemistry, City University of New York – Hunter College and the Graduate School, 695 Park Avenue, New York, New York 10065, USA²Facultad de Farmacia, Departamento Química Orgánica, Universidad de Valencia, Avda Vicente Andres Estelles, 46100-Burjassot (Valencia), Spain

(Received 13 February 2012; accepted 3 July 2012; published online 25 July 2012)

We compare the energetic and structural properties of fully optimized α -helical and antiparallel β -sheet polyanilines and the energetic differences between axial and equatorial conformations of three cyclohexane derivatives (methyl, fluoro, and chloro) as calculated using several functionals designed to treat dispersion (B97-D, ω B97X-D, M06, M06L, and M06-2X) with other traditional functionals not specifically parametrized to treat dispersion (B3LYP, X3LYP, and PBE1PBE) and with experimental results. Those functionals developed to treat dispersion significantly overestimate interaction enthalpies of folding for the α -helix and predict unreasonable structures that contain Ramachandran ϕ and ψ and C = O...N H-bonding angles that are out of the bounds of databases compiled the β -sheets. These structures are consistent with overestimation of the interaction energies. For the cyclohexanes, these functionals overestimate the stabilities of the axial conformation, especially when used with smaller basis sets. Their performance improves when the basis set is improved from D95** to aug-cc-pVTZ (which would not be possible with systems as large as the peptides).
© 2012 American Institute of Physics. [<http://dx.doi.org/10.1063/1.4737517>]

I. INTRODUCTION

The importance of dispersion interactions in biochemical systems, such as peptide structures, has received much recent attention in the literature. Since the more popular older functionals (such as B3LYP) do not properly describe pure dispersion, several newer functionals have been developed with the specific intention of including the contributions of dispersion to non-bonded interactions by means of parametrization to high level calculations. The temptation to use these to calculate the properties of peptides is obvious.

In a recent study of the non-bonded dispersive interactions of rare gases and alkane dimers compared with the induction interactions of polar molecules with rare gases and hydrocarbons, we found that these newer functionals often overestimate the interactions when compared to high level *ab initio* calculations increasing in quality up to CCSD(T)/aug-cc-pV5Z.¹ We have also demonstrated the erratic performance of several of these functionals for the stacking vs H-bonding interactions of pyrimidine and *p*-benzoquinone,² as well as compared their behaviors H-bonding in the water dimer.³

In this paper, we compare several of these newer functionals with those previously used for the determination of the α -helical and β -sheet peptide secondary structures of capped polyanilines and available experimental results. As a complementary test to this study, we examine the relative energies of axial and equatorial conformations of some substituted cyclohexanes for which precise ΔH 's in the gas phase or non-interacting media have been recently reported.⁴⁻⁶ The steric

repulsive strain between the axial H's on carbons 3 and 5 and an axial substituent provides the generally accepted explanation for the preference of the equatorial conformations. If the newer functionals properly account for dispersion, they should better describe the enthalpy differences between axial and equatorial conformations than the older methods.

A. Computational details

All calculations were performed using the GAUSSIAN 09 suite of computer programs.⁷ All structures were optimized with respect to all internal coordinates. Harmonic frequencies were calculated for these optimizations to (a) confirm the existence of a minimum (no imaginary frequencies) and (b) obtain the gas phase enthalpies that correspond to these structures. Counterpoise (CP) corrections⁸ for the β -sheet structures as single point *a posteriori* calculations were performed as optimization on the CP-corrected surface⁹ is too costly. We considered the following dispersion-corrected functionals: B97-D,¹⁰ ω B97X-D,¹¹ M06, M06-2X, M06L¹² (which we shall collectively designate as DF's) and the following traditional functionals (TF's): B3LYP,¹³ PBE1PBE,¹⁴ X3LYP¹⁵ for comparisons. The DF's contain up to 39 fitted parameters used to reproduce the values in several data sets of intermolecular interactions. While X3LYP uses the neon dimer as a reference point, it is basically a small modification of B3LYP, with the values of the three parameters adjusted and one additional parameter that mixes in B88 and PW91 exchange functionals so we include it among the TF's. The default grids and convergence criteria of the GAUSSIAN 09 program were initially used. However, several of the

a)jdannenberg@gc.cuny.edu.

TABLE I. $\Delta H_{\text{Ax-Eq}}$'s (kcal/mol) for fluoro-, chloro-, and methylcyclohexanes and MSE and MUE as calculated using different methods and comparison with best experimental values.

Substituent X	$\Delta H_{\text{Ax-Eq}}$									
	HF	PBE1PBE	B3LYP	X3LYP	B97D	M06	M06-2X	M06L	ω B97X-D	Expt.
D95**										
F	0.16	0.37	0.32	0.30	0.21	0.01	0.24	−0.19	0.14	0.31 (Ref. 6)
Cl	0.72	0.40	0.60	0.55	0.20	−0.25	0.14	−0.56	0.12	0.38 (Ref. 4)
Me	2.45	2.09	2.16	2.13	1.27	1.81	1.54	1.43	1.45	2.03 (Ref. 5)
MSE	0.20	0.05	0.12	0.09	−0.35	−0.38	−0.27	−0.68	−0.34	
MUE	0.30	0.05	0.12	0.09	0.35	0.38	0.27	0.68	0.34	
aug-cc-pvtz										
F	0.29	0.43	0.34	0.32	0.27	0.05	0.30	−0.14	0.17	0.31 (Ref. 6)
Cl	1.07	0.78	0.90	0.87	0.49	−0.29	0.47	−0.93	0.43	0.38 (Ref. 4)
Me	2.65	2.36	2.46	2.44	1.58	1.74	1.92	1.46	1.62	2.03 (Ref. 5)
MSE	0.43	0.28	0.33	0.30	−0.13	−0.41	−0.01	−0.78	−0.17	
MUE	0.44	0.28	0.33	0.30	0.20	0.41	0.07	0.78	0.20	

new functionals required more refined grid – “ultrafine” or “99974” – to produce acceptable results. Re-optimization of the large peptide structures using denser grids became computationally too costly. We recomputed only those systems, which had imaginary frequencies using “ultrafine” grid. This procedure removed low-lying imaginary frequencies due to methyl rotations in the α -helix described with M06, methylcyclohexane with M06L and the β -sheet with M06-2X functionals. When ONIOM¹⁶ calculations were performed with these functionals, the low level AM1¹⁷ was used only for the methyl side chains. We¹ and others¹⁸ have previously noted convergence problems for cases involving weak interactions using some of the dispersion corrected functionals.

II. RESULTS

A. Conformations of substituted cyclohexanes

We compare the calculated and experimental values in Table I. As seen from the table, the DF's consistently predict the $\Delta H_{\text{Ax-Eq}}$'s for the axial conformations to be lower than the other methods and lower than the accurate experimental measurements for the methyl,⁵ fluoro,⁶ and chloro⁴ cyclohexanes in the gas phase and in xenon solutions when used with the D95** basis set. The PBE1PBE functional consistently give experimental enthalpy differences closest to the experimental reports. Both B3LYP and X3LYP overestimate the difference for chlorocyclohexane. M06L, M06, and/or M06-2X predict the opposite sign for the ΔH 's than those predicted by the other methods for the three cases where good experimental data exist. When the basis set is improved to aug-cc-pVTZ, we see that, except for four instances (two each with M06 and M06L), the larger basis set always leads to larger ΔH 's. The mean signed and mean unsigned errors (MSE's and MUE's) significantly decrease upon increasing the size of the basis set for B97D, ω B97x-D, and M06-2X, while they worsen for M06, M06L, B3LYP, X3LYP, and PBE1PBE. Thus, three of the newer functionals, B97D, ω B97x-D, and M06-2X give reasonable results when used with the larger (but not the smaller) basis set.

B. Peptides

1. α -helices

We have previously reported the differential folding energies,¹⁹ enthalpies,²⁰ and solvation free energies²¹ for polyalanine α -helices as calculated by B3LYP/AM1 using ONIOM and the SM5.2 solvation model. The differential values were defined as the change in stabilization of the helix compared to an extended strand upon the addition of one more alanine residue to the peptide. The ΔH 's that we reported were slightly more stabilizing than those determined by Kemp in experiments done subsequent to our report.²²

Here, we compare the ΔH /residue for folding from an extended strand to an α -helix (see Table II). We have chosen the 17-alanine helix for the comparison. The enthalpies were calculated using the estimates for basis set superposition error (BSSE) taken from chains of five H-bonds (six formamides) frozen in the geometries assumed in each of the three H-bonding chains of the optimized

TABLE II. ΔH /residue for the folding of acetyl(Ala)₁₇NHCH₃ from extended β -strand to α -helix.

	Reference
B3LYP	−0.8
X3LYP	−1.2
M06	−3.9
M06-2X	−4.2
B97-D	−3.9
ω -B97x-D	−4.3
PBE1PBE ^a	−1.3
Other theoretical	
ONIOM/B3LYP/AM1	−0.7
Thermodynamic Cycle	−0.8
Experimental	
Kemp 2006	−1.0
Baldwin 1991	−1.1
Makhatadze 2005	−0.9
Rialdi 1966 (poly Glu)	−1.1
Scheraga 1971 (Lys)	−1.2

^aThe optimized helix contains a small 3_{10} segment near the N-terminus.

TABLE III. Geometric parameters for the optimized structures of the acetyl(Ala)₁₇NH₂ α -helix. Only the values for the central residues (4th to 12th H-bonds) are shown to avoid end effects.

	Distances (Å)			Angles (°)		
	H–O	N–O	CH–O	C=O–H	O–H–N	C=O...N
X3LYP	2.01	3.01	2.86	151.2	164.5	155.8
B3LYP	2.03	3.03	2.90	151.0	164.6	155.5
M06	1.97	2.96	2.62	151.2	162.1	156.5
M06-2X	1.98	2.95	2.61	151.2	159.8	157.7
B97-D	2.04	3.00	2.59	146.8	155.7	154.0
ω B97x-D	1.94	2.92	2.64	149.5	161.4	155.3
PBE1PBE ^a	2.03	3.01	2.84	149.3	163.1	154.3
Exper. (Ref. 31)	2.06	2.99 $\pm$ 0.01		149 $\pm$ 7	157 $\pm$ 9	
Exper. (Ref. 32)		2.97 $\pm$ 0.01				153 $\pm$ 2

^aThe optimized helix contains a small 3₁₀ segment near the N-terminus which does not involve these H-bonds.

helices in a manner somewhat similar to that described in an earlier report.²⁰ Both B3LYP and X3LYP predict results that are reasonably close to the several experimental values for polyaniline,^{23–25} polyglutamic acid,²⁶ and polylysine,²⁷ and some previous predictive²⁸ theoretical values²⁰ available. While the experimental values refer to aqueous solution, we note that SM5.2²⁹ estimates of the differences in free energies of solvation for the α -helix and polyproline II structure of polyaniline are quite small.²¹ We have recently reported a more complete analysis of the aqueous solvation of polyaniline α -helices.³⁰ Clearly, the ΔH 's predicted using the DF's are three to four times too large.

We present various structural details for the helices in Table III and compare them with literature values for large numbers of α -helices taken from databases.^{31,32} We note that the database helices contain many different residues other than alanines. Also, the x-ray crystal structures used to create the databases cannot locate positions of the H's, which were estimated by assuming the N–H bonds to be all 1.0 Å long and normal valence angles.³¹ Thus, these details lower our confidence in the geometric parameters that involve the positions of the H's. Furthermore, the estimated errors in these angles

are significantly larger than those for the C = O...N angle, which can be measured directly.

An α -helix contains three H-bonding chains, whose ends do not interact with anything in these gas phase calculations. In solution or in proteins, they would certainly interact with the environment. We have, therefore, considered only the nine H-bonds nearest the center as the other six would each be at a terminus of one of the chains. As the most recent database study breaks down the helices by size, we have used the data for the helices containing 15 H-bonds.³² We further note that recent density functional theory (DFT) calculations of polyaniline α -helices indicate that the H-bonds shorten somewhat when optimized in a continuum solvent or when water is added to the ends or with combinations of both.³⁰ The H-bond lengths for α -helices seem in agreement with the values from the database (taking into account the shortening expected upon solution) for all functionals with the exception of ω B97x-D, which predicts much shorter H-bonds even before consideration of the solvation effect. Analysis of the angles poses more of a problem as the two database studies differ in tabulating the C = O...N or C = O...H angles.

Improving the basis set to aug-cc-pVTZ, as we did for the cyclohexane derivatives, would not be possible for systems this large. However, from the results on the cyclohexanes, one might expect the dispersion interactions to be further overestimated by a larger basis set.

2. β -sheets

We compared the performance normal and dispersion corrected functionals for the calculations of antiparallel β -sheets consisting of four strands of acetyl(Ala)₆NHCH₃. The calculations used both pure DFT methods and ONIOM where the high level is DFT and the low level AM1. Only the methyl side chains were calculated at the low level. Our recent report on several parallel and antiparallel polyaniline (and other) β -sheet peptides used ONIOM with B3LYP and AM1 as the high and low levels.³³ Table IV and Tables S1 and S2 of the supplementary information (SI) and Figures 1 and 2 and S1 of the SI present the relevant data.

TABLE IV. Interaction energy (ΔE_{int}) and distortion energy per interaction in 4A6 AP β -sheet using different functionals with and without ONIOM. "Optimized" indicates energy with respect to optimized strands; "frozen" indicates interaction with strands frozen in the geometry of the optimized sheet and the $\Delta E_{\text{distortion}}$ represents the energy required to distort the optimized strands to the geometries they assume in the sheet.

Method	ONIOM (DFT:AM1)						
	B3LYP	X3LYP	M06	M06-2X	PBE1PBE	B97-D	ω B97x-D
$\Delta E(\text{optimized})$	−41.5	−45.0	−50.0	−53.7	−43.0	−55.2	−54.2
$\Delta E(\text{frozen})$	−54.8	−58.5	−57.1	−60.1	−56.3	−63.4	−61.9
$\Delta E(\text{optimized with CP})$	−30.1	−33.2	−40.2	−44.9	−33.8	−44.2	−44.8
$\Delta E(\text{frozen with CP})$	−43.6	−46.7	−47.3	−51.4	−47.1	−52.3	−52.5
$\Delta E_{\text{distortion}}$	13.3	13.4	7.1	6.4	13.3	8.2	7.7
Full DFT							
$\Delta E(\text{optimized})$	−42.7	−46.3	−57.1	−58.7	−44.2	−61.2	−62.9
$\Delta E(\text{frozen})$	−53.2	−56.8	−64.2	−65.1	−54.3	−66.6	−70.2
$\Delta E(\text{optimized with CP})$	−30.1	−33.2	−46.1	−48.9	−34.0	−48.3	−52.4
$\Delta E(\text{frozen with CP})$	−40.5	−43.8	−53.2	−55.3	−44.0	−53.7	−59.7
$\Delta E_{\text{distortion}}$	10.4	10.5	7.1	6.4	10.0	5.4	7.3

FIG. 1. Three types of geometries for AP β -sheets classified by the calculations that give rise to them. For the ONIOM structures, the low level (AM1) is shown as wire frame. View from edge of sheets.

Table IV presents the energetics of the formation of the β -sheets from individual strands. The DFs predict the sheets to be much more stable than the TFs. The TFs predict CP-corrected inter-strand interactions of about 30–34 kcal/strand, while the DFs predict about 46–49 kcal/strand, about 50% more stabilization. These results parallel those noted above for the α -helices. Since the differential folding energies predicted by these functionals for the helices are much more than determined by experiment, we expect the similarly larger interactions in the β -sheets to be too large.

As seen from Figures 1 and 2, the structures of the sheets fall into three different categories. The calculation performed with the pure TFs converge to pleated sheet structures, such as found in standard textbooks. The ONIOM calculations with these functionals predict the sheets to be pleated, but they are also somewhat twisted as often observed from experiments.³² The calculations using the pure DFs predict sheets that are significantly more flat than normal pleated sheets. When used with ONIOM, three of them maintained this flattened structure while one (B97-D) predicted a more normal pleated sheet.

We compare the average Ramachandran dihedral angles (ϕ and ψ) for the two innermost (to avoid effects due to the

FIG. 2. Three types of geometries for AP β -sheets classified by the calculations that give rise to them. For the ONIOM structures, the low level (AM1) is shown as wire frame. View from end of sheets.

edges of the sheets) strands in Table S1. Creighton suggests standard values of the ϕ and ψ dihedrals for planar β -sheets to be -139° and 135° , respectively.³⁴ Bax studied the correlation between nuclear magnetic resonance (NMR) chemical shifts and the dihedrals for 126 residues in β -sheets taken from high quality crystal structures. The values for ϕ ($-114 \pm 32^\circ$) and ψ ($142 \pm 20^\circ$) indicate substantial variation for these angles.³⁵ Clearly, the TFs predict ϕ and ψ to be closer to those expected for β -sheets in general, and closer to those reported for these dihedrals ($\phi = -135^\circ$, $\psi = 150^\circ$) about alanine residues from a NMR determination of the structure of spider silk,³⁶ which is known to be a large antiparallel β -sheet consisting mostly of alanine and glycine residues. The DFs predict these dihedrals to be significantly larger. The ψ angles are close to high limit and the ϕ 's are out of the range of the experimental values found by Bax.³⁵

Other features of the β -sheets also suggest that the DFs predict unusual structures. As seen from Figure 2, the four DFs predict the methyl groups in adjacent sheets to tilt towards or away from each other, apparently in response to maximizing dispersion interactions between them. Antiparallel β -sheets contain cyclic H-bonding structures that contain 14 (large ring) or 10 (small ring) structures, which have significantly different interactions energies when isolated.³⁷ Figure S1 and Table S2 also show that the DFs predict closer C_α - C_α and C-H...H-C (labeled 4 and 5 in the table and figure) interactions in the large H-bonding rings at the expense of longer O...H and O...N interactions for the H-bonds (6, 7, and 8) than do the TFs. The C-H...O interactions do not differ significantly for those calculated by the DFs and TFs, although those calculated by the DFs tend to have slightly longer distances, especially when used with ONIOM. The reported experimental N...O distances for antiparallel β -sheets are 2.91 ± 0.14 ³¹ and 2.904 ± 0.005 ³² while that for sheets containing four strands is 2.903 ± 0.004 Å. These values are close to those calculated by all methods except ω B97x-D (see Table S2), which also predicted short H-bonds for the helices. Klebe, *et al.* report the C = O...N angle to be $153^\circ \mp$ from 0.2° to 5.28° depending upon the number of strands.³² The TF's and B97-D using ONIOM predict angles (150° – 154°) that agree with this report, while the TF's by themselves predict slightly large values (157° – 160°). On the other hand, the DFs predict larger angles (162° – 167°), which are beyond the range of the database (except for B97-D with ONIOM as noted). Baker and Hubbard only report H-bonding angles that center on the H, which are not reliable due to the uncertainty in the position of the H atoms.³¹

We know of no experimental evidence to support the peculiarities of the geometries predicted with the dispersion corrected functionals. In particular, we know of no structural evidence for the predicted tilting of the methyls, which would probably be the easiest of the geometric differences noted above to be determined experimentally. We do note that x-ray structures of proteins are not generally close to the resolution of small molecules and the positions of hydrogen atoms are not usually obtained directly (this would require very high resolution x-ray or neutron diffraction).

Naturally occurring β -sheets from globular proteins are generally somewhat twisted.^{34,38} While the sheets calculated

using the traditional functions and B97-D/ONIOM calculations (including those previously reported³³) optimize to twisted structures, none of the pure DFT calculations show this feature.

III. DISCUSSION

The ΔH_{Ax-Eq} 's for the three derivatives of cyclohexane chosen for study were selected because of the extreme accuracy of the measurements and the lack of solvent interactions. Clearly, noncovalent interactions occur between solutes and solvents as well as between the substituents and the axial H's (when the substituent is in the axial position). The same reports⁴⁻⁶ contain MP2 calculations using basis sets as large as 6-311+G(2p,2d) which are in reasonable agreement with the experimental observations.

Our results show that these ΔH 's increase upon improving the basis set from d95** to aug-cc-pVTZ for all functionals except M06 and M06L. An increase in basis set should have two opposing effects upon the ΔH 's: (1) reduction of the intramolecular BSSE should cause an increase; and (2) an increase in the calculated dispersion interaction should cause a decrease in ΔH . The first effect removes the non-physical attraction between the axial substituent and the axial H's at the 3- and 5-positions, while the second should increase the dispersion interaction as these interactions depend upon the overlap populations which become more flexible with larger basis sets. These effects are observed for dispersion interactions in small systems when *ab initio* calculations (i.e., MP2 or CCSD(T)) are used. They are not consistently followed by functionals, especially the newer ones designed to treat dispersion, as we report elsewhere.¹ The observation (see Table I) that the ΔH 's generally increase (except for M06 and M06L) upon basis set improvement suggests that removing BSSE produces the larger effect for most cases. The behavior of B97-D, ω B97x-D, and M06-2X all improve so that they provide better agreement with experiment than the older functionals. With the smaller basis set, all of the functionals designed to treat dispersion underestimate all of the ΔH 's, while the older functionals generally provided more accurate results possibly due to the cancellation of errors. From the foregoing, one would expect that using these newer functionals with the smaller basis sets that must be used for large systems, such as peptides would lead to exaggeration of dispersion interactions, as confirmed from our results on the peptides.

The calculations on the ΔH of folding per residue for polyalanine α -helices clearly indicate that these DFs grossly overestimate the magnitude of this interaction, at least when used with a moderate basis set (d95**). As the structure of an α -helix is reasonably rigid, one would not expect these energy differences to cause large changes in geometry and we noticed only a small difference in geometries for the helices. The observation that the experimental folding enthalpies for polylysine and polyglutamic acid do not differ significantly from that for polyalanine is consistent with a relatively rigid α -helix whose basic structure is not significantly perturbed by changes in the amino acid residue. A recent report of calculations of α -helices using M06-2X provides significant overes-

timations of the folding energies in agreement with what we report here.³⁹

For the β -sheets, we again see much larger interaction energies when we used the newer functionals. Unfortunately, we know of no useful experimental data for the corresponding ΔH_{int} 's. However, we assume that the errors should be similar to those encountered in the study of the α -helices. Unlike the α -helices, the β -sheets have much more geometric flexibility. One sees from Figure 1 that all of the dispersion adapted functionals produce sheets that are much closer to planarity and all but B97-D do the same when used with AM1 in ONIOM calculations. Table S2 shows that the DFs have close interactions between C α 's and H α 's (marked 4 and 5, see Figure S1) in the large H-bonding rings of the adjacent middle strands. The TFs and the ONIOM calculations (which treat the methyls at the AM1 level) provide results that better conform to geometric parameters obtained from crystallographic databases. The Ramachandran angles (Table S1), ϕ and ψ , are at or beyond the extremes of accepted values³⁵ for the newer functionals designed to treat dispersion but not for the older functionals or the ONIOM calculations.

We note that the interaction energies for the ONIOM calculations when used with the DFs are significantly (5–9 kcal/mol) less than for the calculations using the same pure functionals (without ONIOM), while the differences for these two kinds of calculations for the more TFs are quite small (~ 1 kcal/mol). Thus, the differences appear to be due to the methyl–methyl attractions between the strands. Figure 2 shows how the methyls tilt towards and away from each other when the dispersion adapted functionals are used without ONIOM. This behavior disappears when ONIOM is used with B97-D, but not with the other DFs.

Table S2 shows significant differences in the distances between the C's of the methyls in the two middle strands (marked 1 and 2 in Figure S1) calculated using the DFs and TFs. The distance between the methyls in the small ring (1) is slightly longer (0.2–0.6 Å) than for those in the large ring (2), using the DFs this difference increases significantly (2.0–2.2 Å). This suggests an overly attractive interaction between the CH₃'s on adjacent strands (see Figure 2). When ONIOM is used, the methyl-methyl distance in the large ring (2) becomes somewhat larger (0.9–1.1 Å) than for the small ring (1) for the TFs and for B97-D probably due to the twisted structure. For the other three DFs, the relative order of these distances reverse, with (1) greater than (2) by 2.4–2.6 Å. While the excessive attractions between the methyls disappear when they are calculated using AM1 (low level in ONIOM), this would not be sufficient to push the methyls farther apart than when calculated using the TFs. We attribute this behavior to the flattening of the sheets that occur when they are calculated using the DFs, consistent with the deviations from the Ramachandran ϕ and ψ angles and the C = O...N angles reported from the ranges of the databases of experimental results noted above. We should note, that more planar structures are possible for the special case of polyglycine, which has more flexibility due to its enantiomorphic nature.^{37,40} Also, another report has found the DF, M05-2X (which we have not considered here), to predict stronger interactions for the foldamers of β -peptides.⁴¹

The results collected above indicate that the dispersion-corrected functionals tested do not generally provide better results for the energies and geometries of axial vs. equatorial cyclohexane derivatives or polyanilines in the α -helical and antiparallel β -sheet forms.

The X3LYP and PBE1PBE functionals generally gave the best results for the larger systems. B3LYP performed almost as well. For the peptide fragments, using AM1 for the side chains in ONIOM calculations gave different results from the pure functionals. The latter calculations might be in better accord with the experimental reports that many β -sheets become twisted in globular proteins. However, the antiparallel β -sheets may be twisted either due to their natural properties or because of the constraints imposed by the tertiary and quaternary structures of the globular proteins in which they are found. The ONIOM calculations used with B97-D clearly give more reasonable results for β -sheets than when used with the other dispersion corrected functionals (see Figure 1).

Dispersion behaves differently when mixed with inductive interactions. We have shown that the interactions of HF or LiF with Ne, methane, and 2-butyne can be described by some of the traditional functionals while those that have been parametrized to account for dispersion tend to overestimate the interaction energies.¹ The observation that functionals, such as B3LYP and X3LYP, perform very well for H-bonds in water dimer³ also suggests that pure dispersion interactions behave differently than when they are coupled with induction. Thus, the better performance of these functionals in describing peptide interactions might be due to the cooperative induction/dispersion nature of these interactions. The fact that the current results show that the DFs behave similarly to each other (except for B97D/ONIOM) suggests that the parametrizations to reproduce the dispersion interaction in the databases used might be at fault.

The data on the cyclohexane derivatives show that the TFs all perform better than the DFs with the smaller (D95**) basis set, but the results for the TFs become worse with the larger (aug-cc-pVTZ) basis. The TFs all overestimate and the DFs all underestimate the ΔH s for D95**. With the larger basis all ΔH s for the TFs increase in value (and error). The TFs, do not all behave similarly with increase in basis size. Three (B97-D, M06-2X, and ω B97x-D) give improved results, while two (M06 and M06L) give poorer results based upon the MSE's and MUE's of Table I. Since the reduction of BSSE that accompanies the increase in basis size would cause ΔH to increase, one may conclude that the TFs do not properly describe dispersion, which has been shown many times. Since the increase in basis size raises ΔH for the three TFs mentioned above, one may conclude that they behave similarly to the TFs with respect to basis set size and attribute the improvement to decrease in BSSE. However, the two other DFs show decreases in the already too low ΔH s indicating an increased dispersion interaction predominates. We note that the interactions in the cyclohexanes are probably closer to pure dispersion (as opposed to induction/dispersion) than for the peptides. One would not expect the axial C–H bonds to be very polarizable in the direction of the substituent (perpendicular to the C–H bond axis). Chlorocyclohexane probably has

the largest induction interaction (the F in fluoro cyclohexane is small and probably too far away from the axial H's).

Perhaps parametrization to databases of good experimental data would be useful. We do not think further speculation about the cause of these problems to be warranted purely based on the observations reported here.

IV. CONCLUSIONS

We have found that the several DFs tend to overestimate the attractive interactions in both axial cyclohexane derivatives and in α -helical and β -sheet peptides when compared to experimental energetic and structural information. In particular, the DFs significantly overestimate the folding enthalpy per residue of the α -helices and predict several geometric parameters to be outside the ranges of the databases cited above. We believe the interaction energies for the β -sheets to be too great when calculated with the DFs, but we have no experimental data available for comparison. On the other hand, the TFs (B3LYP, X3LYP, and PBE1PBE) functionals consistently provide better accord with the same experimental results. We do not generally recommend use of the dispersion correct functionals considered here, B97-D, ω B97x-D, M06, or M06-2X, for determination of peptide structures or interactions. Other functionals that might be developed should be carefully evaluated using the available experimental data on peptides and other systems before they can be used with confidence in this field.⁴²

ACKNOWLEDGMENTS

The work described was supported by Award No: SC1AG034197 from the National Institute on Aging.

¹D. Roy, M. Marianski, N. Maitra, and J. J. Dannenberg, "Comparison of some dispersion-corrected and traditional functionals with CCSD(T) and MP2 *ab initio* methods: Dispersion, induction, and basis set superposition error," J. Chem. Phys. (submitted).

²M. Marianski, A. Oliva, and J. J. Dannenberg, "A reinvestigation of the dimer of para-benzoquinone with pyrimidine with MP2, CCSD(T), and DFT using functionals including those designed to describe dispersion," J. Phys. Chem. A (in press).

³J. A. Plumbley and J. J. Dannenberg, J. Comput. Chem. **32**(8), 1519 (2011).

⁴J. R. Durig, A. M. El Defrawy, R. M. Ward, G. A. Guirgis, and T. K. Gounev, Struct. Chem. **19**(4), 579 (2008).

⁵J. R. Durig, R. M. Ward, G. A. Guirgis, and T. K. Gounev, J. Raman Spectrosc. **40**(12), 1919 (2009).

⁶J. R. Durig, R. M. Ward, K. G. Nelson, and T. K. Gounev, J. Mol. Struct. **976**(1–3), 150 (2010).

⁷M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 09, Revision A.2., Gaussian, Inc., Wallingford, CT, 2009.

⁸H. B. Jansen and P. Ros, Chem. Phys. Lett. **3**(3), 140 (1969); S. F. Boys and F. Bernardi, Mol. Phys. **19**(4), 553 (1970).

⁹S. Simon, M. Duran, and J. J. Dannenberg, J. Chem. Phys. **105**(24), 11024 (1996).

¹⁰S. Grimme, J. Comput. Chem. **27**(15), 1787 (2006).

¹¹J.-D. Chai and M. Head-Gordon, Phys. Chem. Chem. Phys. **10**(44), 6615 (2008).

¹²Y. Zhao and D. G. Truhlar, Theor. Chem. Acc. **120**(1–3), 215 (2008); Acc. Chem. Res. **41**, 157 (2008).

¹³A. D. Becke, J. Chem. Phys. **98**, 5648 (1993); C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

¹⁴C. Adamo and V. Barone, J. Chem. Phys. **110**(13), 6158 (1999).

¹⁵X. Xu and W. A. Goddard, Proc. Natl. Acad. Sci. U.S.A. **101**(9), 2673 (2004).

- ¹⁶T. Vreven and K. Morokuma, *J. Chem. Phys.* **113**(8), 2969 (2000); K. Morokuma, *Bull. Korean Chem. Soc.* **24**(6), 797 (2003).
- ¹⁷M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, *J. Am. Chem. Soc.* **107**(13), 3902 (1985).
- ¹⁸E. R. Johnson, R. A. Wolkow, and G. A. DiLabio, *Chem. Phys. Lett.* **394**(4–6), 334 (2004); E. R. Johnson, A. D. Becke, C. D. Sherrill, and G. A. DiLabio, *J. Chem. Phys.* **131**(3), 034111 (2009).
- ¹⁹R. Wiczorek and J. J. Dannenberg, *J. Am. Chem. Soc.* **126**(43), 14198 (2004).
- ²⁰R. Wiczorek and J. J. Dannenberg, *J. Am. Chem. Soc.* **127**, 14534 (2005).
- ²¹P. Salvador, A. Asensio, and J. J. Dannenberg, *J. Phys. Chem. B* **111**(25), 7462 (2007).
- ²²G. E. Job, R. J. Kennedy, B. Heitmann, J. S. Miller, S. M. Walker, and D. S. Kemp, *J. Am. Chem. Soc.* **128**(25), 8227 (2006).
- ²³J. M. Richardson, M. M. Lopez, and G. I. Makhatadze, *Proc. Natl. Acad. Sci. U.S.A.* **102**(5), 1413 (2005).
- ²⁴J. M. Scholtz, H. Qian, E. J. York, J. M. Stewart, and R. L. Baldwin, *Biopolymers* **31**, 1463 (1991).
- ²⁵R. J. Kennedy, S. M. Walker, and D. S. Kemp, *J. Am. Chem. Soc.* **127**(48), 16961 (2005).
- ²⁶G. Rialdi and J. Hermans, *J. Am. Chem. Soc.* **88**(24), 5719 (1966).
- ²⁷P. Y. Chou and H. A. Scheraga, *Biopolymers* **10**(4), 657 (1971).
- ²⁸T. Ooi and M. Oobatake, *Proc. Natl. Acad. Sci. U.S.A.* **88**(7), 2859 (1991).
- ²⁹G. D. Hawkins, C. J. Cramer, and D. G. Truhlar, *J. Phys. Chem. B* **102**(17), 3257 (1998).
- ³⁰M. Marianski and J. J. Dannenberg, *J. Phys. Chem. B* **116**(4), 1437 (2012).
- ³¹E. N. Baker and R. E. Hubbard, *Prog. Biophys. Mol. Biol.* **44**(2), 97 (1984).
- ³²O. Koch, M. Bocola, and G. Klebe, *Proteins: Struct., Funct., Bioinf.* **61**(2), 310 (2005).
- ³³J. A. Plumley and J. J. Dannenberg, *J. Phys. Chem. B* **115**, 10560 (2011).
- ³⁴T. E. Creighton, *Proteins: Structures and Molecular Principles* (Freeman, New York, 1984).
- ³⁵S. Spera and A. Bax, *J. Am. Chem. Soc.* **113**(14), 5490 (1991).
- ³⁶J. D. van Beek, S. Hess, F. Vollrath, and B. H. Meier, *Proc. Natl. Acad. Sci. U.S.A.* **99**(16), 10266 (2002).
- ³⁷R. Viswanathan, A. Asensio, and J. J. Dannenberg, *J. Phys. Chem. A* **108**(42), 9205 (2004).
- ³⁸C. Chothia, M. Levitt, and D. Richardson, *Proc. Natl. Acad. Sci. U.S.A.* **74**(10), 4130 (1977).
- ³⁹S. Hua, L. Xu, W. Li, and S. Li, *J. Phys. Chem. B* **115**(39), 11462–11469 (2011).
- ⁴⁰S. MuZoz-Guerra, J. Puiggalí, A. Rodríguez, and J. A. Subirana, *J. Mol. Biol.* **167**(1), 223 (1983); F. Colonna-Cesari, S. Premilat, and B. Lotz, *ibid.* **87**(2), 181 (1974); B. Lotz, C. R. Acad. Sci., Ser. C: Sci. Chim **274**(23), 1907 (1972); B. Lotz, *J. Mol. Biol.* **87**(2), 169 (1974); A. Kajava, *Acta Crystallogr.* **55**(2), 436 (1999); Y.-L. Zhao and Y.-D. Wu, *J. Am. Chem. Soc.* **124**(8), 1570 (2002); J. A. Plumley, M. I.-H. Tsai, and J. J. Dannenberg, *J. Phys. Chem. B* **115**(6), 1562 (2011).
- ⁴¹G. Pohl, T. Beke, I. G. Csizmadia, and A. Perczel, *J. Phys. Chem. B* **114**(29), 9338 (2010).
- ⁴²See supplementary material at <http://dx.doi.org/10.1063/1.4737517> for Tables S1 and S2 and Figures S1 and S2.