chemistry proceedings

MDPI

Proceeding Paper

Perturbation Methods for Obtaining Interactions Energies: Can SAPT2+/aug-cc-pVDZ Predict Adequate Interaction Energies When Biochemically Relevant Motifs Are Present? †

Ángel Vidal-Vidal * and Enrique M. Cabaleiro-Lago

Facultade de Ciencias (Dpto. de Química Física), Universidade de Santiago de Compostela, Campus de Lugo. Avda. Alfonso X El Sabio s/n, 27002 Lugo, Galicia, Spain; e-mail@e-mail.com

* Correspondence: angel.vidal.vidal@usc.es

† Presented at the 26th International Electronic Conference on Synthetic Organic Chemistry; Available online: https://ecsoc-26.sciforum.net.

Abstract: In the field of non-covalent interactions, there has always been a great interest in finding the appropriate methodology to analyze bond energies and properties. There are multiple approaches; however, those based on symmetry adapted perturbation theory (SAPT) are interesting for two different reasons: quality of the interaction energy and how it is obtained. Total interaction energies are computed in SAPT as the sum of the electrostatic, repulsive, inductive and dispersive components. This provides enormous information about the intimate nature of intermolecular interactions. The performance of a variety of symmetry adapted perturbation theory (SAPT) methods for describing non-covalent interactions has been tested in several studies. The appropriate level depends to a certain degree on the nature of the interaction and the extent of the database, however, there is a methodological combination that can be considered as a reference. The SAPT2+(3)δMP2 truncation combined with the aug-cc-pVTZ basis set offers an outstanding performance for the majority of non-covalent complexes. This methodology produces interaction energies of excellent quality with low relative errors and little error spread so it can be adopted as a methodology to obtain reference energies for most applications of interest in chemistry and biochemistry. The problem that SAPT2+(3)δMP2/aug-cc-pVTZ faces is the computational resources demand. These requirements grow enormously with size so that it soon becomes unfeasible for most systems of interest in biochemistry. When the computational cost is prohibitively high it has been suggested the use of SAPT2+ level in combination with the jun-cc-pVDZ basis set. This methodology is known to give remarkable results at a reduced computational cost. In this work, the goodness of the SAPT2+ methodology to produce interaction energies of non-covalent systems is explored using the so-called blind database. This database consists on a set of dimers bearing different type of interactions at equilibrium and non-equilibrium distances. Likewise, the SAPT2+/aug-cc-pVDZ methodology is employed to describe a set of prototypical interactions found in biochemical systems involving sugars, proteins and nucleic acids.

Keywords: SAPT, intermolecular interactions, biochemical motifs, perturbation methods

Citation: Vidal-Vidal, Á.; Cabaleiro-Lago, E.M. Perturbation Methods for Obtaining Interactions Energies: Can SAPT2+/aug-cc-pVDZ Predict Adequate Interaction Energies When Biochemically Relevant Motifs Are Present? *Chem. Proc.* **2022**, *4*, x. https://doi.org/10.3390/xxxxx

Academic Editor(s):

Published: 15 November 2022

1. Introduction

Life as we know it is largely conditioned by the characteristics of non-covalent interactions, so a deep knowledge of their workings can provide tools for understanding and harnessing the processes that sustain it. Non-covalent Interactions play a pivotal role in biochemistry [1–3] governing function through the binding between complementary substrates and receptors in crucial processes such as [4–8] protein-ligand recognition [7,9] drug-receptor binding [10–14] or DNA base-pairing [15,16] among others, [17–19] and determining key aspects of biomolecular structure, properties or folding [20–26].

Over the years, many theoretical and experimental studies have been carried out in order to understand and control the intimate nature of the interactions between biologically relevant molecules or fragments [27–35]. Since their early days, quantum mechanics and electronic structure methods have contributed enormously to the study of non-covalent bonds in biomolecules, [36–44] as reflected in the many different approaches that can be found in literature [45–50]: from studies about the factors governing these interactions to those concerned with determining what methodology provides the most reliable and robust information for the study of non-covalent interactions. [51–56] In the latter case, the use of databases providing a set of benchmark/reference values has greatly facilitated the work, allowing a discrete set of systems to be employed for testing the performance of any given method [27,56].

Different databases have been specifically built for the study of non-covalent interactions [57–61], such as those recently developed by Jan Rezác et al. for different purposes (hydrogen bond, repulsive contacts, dispersive systems, and σ-hole interactions, among others) [51,52,62]. Besides, there are more established databases with specific interactions as their target, such as the X40 database for interaction between halogenated molecules [55], the A24 for small non-covalent complexes [63], or even the L7 database [64], which includes large non-covalent systems. Additionally, there are also some databases that focus on interactions relevant to the field of biochemistry, such as the well-known S66 database made up of systems that model biomolecular interactions [27,57], sets of molecular conformations of peptides such as Peptide_FGG54 and Peptide_WGG54, databases of peptides and macrocyclic compounds such as MPCONF [65,196] or the PLF547 database [67] including interactions between systems such as protein fragments and ligands relevant to computer-aided drug design, among others [54,66,67].

These databases provide a set of reference values against which the general performance of different methodologies can usually be tested, as it is often the case of semiempirical and density functional theory methods [68–72]. The behavior of wavefunction methods can also be tested in their different implementations, as it has been the case of MBPT methods [73–76], strategies based on the Local Pair Natural Orbitals approach [77–82] or perturbative methods such as those derived from symmetry-adapted perturbation theory (SAPT) [83,84]. It is worth noting that, among these databases, the S66 database has become the most widely employed for testing a large variety of methods, allowing a direct comparison about the performance of the different approaches used [68–72,85–89].

Among the assortment of methods available for the study of non-covalent interactions, those based in SAPT are specially interesting due to the way they treat interaction energies. SAPT methods obtain the interaction energy as a sum of different contributions that can be related to physical components of the interaction, such as electrostatic, repulsion, induction and dispersion [37,90,91]. By means of these contributions they allow us to delve into the nature of non-covalent interactions and rationalize the effects that different orientations or substitutions have on each component of the interaction energy. This in-depth understanding of how each component contributes to the global interaction energy can be harnessed to provide predictive power in the design of new structures and control in processes.

However, studies that systematically apply these methodologies to biomolecular databases are scarce [92–97]. The largest study of the efficiency and performance of SAPT methods for interaction energies was carried out in 2014 by Parker et al. [90] who analyzed the quality of different SAPT approaches with different basis sets using several databases; namely S22, NBC10, HBC6 and HSG.

In the context of our research on the use of SAPT to describe non-covalent interactions, the performance of different SAPT truncations will be tested using the so-called blind database [98]. This contains a large set of systems bearing C, H, O and N atoms with both equilibrium and non-equilibrium geometries. We especially sought to study whether the SAPT2+/aug-cc-pVDZ methodology is capable of producing adequate interaction energies in various spatial configurations bearing different intermolecular interactions

relevant to the biochemistry field. Additionally, the performance of some SAPT approaches will be assessed using three model complexes representative of the interaction between DNA bases, the interactions in proteins and the interactions between sugars and proteins.

2. Computational Details

Methods based on Symmetry Adapted Perturbation Theory give the interaction energy of a dimer as a sum of different components that can be associated to physical contributions such as electrostatics, repulsion, induction and dispersion [37,90,91]. The interaction energy can be written as:

$$E_{int} = \sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\left(E_{pol}^{(ij)} + E_{exch}^{(ij)}\right), \quad (1)$$

where i indicates the order in the perturbation theory with respect to the intermolecular potential while j indicates the order with respect to the intramolecular electronic correlation. The polarization energies are identical to the corrections obtained in the conventional Rayleigh-Schrödinger theory while the exchange terms arise from using an antisymmetrizer that forces the dimer wavefunction to have the correct symmetry (antisymmetrized wave function) [99,100]. Since the interaction energy in the SAPT formalism is expanded as a series of contributions, it can be truncated including different terms, leading to a series of MBPT-SAPT approaches of different accuracy, though low order truncations are generally used [90]. For a detailed description of the different terms of the MBPT-SAPT approaches employed, the reader is encouraged to consult the work of Parker et al. [90] or the review of Patkowski, K. [91]. All MBPT-SAPT calculations in this work have been performed using PSI4 1.3.2 software [101].

Thus, interaction energies have been obtained using CCSD(T) at the estimated complete basis set (CBS) limit by using the compound method described below:

$$\Delta E_{CCSD(T)}(CBS) = \Delta E_{MP_2}^{corr}(CBS) + \delta CCSD(T) + \Delta E_{HF}^{(CBS)}, \quad (2)$$

In this approach the MP2 correlation energy and the Hartree-Fock energy are extrapolated to the CBS limit using the Halkier and Helgaker extrapolation procedure [102,103] with the aug-cc-pVTZ and aug-cc-pVQZ basis sets [104,105]. The resolution of identity approach is used [106] in all MP2 calculations using the corresponding auxiliary basis sets. The δCCSD(T) term is evaluated as the difference in correlation energy between CCSD(T) and MP2 obtained with the heavy-aug-cc-pVDZ basis set (no diffuse functions on hydrogen atoms). All interaction energies have been corrected from basis set superposition error by using the counterpoise method [107]. These calculations have been carried out with ORCA 5.0 [108].

3. Results

The analysis will start using the blind database created in 2016 [98]. Specifically, this set of compounds is buildup of 10 different dimers with sizes ranging from 6 to 32 atoms. The complete database consists of 80 points obtained by scanning the distances in the analyzed complexes, separating the centers of mass from the position of the energy minimum. In this way, both equilibrium configurations and non-equilibrium spatial arrangements are included considering both shorter and longer distances. The systems used are the following: water dimer, ethanol dimer, nitromethane dimer, methylformate dimer, benzene-water dimer, benzene-methane dimer, imidazole dimer, nitrobenzene dimer, 1,1-diamino-2,2-dinitroethylene (hereinafter FOX7) dimer, and ethylene dinitramine (hereafter EDNA).

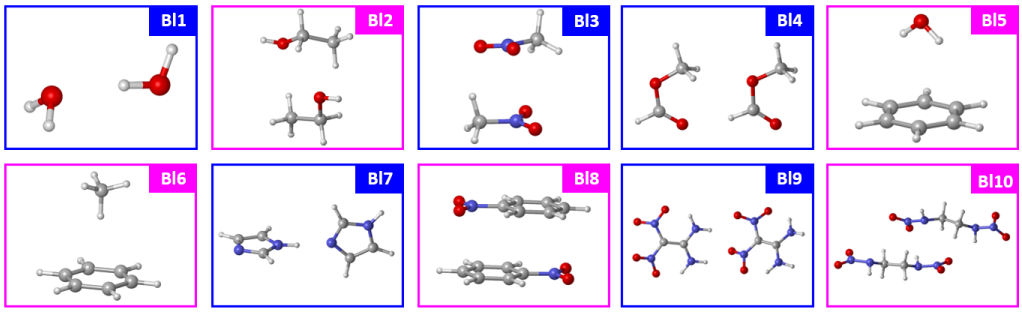

Figure 1. systems included in the blind database. The colored square around each system shows the nature of the complex analyzed in the minimum energy configuration. Electrostatic dominated complexes are shown in blue while dispersion systems are represented in magenta.

The original study [98] includes the comparison of 13 different methodologies (including SAPT-DFT) with respect to CCSD(T)/CBS energies but no mention is made of any of the MBPT-SAPT truncations that are the focus of this work. Thus, we have calculated the interaction energies for these systems with SAPT2+(3)δMP2/aug-cc-pVTZ and SAPT2+/aug-cc-pVDZ and represented the MAD of the resulting 15 methods in Figure 2.

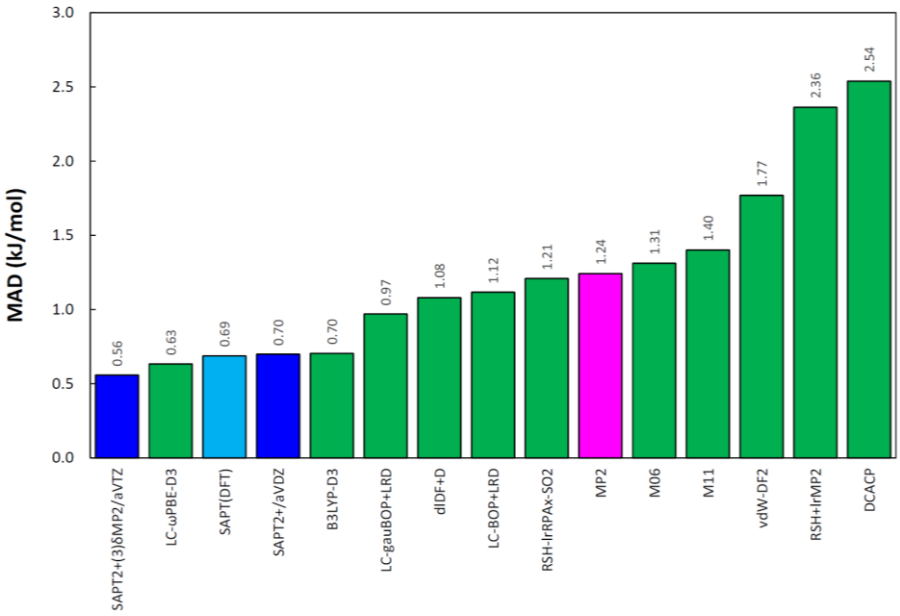

Figure 2. Representation of the MAD (kJ/mol) obtained in the blind database for 15 different methodologies. The color code used serves to group the methodologies according to their nature: dark blue (MBPT-SAPT), light blue (DFT-SAPT), fuchsia (MBPT) and green (DFT). MBPT-SAPT results are obtained in this work while the rest of the values are taken from the literature [98].

As can be seen in Figure 2, with a MAD of 0.56 kJ/mol, SAPT2+(3)δMP2/aug-cc-pVTZ provides the best results, confirming its robustness for the study of non-covalent interactions. Likewise, it should be noted that the SAPT2+/aug-cc-pVDZ methodology, which has a reduced cost compared to the best SAPT truncation, occupies the fourth position with only 0.14 kJ/mol higher MAD. It is important to highlight that among the six methodologies with a MAD under 1 kJ/mol, three of them are SAPT methodologies. In Table 1 the main statistical parameters for the best methodologies can be found.

These results are indicative of the quality of interaction energies provided by SAPT methodologies in general and especially by the SAPT2+(3)δMP2/aug-cc-pVTZ approach.

Of the statistical parameters collected in Table 1, the MAPE values, which are very high, are particularly noteworthy. If only this statistical parameter is analyzed in the 16

methodologies compared in this study, we can see that the MAPE is between 11.44% of the DFT LC-ωPBE-D3 method and 59.14% of the wdW-DF2 method.

Table 1. main statistical parameters obtained in the blind database using different methodologies that produces MAD values lower than 1 kJ/mol. All values are expressed in kJ/mol except for MAPE, whose value is a percentage.

	SAPT2+(3) δMP2 aVTZ	LC-ωPBE-D3	SAPT-(DFT)	SAPT2+ δMP2 VDZ	B3LYP-D3	LC-gauBOP +LRD
MAD	0.559	0.633	0.688	0.699	0.705	0.970
RMSE	0.983	1249	1535	1699	1884	2207
MAPE	15,050	11,443	24,379	21,226	29,506	32,973
MSiE	−0.270	0.073	0.115	−0.187	0.270	0.513

If the best of all methodologies, SAPT2+(3)δMP2/aVTZ is taken and the results are analyzed in detail, it is possible to verify that this statistical parameter is largely affected by the poor values obtained for two points out of the 80 analyzed: benzene-methane dimer R = 3280 Å with an error of 744.1% and in the nitrobenzene dimer R = 2.900 Å with a percentage error of 91.1%. Such high errors may be a consequence of the small values of the reference energies that makes small energy differences being large in percentage.

In any case, the global behavior of SAPT2+(3)δMP2/aVTZ and SAPT2+/aug-cc-pVDZ is very good. Both methodologies have proven to provide very good results in systems of great structural diversity not only in the nature of the functional groups but also in their composition. The good results are not only limited to the minimum energy position, since excellent results are obtained for the interaction energy in non-equilibrium geometries, which gives this methodology even more robustness and makes it the methodology of choice (if the size of the system allows it) to study new unknown complexes.

Application to Systems of Biological Interest

After testing the performance of the different SAPT methods against the blind database, a simple application study has been carried out analyzing the interaction in three complexes representative of the interaction between DNA bases, the interactions in proteins and the interactions between sugars and proteins. The selected systems are the adenine-thymine dimer in a Watson-Crick arrangement (A-T) [53], a phenalanine interaction with a peptide bond (F49-PB) [53], and a fucose-indol (FI) dimer [109]. The structures of these dimers have been taken from literature, and their properties (geometrical arrangement, reference interaction energy at the CCSD(T)/CBS level and non-covalent interaction plot) are shown in Figure 3.

The study is not limited to the application of SAPT2+/aug-cc-pVDZ (the main objective of this work) to the complexes shown above. We settle on to include other SAPT methodologies of different quality (and computational resources demand). In first place, it was decided to continue using the SAPT2+(3)δMP2/aug-cc-pVTZ despite the high computational cost. Additionally, the "low-cost" sSAPT0/jun-cc-pVDZ methodology is also considered since biochemical systems can have huge sizes and is a very low-demanding option. The values of interaction energies and the absolute errors with respect to the reference energy for these three methodologies can be found in Table 2.

In the first place, the A-T system has the highest interaction energy of all the complexes analyzed in this study and also the lowest error values of the three systems considered. All methods overestimate the interaction energy, although to a different extent. For example, the SAPT2+(3)δMP2/aug-cc-pVTZ methodology deviates from the reference value by −0.35 kJ/mol, while SAPT2+/aug-cc-pVDZ deviates by −0.30 kJ/mol. The cheaper sSAPT0/jun-cc-pVDZ shows a larger deviation of −1.75 kJ/mol.

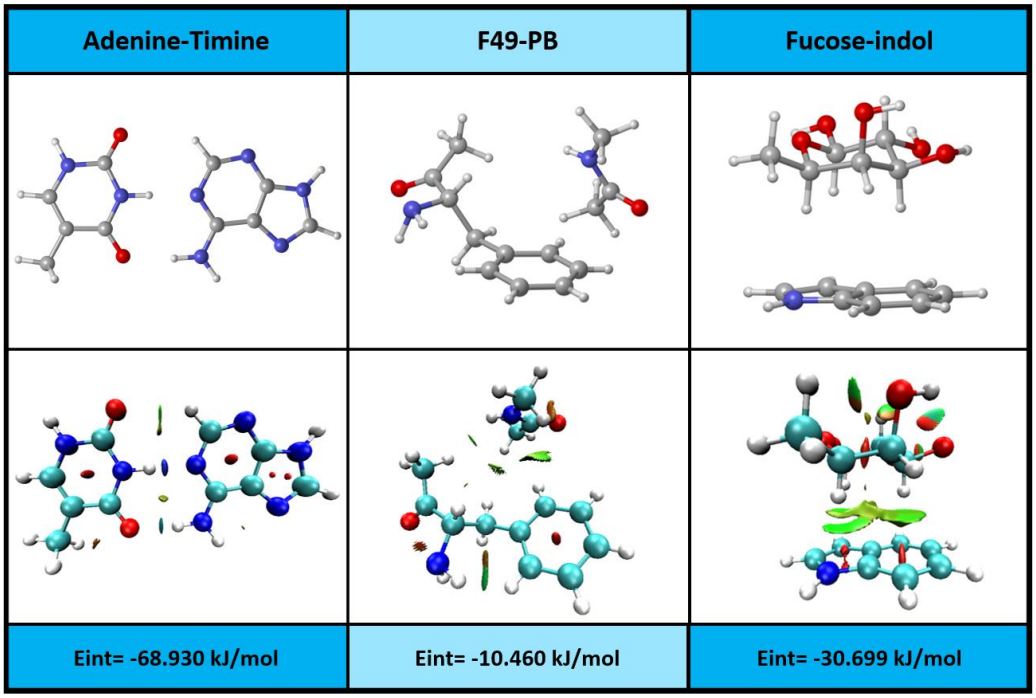

Figure 3. Geometry and non-covalent interactions obtained with the NCI method [110,111] (isosurface of reduced density gradient 0.5 a.u.) for the three biologically relevant models used in this part of the study. Interaction energies are computed at the CCSD(T)/CBS level as described in the computational methods section and are given in kJ/mol.

In the peptide-relevant model, the errors are slightly higher than in the DNA-relevant model. In this case, both SAPT2+(3)δMP2 and SAPT2+ have an error of less than 1 kJ/mol and deviate in the direction of system overstabilization. On the other hand, sSAPT0 has a significantly higher error than the other methodologies but it goes in the opposite direction. The sSAPT0 methodology deviates by more than 2 kJ/mol and provides an energy of −8.29 kJ/mol versus −10.46 kJ/mol for the CCSD(T)/CBS reference. Among the three systems analyzed, this is the only one in which a less stable complex than the reference is predicted.

Table 2. Interaction energies obtained for each of the complexes under study and errors with respect to the CCSD(T)/CBS reference energy. All values are given in kJ/mol.

	CCSD(T)	SAPT2+(3)δMP2	SAPT2+	sSAPT0	ε SAPT2+(3)δMP2	ε SAPT2+	ε sSAPT0
A-T	−68.93	−69.28	−69.23	−70.68	−0.35	−0.30	−1.75
F49-PB	−10.46	−10.95	−11.16	−8.29	−0.49	−0.70	2.17
F-I	−30.70	−32.11	−34.05	−32.86	−1.41	−3.35	−2.16

Lastly, the complex representing the sugar-amino acid interaction exhibits greater deviations with respect to the reference. SAPT2+(3)δMP2 is the methodology that provides the best results with error values around 1 kJ/mol higher than those provided for the rest of the systems. Though the results cannot be said to be bad, it exceeds the desired precision of ±1 kJ/mol by approximately 0.4 kJ/mol. The other methods provide even larger deviations. In view of these results, it can be concluded that SAPT2+(3)δMP2/aug-cc-pVTZ is a very robust methodology for the description of the interaction energy in systems relevant to biomolecular structures. SAPT2+/aug-cc-pVDZ may be a viable alternative depending on the accuracy and quality of the results sought.

Until now only global interaction energies have been considered, but it is also worth asking what happens to the individual components that contribute to the interaction

energy. The overall performance of a methodology can be good, but large errors could be obtained for the energy components. To verify the quality of the interaction energy components, the results provided by SAPT2+/aug-cc-pVDZ and sSAPT0/jun-cc-pVDZ will be compared with those provided by SAPT2+(3)δMP2/aug-cc-pVTZ, which is taken as the reference methodology. The values of the components for each of the methodologies are shown in Table 3.

In the A-T system, notable differences can be seen between the components of the interaction energy at the sSAPT0 and SAPT2+(3)δMP2 levels. The largest ones are located in the repulsion and in the dispersion, with values of −15.49 kJ/mol and 14.09 kJ/mol, respectively. In the opposite direction, these disparate results end up practically compensating each other; however, the percentage errors are close to 12% in repulsion and 32% in dispersion. Electrostatics and induction also have differences that go in the opposite directions with very similar values. While sSAPT0 predicts a higher electrostatic component than SAPT2+(3)δMP2, it falls short in induction. Taking into account the values of the components, an absolute difference of around −3.8 kJ/mol translates into percentage errors of 3.45% in electrostatics and 8.21% in induction. Slight problems with the induction are also observed for the SAPT2+/aug-cc-pVDZ method, which overestimates it by approximately 3.3 kJ/mol, which is associated with a relative error of 7.22%. In the rest of the components, the percentage error does not exceed 3.25% in any case. Therefore, at a global level, for the A-T dimer SAPT2+ provides an acceptable description of the interaction energy components.

The F49-PB system is strongly dominated by the dispersive component, in fact, it accounts for approximately 86% of the total stabilizing interactions. Likewise, electrostatics has a very limited relevance, its contribution not even reaching 5% of the total interaction energy, and both SAPT2+ and sSAPT0 deviate by similar amounts. Also, the deviations are fairly low in repulsion and induction. However, larger differences are observed for the dominant contribution of dispersion. While SAPT2+ has an absolute error of just 0.08 kJ/mol, sSAPT0 deviates by almost 3 kJ/mol, leading to an underestimation of around 20–21% of the dispersion term, which translates to the total interaction energy.

Table 3. Components of the interaction energy and total interaction energy (all in kJ/mol) obtained with the three SAPT methodologies under study in the complexes A-T, F49-PB and F-I. Elec. Stands for SAPT electrostatic energy while Rep., Ind. and Dis. are respectively repulsion, induction and dispersion energies.

A-T					
	Elec.	Rep.	Ind.	Dis.	Tot.
Ref. (SAPT2+(3)δMP2/aug-cc-pVTZ)	−111.14	132.94	−46.67	−44.41	−69.28
SAPT2+/aug-cc-pVDZ	−113.46	137.25	−50.04	−42.99	−69.23
sSAPT0/jun-cc-pVDZ	−114.97	117.45	−42.84	−30.32	−70.68
F49-PB					
Methodology	Elec.	Rep.	Ind.	Dis.	Tot.
Ref. (SAPT2+(3)δMP2/aug-cc-pVTZ)	−0.77	4.96	−1.42	−13.72	−10.95
SAPT2+/aug-cc-pVDZ	−1.18	5.22	−1.56	−13.64	−11.16
sSAPT0/jun-cc-pVDZ	−0.34	4.38	−1.46	−10.87	−8.29
F-I					
	Elec.	Rep.	Ind.	Dis.	Tot.
Ref. (SAPT2+(3)δMP2/aug-cc-pVTZ)	−19.81	36.58	−5.46	−43.42	−32.11
SAPT2+/aug-cc-pVDZ	−21.33	38.26	−7.29	−43.68	−34.05
sSAPT0/jun-cc-pVDZ	−22.16	33.58	−6.82	−37.47	−32.86

Finally, for the fucose-indole complex, there are larger errors in the total interaction energy for SAPT2+ than for sSAPT0. However, this is a result of error compensation, mainly on electrostatics and repulsion with dispersion. The sSAPT0 methodology

overestimates electrostatics by −2.3 kJ/mol while underestimating repulsion by 3.09 kJ/mol. Thus, the interaction energy suffers from an overestimation of around 5 kJ/mol in these two terms that is compensated by an underestimation of dispersion of around 6 kJ/mol, leading to a nice total interaction energy. Nevertheless, the qualitative picture of the contributions of the different components is fairly good considering the saving in computational effort of the cheaper methods. If a proper balance of the different contributions must be ensured it is necessary to invest in computational resources going to at least SAPT2+ or even better to SAPT2+(3)δMP2.

4. Conclusions

Thanks to the use of a blind database (which includes equilibrium and non-equilibrium structures) we can state that SAPT2+(3)δMP2/aug-cc-pVTZ is a very robust methodology, providing the best statistical descriptors. However, it is important to mention that SAPT2+/aug-cc-pVDZ methodology which has a remarkable lower computational cost makes a good job and with MAD less than 1 kJ/mol also gives excellent results. After the application of this methodology to different model biochemical systems, results led as to conclude that SAPT2+ is able to predict fairly good intermolecular interaction energies. Moreover, SAPT2+/aug-cc-pVDZ provides an excellent picture of the contributions of the different energy components considering the associated computational cost. Finally, it should be highlighted that the SAPT2+/aVDZ methodology is a good choice for the study of biochemical systems.

Author Contributions:

Funding:

Institutional Review Board Statement:

Informed Consent Statement:

Data Availability Statement:

Acknowledgments: The authors thank the financial support from the Consellería de Cultura, Educación e Ordenación Universitaria e da Consellería de Economía, Emprego e Industria (Axuda para Consolidación e Estruturación de unidades de investigación competitivas do Sistema Universitario de Galicia, Xunta de Galicia ED431C 2021/40). The authors also want to express their gratitude to the CESGA (Centro de Supercomputación de Galicia) for the generous allocation of computation resources. Á. V-V. thanks the University of Vigo for granting a "Margarita Salas" postdoctoral fellowship.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Riley, K.E.; Hobza, P. Title. *WIREs Comput. Mol. Sci.* **2011**, *1*, 3–17.
2. Černý, J.; Hobza, P. Title. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5291–5303.
3. Kollman, P. Title. In *New Compr. Biochem.*; Page, M.I., Ed.; Elsevier: Amsterdam, The Netherlands, 1984; Volume 6, pp. 55–71.
4. Wlodarski, T.; Zagrovic, B. Title. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 19346–19351.
5. Mazik, M. Title. *Chem. Soc. Rev.* **2009**, *38*, 935–956.
6. Davis, A.P.; Wareham, R.S. Title. *Angew. Chem. Int. Ed.* **1999**, *38*, 2978–2996.
7. Williams, D.H.; Stephens, E.; O'Brien, D.P.; Zhou, M. Title. *Angew. Chem. Int. Ed.* **2004**, *43*, 6596–6616.
8. Bordenave, N.; Hamaker, B.R.; Ferruzzi, M.G. Title. *Food Funct.* **2014**, *5*, 18–34.
9. Zhou, P.; Huang, J.; Tian, F. Title. *Curr. Med. Chem.* **2012**, *19*, 226–238.
10. Cho, Y.; Min, S.K.; Yun, J.; Kim, W.Y.; Tkatchenko, A.; Kim, K.S. Title. *J. Chem. Theory Comput.* **2013**, *9*, 2090–2096.
11. Cusumano, M.; Pietro, M.L.D.; Giannetto, A.; Nicolò, F.; Rotondo, E. Title. *Inorg. Chem.* **1998**, *37*, 563–568.
12. Smith, A.J.T.; Zhang, X.; Leach, A.G.; Houk, K.N. Title. *J. Med. Chem.* **2009**, *52*, 225–233.
13. Blanco, F.; Kelly, B.; Sánchez-Sanz, G.; Trujillo, C.; Alkorta, I.; Elguero, J.; Rozas, I. Title…. *J. Phys. Chem. B* **2013**, *117*, 11608–11616.
14. Ferguson, L.R.; Denny, W.A. *Mutat. Res.—Fundam. Mol. Mech. Mutagen.* **2007**, *623*, 14–23.
15. Rehman, S.U.; Sarwar, T.; Husain, M.A.; Ishqi, H.M.; Tabish, M. *Arch. Biochem. Biophys.* **2015**, *576*, 49–60.
16. Strekowski, L.; Wilson, B. *Mutat. Res.—Fundam. Mol. Mech. Mutagen.* **2007**, *623*, 3–13.

17. Stasyuk, O.A.; Jakubec, D.; Vondrášek, J.; Hobza, P. J. *Chem. Theory Comput.* **2017**, *13*, 877–885.
18. Lejeune, D.; Delsaux, N.; Charloteaux, B.; Thomas, A.; Brasseur, R. *Proteins: Struct. Funct. Genet.* **2005**, *61*, 258–271.
19. Hinson, J.A.; Roberts, D.W. *Annu. Rev. Pharmacool. Toxicol.* **1992**, *32*, 471–510.
20. Frieden, E. J. *Chem. Educ.* **1975**, *52*, 754.
21. Gurusaran, M.; Shankar, M.; Nagarajan, R.; Helliwell, J.R.; Sekar, K. *IUCrJ* **2014**, *1*, 74–81.
22. Mundlapati, V.R.; Sahoo, D.K.; Bhaumik, S.; Jena, S.; Chandrakar, A.; Biswal, H.S. *Angew. Chem. Int. Ed.* **2018**, *57*, 16496–16500.
23. Romano, C.; Miller, J.K.; Hyrc, K.; Dikranian, S.; Mennerick, S.; Takeuchi, Y.; Goldberg, M.P.; Malley, K.L. *Mol. Pharmacol.* **2001**, *59*, 46.
24. Zhang, Q.; Cheng, Z.; Chen, R.; Wang, Y.; Miao, S.; Li, Z.; Wang, S.; Fu, L. *Food Funct.* **2021**, *12*, 10107–10120.
25. Prigent, S.V.E.; Gruppen, H.; Visser, A.J.W.G.; van Koningsveld, G.A.; de Jong, G.A.H.; Voragen, A.G.J. *J. Agric. Food. Chem.* **2003**, *51*, 5088–5095.
26. Rajabi, M.; Farhadian, S.; Shareghi, B.; Asgharzadeh, S.; Momeni, L. *Colloids Surf. B* **2019**, *183*, 110287.
27. Řezáč, J.; Riley, K.E.; Hobza, P. *J. Chem. Theory Comput.* **2011**, *7*, 2427–2438.
28. Guan, L.; Grigoriev, A. *Nucleic Acids Res.* **2021**, *49*, 4085–4103.
29. Priyadharshini, R.D.; Ponkarpagam, S.; Vennila, K.N.; Elango, K.P. *Spectrochim. Acta Part A Mol. Biomol.Spectrosc.* **2022**, *271*, 120888.
30. Rangel-Galván, M.; Castro, M.E.; Perez-Aguilar, J.M.; Caballero, N.A.; Rangel-Huerta, A.; Melendez, F.J. *Molecules* **2022**, 27.
31. Shahroz, M.M.; Sharma, H.K.; Altamimi, A.S.A.; Alamri, M.A.; Ali, A.; Ali, A.; Alqahtani, S.; Altharawi, A.; Alabbas, A.B.; Alossaimi, M.A.; Riadi, Y.; Firoz, A.; Afzal, O. *Molecules* **2022**, *27*.
32. Espinosa, J.F.; Asensio, J.L.; García, J.L.; Laynez, J.; Bruix, M.; Wright, C.; Siebert, H.-C.; Gabius, H.-J.; Cañada, F.J.; Jiménez-Barbero, J. *Eur. J. Biochem.* **2000**, *267*, 3965–3978.
33. Riveras, J.A.F.; Frontera, A.; Bauzá, A. *Phys. Chem. Chem. Phys.* **2021**, *23*, 17656–17662.
34. Cao, Y.; Park, S.-J.; Im, W. *Glycobiology* **2021**, *31*, 126–136.
35. Hudson, K.L.; Bartlett, G.J.; Diehl, R.C.; Agirre, J.; Gallagher, T.; Kiessling, L.L.; Woolfson, D.N. *J. Am. Chem. Soc.* **2015**, *137*, 15152–15160.
36. Trujillo, C.; Sánchez-Sanz, G.; Alkorta, I.; Elguero, J. *New J. Chem.* **2015**, *39*, 6791–6802.
37. Vidal-Vidal, Á.; Cabaleiro-Lago, E.M.; López, C.S.; Faza, O.N. *ACS Omega* **2018**, *3*, 16976–16988.
38. Josa, D.; Rodríguez-Otero, J.; Cabaleiro-Lago, E.M.; Rellán-Piñeiro, M. *Chem. Phys. Lett.* **2013**, *557*, 170–175.
39. Alkorta, I.; Hill, J.G.; Legon, A.C. *Phys. Chem. Chem. Phys.* **2020**, *22*, 16421–16430.
40. Vos, E.; Montero-Campillo, M.M.; Corral, I.; Yáñez, M.; Alkorta, I.; Elguero, J. *ChemPhysChem* **2020**, *21*, 2701–2708.
41. Bauzá, A.; Ramis, R.; Frontera, A. *J. Phys. Chem. A* **2014**, *118*, 2827–2834.
42. Bauzá, A.; Frontera, A. *ChemPhysChem* **2020**, *21*, 26–31.
43. Bauzá, A.; Quiñonero, D.; Deyà, P.M.; Frontera, A. *CrystEngComm* **2013**, *15*, 3137–3144.
44. Malloum, A.; Conradie, J. *J. Mol. Liq.* **2022**, *350*, 118522.
45. Andersson, M.P.; Jones, M.N.; Mikkelsen, K.V.; You, F.; Mansouri, S.S. *Curr. Opin. Chem. Eng.* **2022**, *36*, 100754.
46. Cui, Q.; Pal, T.; Xie, L. *J. Phys. Chem. B* **2021**, *125*, 689–702.
47. Schlick, T.; Portillo-Ledesma, S. *Nat. Comput. Sci.* **2021**, *1*, 321–331.
48. van der Spoel, D. *Curr. Opin. Struct. Biol.* **2021**, *67*, 18–24.
49. Schlick, T.; Portillo-Ledesma, S.; Myers, C.G.; Beljak, L.; Chen, J.; Dakhel, S.; Darling, D.; Ghosh, S.; Hall, J.; Jan, M.; Liang, E.; Saju, S.; Vohr, M.; Wu, C.; Xu, Y.; Xue, E. *Annu. Rev. Biophys.* **2021**, *50*, 267–301.
50. Tolbatov, I.; Marrone, A. *Inorg. Chim. Acta* **2022**, *530*, 120686.
51. Kříž, K.; Nováček, M.; Řezáč, J. J. *Chem. Theory Comput.* **2021**, *17*, 1548–1561.
52. Řezáč, J. J. *Chem. Theory Comput.* **2020**, *16*, 2355–2368.
53. Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.
54. Valdes, H.; Pluháčková, K.; Pitonák, M.; Řezáč, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2747–2757.
55. Řezáč, J.; Riley, K.E.; Hobza, P. J. *Chem. Theory Comput.* **2012**, *8*, 4285–4292.
56. Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215.
57. Řezáč, J.; Riley, K.E.; Hobza, P. *J. Chem. Theory Comput.* **2011**, *7*, 3466–3470.
58. Mintz, B.J.; Parks, J.M. *J. Phys. Chem. A* **2012**, *116*, 1086–1092.
59. Miriyala, V.M.; Řezáč, J. *J. Phys. Chem. A* **2018**, *122*, 2801–2808.
60. Miriyala, V.M.; Řezáč, J. *J. Phys. Chem. A* **2018**, *122*, 9585–9586.
61. Sparrow, Z.M.; Ernst, B.G.; Joo, P.T.; Lao, K.U.; DiStasio, R.A. *J. Chem. Phys.* **2021**, *155*, 184303.
62. Řezáč, J. *J. Chem. Theory Comput.* **2020**, *16*, 6305–6316.
63. Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2013**, *9*, 2151–2155.
64. Sedlak, R.; Janowski, T.; Pitoňák, M.; Řezáč, J.; Pulay, P.; Hobza, P. *J. Chem. Theory Comput.* **2013**, *9*, 3364–3374.
65. Řezáč, J.; Bím, D.; Gutten, O.; Rulíšek, L. *J. Chem. Theory Comput.* **2018**, *14*, 1254–1266.
66. Řeha, D.; Valdés, H.; Vondrášek, J.; Hobza, P.; Abu-Riziq, A.; Crews, B.; de Vries, M.S. *Chem. Eur. J.* **2005**, *11*, 6803–6817.
67. Kříž, K.; Řezáč, J. *J. Chem. Inf. Model. JCIM* **2020**, *60*, 1453–1460.
68. Brémond, E.; Li, H.; Sancho-García, J.C.; Adamo, C. *J. Phys. Chem. A* **2022**, *126*, 2590–2599.
69. Goerigk, L.; Kruse, H.; Grimme, S. *ChemPhysChem* **2011**, *12*, 3421–3433.

70. Nascimento, D.R.; DePrince, A.E. *J. Chem. Theory Comput.* **2014**, *10*, 4324–4331.
71. Kesharwani, M.K.; Karton, A.; Sylvetsky, N.; Martin, J.M.L. *Aust. J. Chem.* **2018**, *71*, 238–248.
72. Yu, F.; Wang, Y. *J. Comput. Chem.* **2020**, *41*, 1018–1025.
73. Liakos, D.G.; Izsák, R.; Valeev, E.F.; Neese, F. *Mol. Phys.* **2013**, *111*, 2653–2662.
74. Sedlak, R.; Riley, K.E.; Řezáč, J.; Pitoňák, M.; Hobza, P. *ChemPhysChem* **2013**, *14*, 698–707.
75. Matveeva, R.; Erichsen, M.F.; Koch, H.; Høyvik, I.-M. *J. Comput. Chem.* **2022**, *43*, 121–131.
76. Shee, J.; Loipersberger, M.; Rettig, A.; Lee, J.; Head-Gordon, M. *J. Phys. Chem. Lett.* **2021**, *12*, 12084–12097.
77. Altun, A.; Izsák, R.; Bistoni, G. *Int. J. Quantum Chem*. **2021**, *121*, e26339.
78. Chen, J.-L.; Sun, T.; Wang, Y.-B.; Wang, W. *J. Comput. Chem.* **2020**, *41*, 1252–1260.
79. Ma, Q.; Werner, H.-J. *J. Chem. Theory Comput.* **2019**, *15*, 1044–1052.
80. Schmitz, G.; Hättig, C.; Tew, D.P. *Phys. Chem. Chem. Phys.* **2014**, *16*, 22167–22178.
81. Calbo, J.; Sancho-García, J.C.; Ortí, E.; Aragó, J. *J. Compote. Chem.* **2017**, *38*, 1869–1878.
82. Liakos, D.G.; Sparta, M.; Kesharwani, M.K.; Martin, J.M.L.; Neese, F. *J. Chem. Theory Comput.* **2015**, *11*, 1525–1539.
83. Heßelmann, A. *J. Chem. Theory Comput.* **2018**, *14*, 1943–1959.
84. Lao, K.U.; Herbert, J.M. *J. Phys. Chem.Lett.* **2012**, *3*, 3241–3248.
85. Schwilk, M.; Mezei, P.D.; Tahchieva, D.N.; von Lilienfeld, O.A. *Electron.Struct.* **2022**, *4*, 014005.
86. Gao, T.; Li, H.; Li, W.; Li, L.; Fang, C.; Li, H.; Hu, L.; Lu, Y.; Su, Z.-M. *J. Cheminform.* **2016**, *8*, 24.
87. Yu, F.; Fu, L.-X.; Yang, Y. *Int. J. Quantum Chem* **2017**, *117*, e25417.
88. Suárez, D.; Díaz, N.; Francisco, E.; Pendás, A.M. *ChemPhysChem* **2018**, *19*, 973–987.
89. Shao, X.; Mi, W.; Pavanello, M. *Proc. Natl. Acad. Sci. USA* **2021**, *17*, 3455–3461.
90. Parker, T.M.; Burns, L.A.; Parrish, R.M.; Ryno, A.G.; Sherrill, C.D. *J. Chem. Phys.* **2014**, *140*, 094106.
91. Patkowski, K. *WIREs Comput. Mol. Sci.* **2020**, *10*, e1452.
92. Heßelmann, A.; Korona, T. J. *Chem. Phys.* **2014**, *141*, 094107.
93. Stasyuk, O.A.; Sedlak, R.; Guerra, C.F.; Hobza, P. J. *Chem. Theory Comput.* **2018**, *14*, 3440–3450.
94. Lao, K.U.; Herbert, J.M. J. *Chem. Phys.* **2014**, *140*, 044108.
95. Li, A.; Muddana, H.S.; Gilson, M.K. J. *Chem. Theory Comput.* **2014**, *10*, 1563–1575.
96. Lao, K.U.; Herbert, J.M. J. *Chem. Theory Comput.* **2018**, *14*, 2955–2978.
97. Hesselmann, A. *J. Phys. Chem. A* **2011**, *115*, 11321–11330.
98. Taylor, D.E.; Ángyán, J.G.; Galli, G.; Zhang, C.; Gygi, F.; Hirao, K.; Song, J.W.; Rahul, K.; von Lilienfeld, O.A.; Podeszwa, R.; Bulik, I.W.; Henderson, T.M.; Scuseria, G.E.; Toulouse, J.; Peverati, R.; Truhlar, D.G.; Szalewicz, K. *J. Chem. Phys.* **2016**, *145*, 124105.
99. Hohenstein, E.G.; Sherrill, C.D. *WIREs Comput. Mol. Sci.* **2012**, *2*, 304–326.
100. Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887–1930.
101. Turney, J.M.; Simmonett, A.C.; Parrish, R.M.; Hohenstein, E.G.; Evangelista, F.A.; Fermann, J.T.; Mintz, B.J.; Burns, L.A.; Wilke, J.J.; Abrams, M.L.; Russ, N.J.; Leininger, M.L.; Janssen, C.L.; Seidl, E.T.; Allen, W.D.; Schaefer, H.F.; King, R.A.; Valeev, E.F.; Sherrill, C.D.; Crawford, T.D. *WIREs Comput. Mol. Sci.* **2012**, *2*, 556–565.
102. Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A.K. *Chem. Phys. Lett.* **1998**, *286*, 243–252.
103. Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639–9646.
104. Woon, D.E.; Dunning, T.H. *J. Chem. Phys.* **1993**, *98*, 1358–1371.
105. Kendall, R.A.; Dunning, T.H.; Harrison, R.J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
106. Feyereisen, M.; Fitzgerald, G.; Komornicki, A. *Chem. Phys. Lett.* **1993**, *208*, 359–363.
107. Boys, S.F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–566.
108. Neese, F. *WIREs Comput. Mol. Sci.* 2022, n/a, e1606.
109. Tsuzuki, S.; Uchimaru, T.; Mikami, M. *J. Phys. Chem. A* **2011**, *115*, 11256–11262.
110. Johnson, E.R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A.J.; Yang, W. J. *Am. Chem. Soc.* **2010**, *132*, 6498–6506.
111. Contreras-García, J.; Johnson, E.R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D.N.; Yang, W. J. *Chem. Theory Comput.* **2011**, *7*, 625–632.