

Article

Microsolvation of Histidine—A Theoretical Study of Intermolecular Interactions Based on AIM and SAPT Approaches

Beata Kizior, Jarosław J. Panek  and Aneta Jezierska * 

Faculty of Chemistry, University of Wrocław, ul. F. Joliot-Curie 14, 50-383 Wrocław, Poland; 284184@uwr.edu.pl (B.K.); jaroslaw.panek@chem.uni.wroc.pl (J.J.P.)

* Correspondence: aneta.jezierska@chem.uni.wroc.pl; Tel.: +48-71-375-7224

Received: 4 June 2020; Accepted: 8 July 2020; Published: 10 July 2020



Abstract: Histidine is unique among amino acids because of its rich tautomeric properties. It participates in essential enzymatic centers, such as catalytic triads. The main aim of the study is the modeling of the change of molecular properties between the gas phase and solution using microsolvation models. We investigate histidine in its three protonation states, microsolvated with 1:6 water molecules. These clusters are studied computationally, in the gas phase and with water as a solvent (Polarizable Continuum Model, PCM) within the Density Functional Theory (DFT) framework. The structural analysis reveals the presence of intra- and intermolecular hydrogen bonds. The Atoms-in-Molecules (AIM) theory is employed to determine the impact of solvation on the charge flow within the histidine, with emphasis on the similarity of the two imidazole nitrogen atoms—topologically not equivalent, they are revealed as electronically similar due to the heterocyclic ring aromaticity. Finally, the Symmetry-Adapted Perturbation Theory (SAPT) is used to examine the stability of the microsolvation clusters. While electrostatic and exchange terms dominate in magnitude over polarization and dispersion, the sum of electrostatic and exchange term is close to zero. This makes polarization the factor governing the actual interaction energy. The most important finding of this study is that even with microsolvation, the polarization induced by the presence of implicit solvent is still significant. Therefore, we recommend combined approaches, mixing explicit water molecules with implicit models.

Keywords: histidine; microsolvation model; hydrogen bonding; DFT; PCM; AIM; SAPT

1. Introduction

Intermolecular interactions play an important role in many processes at the molecular level, being involved in, for example, the formation of organometallic structures and supramolecular complexes [1,2]. Other important issues are the protein–ligand interactions, where weak but numerous non-covalent interactions stabilize the conformation, the tautomeric equilibria, and the distribution of atomic charges [3]. Another aspect that should be discussed is the solute–solvent mutual relation, which can be decisive in many processes in nature [4,5]. In general, the intermolecular forces are divided into several classes, depending on the strength of the interaction: ion–ion (electrostatic), hydrogen bonding, dipole–dipole (dipole-induced dipole), van der Waals and dispersion forces [6]. The theoretical description of such interactions has been developed, but there are still open questions and problems, which should be studied [7,8]. Moreover, new intermolecular interactions have been discovered and introduced, e.g., halogen bonds [9,10] and charge inverted hydrogen bonds [11,12]. Here, we discuss the intermolecular forces within the solute–solvent interplay.

The theoretical solvation models are divided into two groups: implicit and explicit [13]. The implicit models are usually used to analyze the structural changes and estimate the free energy profiles of

solute–solvent interactions in chemical processes. Among the most popular continuum solvation models, one can find the Polarizable Continuum Model (PCM) [14] and the Conductor-like Screening Model (COSMO) [15]. The advantage of the application of such models in *ab initio* simulations is the feasibility of simulations, taking into account the CPU time and difficulties related to the model construction. The explicit solvation models are more demanding computationally, and their complexity increases when the number of solvent molecules grows. Many explicit solvation models have been developed describing the specific chemical nature of the molecules. In general, they consider the number of interactions, the flexibility or rigidity of the model, and the inclusion of the polarity, as the main components [16]. Besides the solvation models, one can find gas phase studies, wherein the solvent (water) molecules are added to monitor the geometric and electronic structural changes in molecules (microsolvation models) [17]. Another solution involves the combination of microsolvation with the Polarizable Continuum Model [18]. In the current study, the histidine molecule is discussed in the view of solute–solvent interactions, using quantum-mechanical methods.

Amino acids are the basic building blocks of proteins. They share a common core structure, but differ in their properties due to the diversity of their side chains. Among the amino acids, histidine is unique due to its rich tautomeric equilibria. The two nitrogen atoms of the histidine imidazole ring, $N\delta$ and $N\epsilon$, can be independently protonated, which, together with the acid–base equilibrium of the backbone groups, gives rise to the many tautomers found in molecular histidine [19] and especially in proteins, where different tautomers are often found in differing positions of a single protein chain [20–22]. The pK_a for the imidazole nitrogen sites is ca. 6, and this is close enough to the physiological or neutral pH such that the tautomeric forms can coexist. These facts make the histidine suitable for participating in essential enzymatic centers, such as catalytic triads of subtilisin and trypsin families [23]. On the other hand, the imidazole nitrogen atoms readily coordinate metal ions, which introduces diverse biochemical effects, e.g., the mercury poisoning of serine protease [24,25], the modulation of the apoptotic activity of alloferone peptides by copper (II) [26], or the best-known example—the formation of the Fe-N(His) coordination bond and its participation in the hydrogen bonding network in many heme-containing proteins [27]. The $N\delta$ -H tautomer is not as common in proteins as the $N\epsilon$ -H tautomer, but the relative ease of the histidine adopting one of the two neutral tautomeric forms, together with the large variation of the tautomers' preferences [28,29], suggests that the two imidazole nitrogen atoms (topologically not equivalent) are however electronically similar to a large extent. It is expected that this effect is due to the imidazole heterocyclic ring's aromaticity [30]. One of the aims of our study is an estimation of the degree of this apparent similarity of the electronic properties within the imidazole ring. For this purpose, we also consider the effect of water as a solvent, and the resulting charge flow within the histidine residue. The transition from the gas to the solution phase can be modeled by the explicit microsolvation, and in this study we investigate the histidine in its three protonation states, microsolvated with 1:6 water molecules. These clusters are studied computationally in the gas phase, and are based on implicit solvent (water, using the Integral Equation Formalism variant of the Polarizable Continuum Model, IEF-PCM) [14] within the density functional theory (DFT) framework [31,32]. The points of our interest include the following: structural analysis with emphasis on the presence of intra- and intermolecular hydrogen bonds and short contacts; the analysis of the electron density within the Atoms-in-Molecules (AIM) theory [33]; and, finally, the Symmetry-Adapted Perturbation Theory (SAPT) interaction energy [34] analysis, which is used to determine and explain the stability of the clusters containing microsolvated histidine. We have recently used the SAPT methodology to investigate the microsolvation of biotin [35], and the current study continues this type of analysis.

The microsolvation of histidine was the subject of two recent studies [36,37]. One of them [36] is concerned with a thorough exploration of the microsolvation patterns, but the overwhelming number of possible orientational isomeric forms led the authors to consider only one or two water molecules as the microsolvation shell. On the other hand, the more recent study [37] explores the neutral and zwitterionic forms of histidine surrounded by up to six water molecules. Both these

studies are concerned mostly with structures, tautomeric transitions, vibrational signatures and—to some extent—the AIM analysis (in [36]). Our intent is different: taking a small subset of the rich conformational space of microsolvated clusters, we would like to observe the effect and nature of the changes induced in the electronic structure of histidine by the presence of a water micro-environment.

2. Materials and Methods

The model of histidine HIP, HIE and HID forms (protonated respectively at both imidazole sites for HIP, at N δ for HID, at N ϵ for HIE, see Figure 1) was constructed on the basis of its derivative deposited in the Cambridge Crystallographic Data Centre (CCDC) [38], denoted as ACHIST21 [39] with the deposition number 1840707. Next, the Density Functional Theory (DFT) [31,32] was applied for the quantum-mechanical simulations. The energy minimizations were performed, on the basis of the DFT functional ω B97XD [40] with the triple-zeta split valence basis set 6-311+G(d,p), by Pople [41]. The energy minimizations were performed in the gas phase and with the solvent reaction field for the neutral (protonated at either N δ or N ϵ) and protonated (cationic) forms of histidine. The standard structural convergence criteria (maximum force component less than 0.00045 a.u., root-mean-square force less than 0.0003 a.u., maximum step component less than 0.0018 a.u., and root-mean-square step less than 0.0012 a.u.) were employed. In order to reproduce the solvent's influence on the molecular structure of histidine, the Polarizable Continuum Model (PCM) [14], in its IEF-PCM formulation, and water as a solvent were used. Beside the geometry optimization, the harmonic frequencies were calculated for both phases to confirm that the obtained structures correspond to the minimum on the Potential Energy Surface (PES).

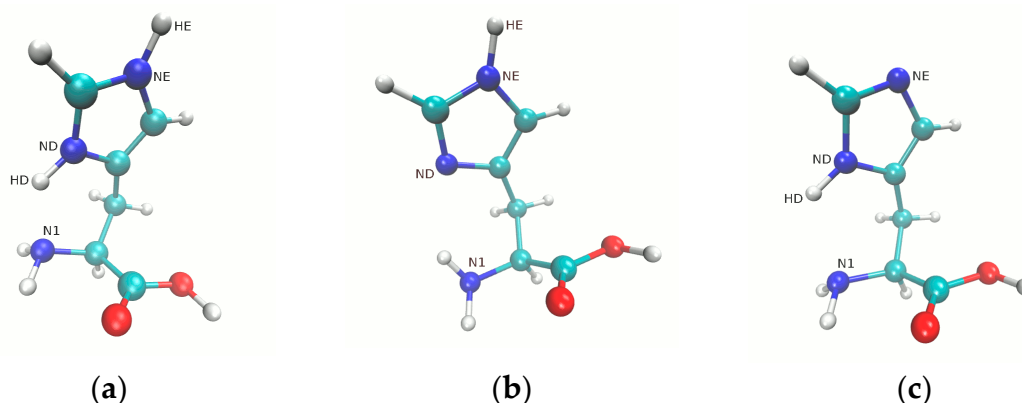


Figure 1. Structures of the studied forms of histidine: (a) N δ -H and N ϵ -H protonation, HIP; (b) N ϵ -H protonation, HIE; (c) N δ -H protonation, HID. Color coding: carbon, cyan; nitrogen, dark blue; oxygen, red; hydrogen, white.

The second part of the study is devoted to the investigation of the intermolecular interactions. The three series of histidine–water (one to six water molecules) complexes were constructed for this purpose, assuming the cationic (protonated) form of histidine or two neutral forms of histidine protonated at either N δ or N ϵ . The models were prepared with the Molden program's assistance [42] by addition of water molecules to the histidine in the places indicated by the presence of hydrogen bond donors/acceptors, in a manner similar to the procedure described in [36]. The energy minimizations of the obtained models were performed according to the procedure described above. The simulations were carried out in two phases as well. The Atoms-in-Molecules (AIM) theory [33] was applied to study the topology of the complexes, as well as to confirm the presence of the intra- and intermolecular short contacts and hydrogen bonds. The wavefunctions for the Atoms-in-Molecules (AIM) investigations were computed using ω B97XD [40] functional and 6-311+G(d,p) basis set [41].

The Supplementary Material contains also the validation of the computational model; the results of optimizations with the 6-311+G(d,p) basis set and a rich set of DFT functionals; the B3LYP [43,44]

and PBE [45] for the gas phase, followed by PCM studies with dispersion-corrected B3LYP-D3 [46] and the Minnesota M06-2X functional [47]. These data serve only as a test proving that the choice of the DFT functional induces quantitative, not qualitative, changes in the solvation shell.

The interaction energy decomposition in the investigated complexes was performed using the Symmetry-Adapted Perturbation Theory (SAPT) [34] at the SAPT2 level [48]. The interaction energy at the SAPT2 level is defined in the following manner—see also Equations (1) to (8) in [48]:

$$E_{int,SAPT2} = E_{elst}^{10} + E_{exch}^{10} + E_{ind,r}^{20} + E_{ex-ind,r}^{20} + \delta E^{HF} + E_{disp}^{20} + E_{ex-disp}^{20} + E_{elst,r}^{12} + E_{exch}^{11} + E_{exch}^{12} + {}^tE_{ind}^{22} + {}^tE_{ex-ind}^{22} \quad (1)$$

where perturbational notation is used in the superscript, e.g., the $E_{elst,r}^{12}$ term is an electrostatic term arising from the first-order intermolecular and second-order intramolecular perturbations. The types of terms are electrostatics, exchange (Pauli repulsion), induction (polarization) and dispersion, denoted in the subscripts as *elst*, *exch* (*ex*), *ind* and *disp* respectively. The δE^{HF} term contains mostly higher-order corrections (especially for the interactions of induced multipoles), and together with the preceding four contributions, yields the Hartree–Fock interaction energy. The subscript *r* indicates that this particular term is calculated taking into account orbital relaxation (response) effects—see Equation (118) in [34].

The SAPT2 calculations were preceded by reoptimizations of the histidine–water complexes at the MP2/aug-cc-pVTZ level [49,50], using the density fitting acceleration of the MP2 scheme (DF-MP2) [51]. The optimizations were concluded with the interaction energy estimation at the DF-MP2/aug-cc-pVTZ level using the supermolecular approach, and the Boys–Bernardi counterpoise correction [52] for the basis set superposition error, taking the histidine as one “monomer” and a set of water molecules as the second “monomer” of the “dimer”. Unfortunately, the computational demands of the SAPT2 scheme did not allow us to compute the interaction energies for the complexes with 5 or 6 water molecules. The structures optimized at the DF-MP2/aug-cc-pVTZ level served then to calculate supermolecular DF-MP2 and SAPT2 interaction energies, using a smaller aug-cc-pVDZ basis set. The calculations with the smaller basis set were performed for all complexes, so that we could estimate the basis set effect on the SAPT2 results.

The quantum-mechanical simulations were performed using the Gaussian 16 (G16) suite of programs [53], while the AIM calculations were done on the basis of AimAll [54] program. The Psi4.1.2.1 [55] program was applied to carry out the SAPT calculations and preceding DF-MP2 optimizations. The visualizations of the structures of histidine and its complexes were prepared using the VMD 1.9.3. [56] and Gimp [57] programs. Remark: the N δ and N ϵ imidazole nitrogen atoms of histidine are further denoted as ND and NE for simplicity, in a manner similar to the standard PDB atom naming scheme. The hydrogen atoms connected to them are denoted as HD and HE in the Figures 1 and 2 presented in the study. Only atoms of interest are indicated in the Figures 1 and 2, and their numbering scheme was prepared only for the study. Depending on the protonation and tautomeric state, the histidine moiety is labeled as HIP (cationic protonated state), HID (N δ -H protonation) and HIE (N ϵ -H protonation), using the convention of the Amber force fields.

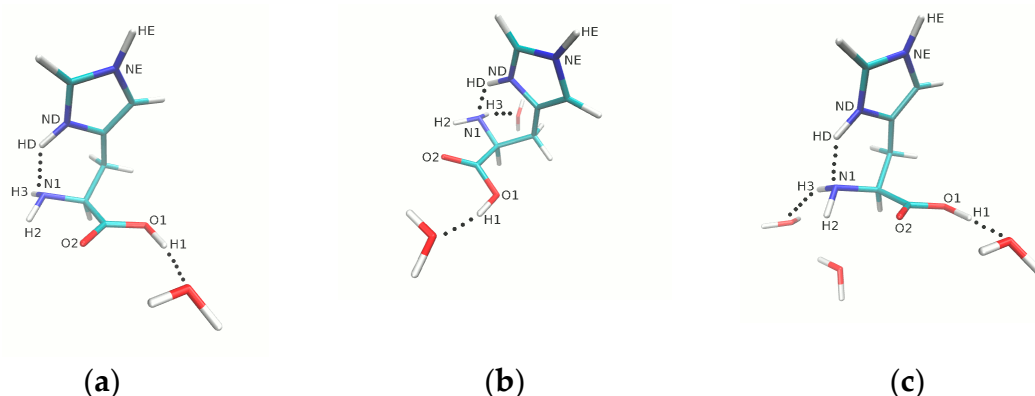


Figure 2. Schematic presentation of microsolvation models for histidine HIP form. For clarity, only the models with (a) 1, (b) 2, and (c) 3 water molecules are shown. The dotted lines indicate the presence of selected intra- and intermolecular short contacts. Color coding: carbon, cyan; nitrogen, dark blue; oxygen, red; hydrogen, white.

3. Results and Discussion

3.1. Metric Parameters Analysis

The structures of the investigated histidine forms (HIP, HIE and HID) are presented in Figure 1. The DFT simulations for a single molecule of the analyzed three forms were performed in the gas phase, and using the PCM solvation model. The ω B97XD/6-311+G(d,p) level of theory used in this study is very similar to the level used by Lee et al. [37], who chose the 6-311++G(d,p) basis set. The triple-zeta valence basis set, with diffuse and polarization functions, is, on the one hand, already sufficient for the DFT (a more flexible basis set would be chosen for post-Hartree–Fock approaches), and on the other hand, small enough to allow massive calculations in the future studies, for example when analyzing hundreds of snapshots from molecular dynamics trajectories. Further, the ω B97XD functional was chosen due to its performance in the description of non-covalent interactions, especially in the field of the microsolvation of amino acids [37].

The summary of the short contacts found between the imidazole ring and the backbone nitrogen (N1) atom forming a quasi-ring is presented in Table 1, while Table S1I presents an additional dataset obtained with diverse DFT functionals.

In the case of histidine HIP and HID forms, an intramolecular hydrogen bond is present (which was also confirmed by the AIM theory; see text below for more details). However, because of the steric effects, the hydrogen bond is not a linear one. The HIE form does not contain an intramolecular hydrogen bond due to a lack of the corresponding hydrogen atom (HD atom) in the imidazole ring. There is a possible interaction between the nitrogen (ND) atom from the imidazole ring with the NH_2 group. The ND...N1 interatomic distance is elongated upon the presence of the polar environment in the case of the histidine HIP form, while for the neutral HID and HIE tautomers, it is shortened upon the transition from the gas phase to PCM. There is no doubt that the presence of the intramolecular short contacts stabilizes the conformation of the histidine residue.

The second part of the analysis is devoted to the histidine (HIP, HIE and HID forms) surrounded by water molecules. The graphical representation of the microsolvation models is presented in the following figures: Figure 2 of the text body, and Figures S2I–S5I of the Supplementary Material. The metric parameters of histidine interacting with water molecules are listed in Table 2 and Tables S2I–S4I. The interactions between water molecules were not taken into account in these tables.

Table 1. Metric parameters (interatomic distances are given in [Å]) of intramolecular short contacts obtained as a result of DFT gas phase and with solvent reaction field (PCM with water as a solvent) simulations with ω B97XD functional and 6-311+G(d,p) basis set. For geometric details see Figure 1.

Metric Parameters	Gas Phase	PCM
HIP		
ND ... N1	2.713	2.769
ND-HD	1.036	1.026
HD ... N1	1.917	2.030
\angle NHN [°]	131.2	126.9
HIE		
ND ... N1	3.219	3.160
N1-H	1.012	1.017
H ... ND	2.816	2.411
\angle NHN [°]	104.2	129.8
HID		
ND ... N1	2.905	2.872
ND-HD	1.011	1.014
HD ... N1	2.247	2.197
\angle NHN [°]	121.4	122.5

Table 2. The metric parameters values obtained for the microsolvation models. The simulations were performed in the gas phase with 1–6 water molecules and with the continuum PCM model (water as a solvent) at the ω B97XD/6-311+G(d,p) DFT level of theory. Only selected interactions with histidine HIP form are presented. Interatomic distances are given in [Å], angles in degrees.

Metric Parameters	Histidine HIP Form	
	HIP with 1–6 Water Molecules	HIP with 1–6 Water Molecules and with PCM
One water molecule (HIP-1)		
Intramolecular HB		
ND...N1	2.706	2.763
ND-HD	1.038	1.027
HD...N1	1.901	2.015
\angle NHN [°]	131.9	127.5
Intermolecular HB		
O1...O	2.644	2.671
O1-H1	0.992	0.992
H1...O	1.671	1.679
\angle O1H1O [°]	165.7	178.0
Two water molecules (HIP-2)		
Intramolecular HB		
ND...N1 [Å]	2.668	2.737
ND-HD	1.048	1.030
HD...N1	1.824	1.966
\angle NHN [°]	134.7	129.3
Intermolecular HB		
O1...O	2.651	2.676
O1-H1	0.991	0.991
H1...O	1.685	1.686
\angle O1H1O [°]	163.6	178.1
N1...O	2.996	3.005
N1-H3	1.018	1.018
H3...O	2.067	2.052
\angle N1H3O [°]	150.5	154.9

Table 2. Cont.

Metric Parameters	Histidine HIP Form	
	HIP with 1–6 Water Molecules	HIP with 1–6 Water Molecules and with PCM
Three water molecules (HIP-3)		
Intramolecular HB		
ND...N1	2.644	2.739
ND-HD	1.055	1.030
HD...N1	1.780	1.968
\angle NDHDN [°]	136.2	129.3
Intermolecular HB		
O1...O	2.661	2.677
O1-H1	0.989	0.991
H1...O	1.703	1.686
\angle O1H1O [°]	161.7	178.2
N1...O	3.225	3.081
N1-H3	1.017	1.018
H3...O	2.300	2.091
\angle N1H3O [°]	150.6	163.6
Four water molecules (HIP-4)		
Intramolecular HB		
ND...N1 [Å]	2.699	2.750
ND-HD	1.039	1.028
HD...N1	1.881	1.990
\angle NDHDN1 [°]	133.0	128.5
Intermolecular HB		
O1...O	2.664	2.676
O1-H1	0.989	0.991
H1...O	1.707	1.686
\angle O1H1O [°]	161.6	177.3
N1...O	3.281	3.092
N1-H3	1.016	1.018
H3...O	2.291	2.074
\angle N1H3O [°]	164.6	180.0
NE...O	2.781	2.773
NE-HE	1.029	1.033
HE...O	1.753	1.741
\angle NEHEO [°]	178.2	177.8
Five water molecules (HIP-5)		
No intramolecular HB		
Intermolecular HB		
O1...O	2.664	2.655
O1-H1	0.990	0.995
H1...O	1.687	1.661
\angle O1H1O [°]	168.1	177.7
N1...O	2.751	2.764
O-H	0.988	0.991
H...N1	1.789	1.785
\angle N1HO [°]	163.5	169.0
NE...O	2.785	2.769
NE-HE	1.029	1.033
HE...O	1.756	1.736
\angle NEHEO [°]	178.1	178.1
ND...O	2.676	2.742
ND-HD	1.051	1.039
HD...O	1.640	1.721
\angle NDHDO [°]	167.6	166.7
O...O2	2.869	2.816
O-H	0.970	0.971
H...O2	1.910	1.851
\angle OHO2 [°]	169.5	171.7

Table 2. Cont.

Metric Parameters	Histidine HIP Form	
	HIP with 1–6 Water Molecules	HIP with 1–6 Water Molecules and with PCM
	Six water molecules (HIP-6)	
	No intramolecular HB	
Intermolecular HB		
O1...O	2.696	2.643
O1-H1	0.986	0.998
H1...O	1.741	1.659
\angle O1H1O [°]	161.7	168.1
N1...O	2.754	2.780
O-H	0.990	0.990
H...N1	1.789	1.801
\angle N1HO [°]	163.8	169.3
NE...O	2.811	2.773
NE-HE	1.023	1.022
HE...O	1.884	1.910
\angle NEHEO [°]	149.2	140.1
ND...O	2.673	2.743
ND-HD	1.050	1.037
HD...O	1.652	1.746
\angle NDHDO [°]	162.6	160.0
O...O2	2.837	2.819
O-H	0.971	0.971
H...O2	1.880	1.855
\angle OHO2 [°]	168.0	171.7

Let us start the discussion of the microsolvation from the histidine HIP form. The selected metric parameters obtained as a result of the DFT simulations are presented in Table 2 and Table S2I. As is visible, the number of water molecules, and the inclusion of the polar environment due to the application of the continuum solvation model, influences the values of the metric parameters. The intramolecular hydrogen bond is not present when we introduced five and six water molecules into the vicinity of the ND-HD bond and NH₂ group. This critical number of water molecules corresponds well with the observation in [37] that six water molecules are necessary to stabilize the zwitterionic form. The study of Rai et al. [36] only includes up to two water molecules, but already significant polarization effects are visible in the histidine residue. The introduction of water molecules to the gas phase models shortened the ND ... N1 interatomic distance. The intramolecular HD ... N1 hydrogen bond was shortened when the number of water molecules was added to the model (one to four water molecules). Concerning the results of the simulations with the application of the PCM model, the intramolecular ND-HD ... N1 interaction was not observed with five or six water molecules in the models. The ND ... N1 interatomic distance and the HD ... N1 hydrogen bond were both elongated.

The data in Table 2 reveal also some peculiarities in the network of hydrogen bonds connecting the protonated HIP residue with water molecules. A monovalent ion dissolved in water acts as a factor governing the orientation and location of the surrounding water shell. For cations, we expect that the water oxygen atoms are directed towards the cation, while the hydrogen atoms should face the dissolved anionic moiety. A simple (spherical) cation, such as Na⁺, would form contacts of similar lengths with diverse water molecules. The irregularly-shaped imidazole ring does not have to follow this line of reasoning, and indeed, the ND ... O and NE ... O distances differ quite significantly in the gas phase HIP-5 and HIP-6 complexes. This behavior changes sharply when the PCM model is used in addition to microsolvation—the two above-mentioned distances become quite similar. This symmetrization of the microsolvation shell is connected with the symmetrization of the imidazole ring's electronic properties, but it also signifies that even six water molecules is not enough to saturate the polarization of the HIP residue.

Table S3I summarizes selected metric parameters of the histidine HIE tautomer. As we could expect, there is a difference in the interatomic distances and angle values of the atoms involved in the intermolecular hydrogen bonds' formation. The PCM's implicit solvation influence is noticeable

when comparing the gas phase and solvent reaction field simulations. For example, as is shown in the Table S3I, the O-H ... O2 intermolecular hydrogen bond was noted in the gas phase simulations, but it was not present for the optimized model with the assistance of the PCM. A similar conclusion could be drawn for the histidine HID tautomer (Table S4I). Comparing the values of the metric parameters obtained after two-phase simulations with various numbers of water molecules, one can observe changes in the interatomic distances and the valence angle value. Moreover, as is shown in Tables S2I–S4I, some interactions (intermolecular hydrogen bonds) are present in the gas phase models, but they were not recognized after the geometry optimization with the PCM solvation model. This signifies that even six water molecules do not saturate the polarization of the histidine residue, and either more explicit water molecules or some form of implicit model (e.g., PCM) is required. Thus, our computational results and metric parameters analysis provided a deeper insight into the histidine residue–water molecules interactions.

3.2. Atoms-In-Molecules (AIM) Analysis

The analysis of the electron density via the framework of the AIM theory was performed in this study for several reasons. First, the topology of the electron density field reveals the presence of intra- and intermolecular interactions, especially hydrogen bonds. Second, the electron density flow can be evaluated by comparing the AIM net atomic charges in the gas phase and solvent reaction field (PCM model) simulations. Third, this also allows an estimation of the histidine charge transfer from or towards the solvent. These three issues will be discussed below.

The topology of the electron density field is shown schematically in the Figures S8I–S10I in the Supplementary Material. The depicted elements of the molecular graphs are bond paths (solid or dashed), bond critical points (BCPs—green spheres), and less frequent ring critical points (RCPs) or cage critical points (CCPs). The bond paths connect the interacting atoms, and it is important to note that the use of the solid vs. the dashed line is based on the simple criterion of the electron density at the BCP. Thus, one can appreciate that some of the hydrogen bonds paths are marked as a solid bond path, which suggests that these bonds are stronger and important, and exemplified in molecular conformations. This especially concerns the intramolecular ND...N1 contacts, but also some of the carboxylic O1-H1...O (water) interactions. On the other hand, some of the bond paths are significantly non-linear (bent), and they connect rather surprising atom pairs (e.g., two oxygen atoms) (see especially the interactions of the carboxylic O2 atom in the first two panels of Figure S8I). Moreover, in these cases the corresponding BCPs and RCPs are located in close proximity. Even a small structural change could potentially lead to their coalescence (annihilation) and the topological change of the molecular graph. A canonical study of Koch and Popelier on the issue of the AIM-based detection of hydrogen bonds [58] includes correct topology of the electron density as the first criterion to be met by a hydrogen-bonded atom pair. This is an indication that in our case these interactions are, from the point of view of the AIM theory, rather weak, and not of the hydrogen bond type. Their presence stems rather from the purely mathematical necessity of fulfilling the Poincaré–Hopf formula, which is also the case for sterically crowded H ... H interactions [59].

Now, we will cover the results of the numerical integration of electronic density within atomic basins, which yields AIM net atomic charges. First, let us note that in all the covered cases (18 structures in both gas phase and PCM with water as a solvent), the sum of all the atomic charges deviates from the ideal value (either +1 or 0) by no more than 0.003 e. Such accuracy of the total integration assures us that the charge flow of magnitude larger than 0.01 e is significant, and not due to numerical integration errors. With this in mind, we will turn to the results gathered in Tables 3–5. The last column of these tables, which is the sum of the AIM net charges for the histidine residue, is an indicator of the charge transfer between the histidine and the microsolvation shell of water molecules. In the case of the cationic histidine HIP (Table 3), the cation acts as an electrophile, and the electron density is transferred from the water molecules via hydrogen bonds to the amino acid. The magnitude of this charge transfer is between 0.04 and 0.11 e, above the assumed “significance

threshold" of 0.01 e, and it changes with no obvious correlation with the microsolvated cluster size or the number of hydrogen bonds.

Table 3. AIM-derived net atomic charges (in the elementary charge units, e) for the selected interaction centers of microsolvated histidine in its protonated form (HIP). Values in bold typeface indicate atoms engaged in a hydrogen bond. Last column: the sum of AIM net charges for the histidine residue.

HIP-(H ₂ O) _n	Simulation Environment	ND	HD	NE	HE	O2	Mol. Charge
HIP-1 n = 1	Gas phase	−1.246	0.537	−1.206	0.476	−1.179	0.958
	PCM H ₂ O	−1.240	0.528	−1.209	0.496	−1.203	0.952
	Δ(PCM-gas)	0.006	−0.009	−0.003	0.020	−0.024	
HIP-2 n = 2	Gas phase	−1.251	0.544	−1.207	0.473	−1.182	0.946
	PCM H ₂ O	−1.244	0.531	−1.209	0.495	−1.205	0.939
	Δ(PCM-gas)	0.007	−0.013	−0.002	0.022	−0.023	
HIP-3 n = 3	Gas phase	−1.252	0.549	−1.207	0.472	−1.184	0.953
	PCM H ₂ O	−1.244	0.531	−1.209	0.495	−1.206	0.933
	Δ(PCM-gas)	0.008	−0.018	−0.002	0.024	−0.022	
HIP-4 n = 4	Gas phase	−1.254	0.533	−1.249	0.539	−1.190	0.916
	PCM H ₂ O	−1.246	0.526	−1.249	0.545	−1.206	0.893
	Δ(PCM-gas)	0.008	−0.008	0.000	0.006	−0.016	
HIP-5 n = 5	Gas phase	−1.253	0.553	−1.256	0.538	−1.207	0.927
	PCM H ₂ O	−1.252	0.541	−1.255	0.547	−1.212	0.943
	Δ(PCM-gas)	0.001	−0.012	0.001	0.008	−0.004	
HIP-6 n = 6	Gas phase	−1.251	0.552	−1.246	0.517	−1.201	0.938
	PCM H ₂ O	−1.248	0.540	−1.239	0.529	−1.211	0.956
	Δ(PCM-gas)	0.003	−0.012	0.007	0.012	−0.010	

The erratic behavior observed for the cationic HIP is also present for both tautomeric forms of neutral histidine (Tables 4 and 5), but in these cases the amino acid acts most frequently as a weak donor of the electron density (nucleophile), and transfers between 0.02 and 0.09 e to the water shell. Only for the cases where one or two water molecules are in contact with the HID tautomer does the histidine act as an electrophile. This diversity and the ease of modification of the electron-donating or -accepting properties is in good agreement with the role of histidine as a versatile amino acid, capable of acting as, for example, a proton relay in catalytic triads of many enzymes.

A detailed understanding of the interplay between the solvent and the histidine requires an analysis of the charge flow around the selected interaction centers of the histidine. We have chosen to investigate more deeply three interaction centers: the vicinity of the imidazole nitrogen atoms, and the carboxylic oxygen atom O2 (retaining its C=O double bond, thus more similar to the situation in a protein backbone). By this method, we also investigate a possible similarity between the properties of the two imidazole nitrogen atoms of histidine. These atoms, N δ and N ϵ (denoted ND and NE, respectively), are not symmetrically equivalent from the point of view of the molecular topology, but the presence of a strong electron delocalization in the heterocyclic aromatic ring allows us to expect the electronic behavior of the ND and NE atoms to be very similar. This similarity should be best seen in the protonated histidine (HIP, Table 3), where ND and NE are in the same protonated state. However, the microsolvation with one, two or three water molecules results in the formation of an intramolecular hydrogen bond between the ND-HD moiety and the backbone amine nitrogen atom, while the NE-HE atoms are not hydrogen-bonded. We can see that the presence of the hydrogen bonding results in a significant polarization; for example, in the gas phase HIP-1, the net charges for ND and NE are -1.246 e and -1.206 e, respectively, while the net charges for HD and HE are 0.537 e and 0.476 e. An inclusion of the PCM solvent reaction field reduces this polarization, and the charges of the corresponding atoms are more similar to each other. In agreement with the notion that explicit solvation saturates the capabilities of the donor/acceptor moiety, we see that when a particular moiety (ND-HD or NE-HE) is engaged in a hydrogen bond, it is less affected by the PCM implicit solvent. The most external atoms, in this case HD and HE, are most affected by the PCM solvation, but—roughly summarizing—the effect of PCM

implicit water is a 0.02 e net charge decrease of the unsaturated (not-hydrogen-bonded) hydrogen atom, while this decrease is only ca. 0.01 e when such an atom is between explicit donor and acceptor atoms. This is also visible for the carboxyl O2 atom, which, when not engaged in explicit hydrogen bonding, is made more negative (by ca. 0.02 e) under the PCM solvation, but this polarization is not greater than 0.01 e in HIP-5 and HIP-6, where this atom participates in an explicit hydrogen bonding.

Table 4. AIM-derived net atomic charges (in the elementary charge units, e) for the selected interaction centers of the microsolvated neutral HIE tautomer of histidine. Values in bold typeface indicate atoms engaged in a hydrogen bond. Last column: the sum of AIM net charges for the histidine residue.

HIE-(H ₂ O) _n	Simulation Environment	ND	NE	HE	O2	Mol. Charge
HIE-1 n = 1	Gas phase	−1.150	−1.227	0.424	−1.183	0.047
	PCM H ₂ O	−1.185	−1.231	0.459	−1.207	0.059
	Δ(PCM-gas)	−0.036	−0.004	0.035	−0.024	
HIE-2 n = 2	Gas phase	−1.156	−1.228	0.424	−1.174	0.058
	PCM H ₂ O	−1.188	−1.232	0.458	−1.198	0.067
	Δ(PCM-gas)	−0.031	−0.003	0.035	−0.024	
HIE-3 n = 3	Gas phase	−1.164	−1.263	0.486	−1.176	0.041
	PCM H ₂ O	−1.194	−1.269	0.508	−1.199	0.039
	Δ(PCM-gas)	−0.030	−0.005	0.023	−0.023	
HIE-4 n = 4	Gas phase	−1.188	−1.263	0.492	−1.176	0.062
	PCM H ₂ O	−1.203	−1.267	0.513	−1.199	0.089
	Δ(PCM-gas)	−0.015	−0.003	0.021	−0.023	
HIE-5 n = 5	Gas phase	−1.180	−1.255	0.493	−1.222	0.051
	PCM H ₂ O	−1.198	−1.259	0.505	−1.225	0.058
	Δ(PCM-gas)	−0.017	−0.004	0.012	−0.003	
HIE-6 n = 6	Gas phase	−1.176	−1.253	0.471	−1.232	0.021
	PCM H ₂ O	−1.196	−1.256	0.498	−1.229	0.052
	Δ(PCM-gas)	−0.020	−0.003	0.027	0.003	

Let us conclude with the possibility of similarity between the electronic properties of ND-HD and NE-HE; indeed, for the cluster sizes of four or more water molecules, when both moieties are engaged in explicit hydrogen bonding, the atomic charges of the nitrogen atoms are very similar (differing by no more than 0.009 e), indicating an equalization of their electronic properties. The HD and HE hydrogen atoms are not as similar; the net charge differences between them are on average 0.015 e, and can be as large as 0.035 e, but the screening effect of the explicit solvation is still visible. These results agree well with the structural equalization of the N-H ... O distances between the histidine imidazole ring and the water shell, induced by the use of the PCM model, as discussed in Section 3.1.

The two neutral tautomers, protonated at either the NE (HIE tautomer, Table 4) or ND (HID tautomer, Table 5) nitrogen atom, mostly follow the rules of the behavior outlined above for the cationic HIP. In these two cases, however, it is possible to examine the polarizability of the nitrogen atom with its lone pair, a potential hydrogen bond acceptor. Due to a very flexible electron distribution, enhanced by the lone pair, the non-protonated nitrogen atoms (ND for HIE, NE for HID) under the PCM water solvent model become more negatively charged: by ca. 0.035 e for ND in HIE, and by ca. 0.06 e for NE in HID. This almost twofold difference between the properties of ND and NE shows that, despite the aromatic delocalization of the electron density in the imidazole ring, these nitrogen atoms are not fully equivalent in the electrostatic sense. Finally, it is evident that the protonation of one of the imidazole nitrogen atoms makes this atom polarized; for example, for the HIE-1 structure (Table 4, no hydrogen bonds to the considered atoms), the protonated NE atom has a net charge more negative by 0.08 e than the non-protonated ND atom. The corresponding difference between ND(-HD) and NE in HID-1 (Table 5) is even larger (0.11 e). Summarizing the AIM part of the study, we would like to stress that the PCM model does not provide the same large degree of polarization, which is obtained with the explicit microsolvation and hydrogen bonding. On the other hand, even with six water molecules, the histidine residue is still visibly polarized by the PCM environment. Thus,

it seems advisable to use in further studies a combination of explicit microsolvation and the PCM implicit solvent.

Table 5. AIM-derived net atomic charges (in the elementary charge units, e) for the selected interaction centers of the microsolvated neutral HID tautomer of histidine. Values in bold typeface indicate atoms engaged in a hydrogen bond. Last column: the sum of AIM net charges for the histidine residue.

HID-(H ₂ O) _n	Simulation Environment	ND	HD	NE	O2	Mol. Charge
HID-1 n = 1	Gas phase	−1.246	0.459	−1.134	−1.194	−0.025
	PCM H ₂ O	−1.252	0.480	−1.192	−1.208	−0.043
	Δ(PCM-gas)	−0.006	0.021	−0.058	−0.014	
HID-2 n = 2	Gas phase	−1.266	0.478	−1.133	−1.193	−0.014
	PCM H ₂ O	−1.254	0.482	−1.193	−1.210	−0.051
	Δ(PCM-gas)	0.012	0.004	−0.060	−0.017	
HID-3 n = 3	Gas phase	−1.254	0.475	−1.165	−1.195	0.015
	PCM H ₂ O	−1.250	0.488	−1.210	−1.207	0.017
	Δ(PCM-gas)	0.004	0.013	−0.045	−0.012	
HID-4 n = 4	Gas phase	−1.260	0.495	−1.166	−1.193	0.032
	PCM H ₂ O	−1.258	0.495	−1.192	−1.209	0.010
	Δ(PCM-gas)	0.002	0.001	−0.026	−0.016	
HID-5 n = 5	Gas phase	−1.260	0.497	−1.168	−1.209	0.035
	PCM H ₂ O	−1.251	0.488	−1.194	−1.212	0.023
	Δ(PCM-gas)	0.008	−0.009	−0.026	−0.002	
HID-6 n = 6	Gas phase	−1.262	0.507	−1.174	−1.206	0.022
	PCM H ₂ O	−1.249	0.490	−1.192	−1.213	0.028
	Δ(PCM-gas)	0.013	−0.017	−0.018	−0.008	

3.3. SAPT Analysis of Interaction Energies in the Microsolvated His-(H₂O)_n Complexes

The reliable assessment of the interaction energies, E_{int} , within the histidine–water complexes has been carried out using the Symmetry-Adapted Perturbation Theory treatment. This approach, although it cannot account for the PCM solvation, is one of the important tools for studying the nature of intermolecular forces. The SAPT level employed in this study is SAPT2, which is roughly equivalent to the MP2 calculus, but it allows for the calculation of separate contributions to the interaction energy. First, we have reoptimized the DFT structures using the MP2/aug-cc-pVTZ level of the theory, and then we used the MP2 structures for the SAPT2 calculations. However, we have found that the calculations at the SAPT2/aug-cc-pVTZ level were not possible for the systems with five or six water molecules. Thus, we have retained the MP2/aug-cc-pVTZ geometries, but we have estimated the interaction energies with the MP2 and SAPT2 techniques with smaller aug-cc-pVDZ basis sets. All these results are summarized in Table S3I. In view of the need to resort to a smaller basis set, it is necessary to verify that the results with smaller and larger basis sets are compatible. This is true in the case of the microsolvated histidine—the MP2 interaction energies obtained with the double-zeta and triple-zeta basis sets are correlated with an r^2 correlation coefficient larger than 0.9998, and the larger basis set yields interactions ca. 8% stronger than the smaller basis. The same degree of correlation is found for the SAPT2 results ($r^2 > 0.9995$), and the triple-zeta basis set yields 10% stronger interactions than does the double-zeta SAPT2 level. Moreover, the double-zeta MP2 and SAPT2 interaction energies are very strongly correlated ($r^2 > 0.9999$), and the corresponding MP2 and SAPT2 results differ by no more than 2.3% (0.35% on average). These results confirm that the total SAPT2 energy corresponds closely to the MP2 result, and that the smaller basis set yields results strongly compatible with the aug-cc-pVTZ basis. It should thus be safe to assume that the SAPT2/aug-cc-pVDZ results are also of good quality in case of the His-(H₂O)₅ and His-(H₂O)₆ clusters, where we do not have the triple-zeta results.

Establishing the consistency of our procedure, we will now focus on the effect of enlarging the microsolvated His-(H₂O)_n cluster. For this purpose, Table 6 contains the values of the interaction energy divided by the number of water molecules in a given system. In all the three protonation states

of histidine, the first H₂O molecule contributes the most to the stability of the complex, and further water molecules bring smaller energetic gains. However, this trend is not monotonic, and deviations are most prominent for the HIP and HIE at $n = 5$ water molecules. At $n = 6$ we already observe the saturation of the hydrogen bond donors/acceptors in the histidine residue, and we do not expect the “per water molecule” contributions to grow significantly. However, due to the inability of the SAPT scheme to deal with the PCM formalism, we cannot establish the degree of polarization incurred by the continuum solvent model.

Table 6. The Basis Set Superposition Error-corrected MP2 interaction energies and SAPT2 interaction energies (total per complex, and per one water molecule) calculated for the structures of histidine-(H₂O)_{*n*} ($n = 1$ –6) complexes optimized at the MP2/aug-cc-pVTZ level of theory. Nomenclature: HIP-*n*—protonated histidine, HID-*n*—the N δ -H tautomer, HIE-*n*—the N ϵ -H tautomer, *n*—number of water molecules. All values in kcal/mol.

	MP2/aug-cc-pVDZ		MP2/aug-cc-pVTZ		SAPT2/aug-cc-pVDZ		SAPT2/aug-cc-pVTZ	
	E _{int}	per 1 H ₂ O	E _{int}	per 1 H ₂ O	E _{int}	per 1 H ₂ O	E _{int}	per 1 H ₂ O
HIP-1	−12.50	−12.50	−13.49	−13.49	−12.57	−12.57	−13.79	−13.79
HIP-2	−21.03	−10.52	−22.47	−11.23	−21.02	−10.51	−22.76	−11.38
HIP-3	−23.87	−7.96	−25.43	−8.48	−24.11	−8.04	−25.99	−8.66
HIP-4	−37.99	−9.50	−40.23	−10.06	−37.88	−9.47	−40.54	−10.13
HIP-5	−57.64	−11.53	−61.90	−12.38	−57.87	−11.57		
HIP-6	−58.74	−9.79	−62.72	−10.45	−58.78	−9.80		
HIE-1	−8.02	−8.02	−8.74	−8.74	−7.98	−7.98	−8.89	−8.89
HIE-2	−12.84	−6.42	−14.16	−7.08	−12.78	−6.39	−14.44	−7.22
HIE-3	−18.87	−6.29	−20.58	−6.86	−18.70	−6.23	−20.85	−6.95
HIE-4	−23.38	−5.85	−25.63	−6.41	−23.43	−5.86	−26.26	−6.57
HIE-5	−38.97	−7.79	−42.25	−8.45	−39.14	−7.83		
HIE-6	−40.74	−6.79	−44.39	−7.40	−41.33	−6.89		
HID-1	−10.09	−10.09	−11.19	−11.19	−10.31	−10.31	−11.70	−11.70
HID-2	−14.93	−7.46	−16.56	−8.28	−15.16	−7.58	−17.15	−8.58
HID-3	−21.62	−7.21	−23.64	−7.88	−21.87	−7.29	−24.41	−8.14
HID-4	−29.12	−7.28	−31.80	−7.95	−29.05	−7.26	−32.29	−8.07
HID-5	−36.72	−7.34	−40.10	−8.02	−36.53	−7.31		
HID-6	−43.71	−7.28	−47.70	−7.95	−43.87	−7.31		

Let us note that, as one could expect from the usual magnitudes of the ion–dipole and dipole–dipole forces, the interactions of the protonated histidine (HIP-1 to HIP-6) are stronger (by 10% to 60%) than their counterparts with neutral tautomers HID and HIE; for example, with one water molecule we obtain an SAPT2 E_{int} for HIP-1, HIE-1 and HID-1 equal respectively to −12.57, −7.98 and −10.31 kcal/mol. This charge-induced strengthening is, curiously, not correlated with the system size.

Unfortunately, the previous studies on the microsolvation of histidine [36,37] are mostly concerned with the tautomeric equilibria, not with the nature of the interactions. We cannot directly compare our calculated interaction energy values with the results of these studies. However, [36] contains data on the estimated strengths of particular hydrogen bonds, evaluated from the electron density using the Espinosa formula [60]. The resulting interactions seem to be stronger than those of our study; for example, HID and HIE hydrated with one water molecule should have a stabilization of ca. −8 to −10.3 kcal/mol at the SAPT2 level, but the estimate given in [36] is close to −16 kcal/mol. The same happens with two water molecules: our SAPT2 results predict E_{int} equal to −13 to −15 kcal/mol, while the data in [36] provide values ranging from −13 to −25 kcal/mol. This comparison cannot be taken too far; the Espinosa formula yields a local parameter, which is an estimate of the energy for a particular bond, with no guarantee that some other region of the interface between molecules is not strongly repulsive.

As noted in the Materials and Methods section, the SAPT approach allows us to identify the physical phenomena behind the interaction energy. A detailed partitioning of E_{int} is given in

Table 7 and Figure 3 (for a quick visual examination of the importance of diverse terms and their relative changes), while Figure S11I shows the percentages of the contributions of four principal terms to the E_{int} , conserving their sign (thus the dimensionless E_{int} is always -1). These terms are: ELST (Coulombic interaction of nuclei and frozen electron densities of the monomers), EXCH (Pauli repulsion of the frozen monomer densities), IND (induction/mutual polarization, taking into account a relaxation of monomer electron densities and associated changes to the Pauli repulsion) and DISP (dispersion contributions)—see Equations (2)–(5) in [48] for a detailed list of the contributions included in these four general interaction energy components. The last term, DISP, is the least sensitive to the histidine protonation state. The permanent and frozen electric moments contribute to the ELST term, but in the studied case of the relatively “soft”, polarizable residue of histidine, this term is mostly cancelled out by the Pauli repulsion EXCH. In absolute values, ELST and EXCH dominate over the effects of polarization and dispersion. On the other hand, a rather good balance between ELST and EXCH makes polarization the most important factor for the actual interaction magnitude. This is especially true for the protonated HIP residue, wherein IND can be two times larger than the dispersion contribution. On average, for the HIP- n complexes, the magnitude of the dispersion term is ca. 60% of the induction, while for the neutral systems DISP is ca. 80% of the IND term. Surprisingly, the ELST term is not dramatically different between the charged and neutral complexes (although HIP-5 and HIP-6 exhibit very large ELST values), and we conclude that the increased stability of the charged complexes results from the ability of the HIP cation to polarize the surrounding network of the water molecules.

Table 7. Interaction energy partitioning at the SAPT2/aug-cc-pVDZ level for the structures of histidine-(H_2O) $_n$ ($n = 1$ –6) complexes optimized at the MP2/aug-cc-pVTZ level. All values in kcal/mol.

	ELST	EXCH	IND	DISP	Total SAPT2
HIP-1	−22.19	24.75	−10.39	−4.74	−12.57
HIP-2	−34.12	33.95	−12.95	−7.89	−21.02
HIP-3	−37.44	36.04	−13.01	−9.70	−24.11
HIP-4	−55.27	50.20	−20.16	−12.66	−37.88
HIP-5	−100.08	110.86	−44.28	−24.38	−57.87
HIP-6	−96.58	102.61	−40.69	−24.12	−58.78
HIE-1	−17.19	20.65	−6.48	−4.96	−7.98
HIE-2	−26.68	32.45	−9.27	−9.28	−12.78
HIE-3	−36.61	42.29	−12.79	−11.59	−18.70
HIE-4	−46.72	55.55	−16.98	−15.28	−23.43
HIE-5	−72.60	84.21	−30.90	−19.85	−39.14
HIE-6	−78.58	94.95	−34.93	−22.78	−41.33
HID-1	−23.95	29.14	−9.87	−5.63	−10.31
HID-2	−31.15	37.99	−11.90	−10.10	−15.16
HID-3	−45.01	52.18	−17.34	−11.69	−21.87
HID-4	−52.01	61.66	−21.06	−17.65	−29.05
HID-5	−65.31	76.76	−26.04	−21.94	−36.53
HID-6	−74.77	88.46	−30.94	−26.63	−43.87

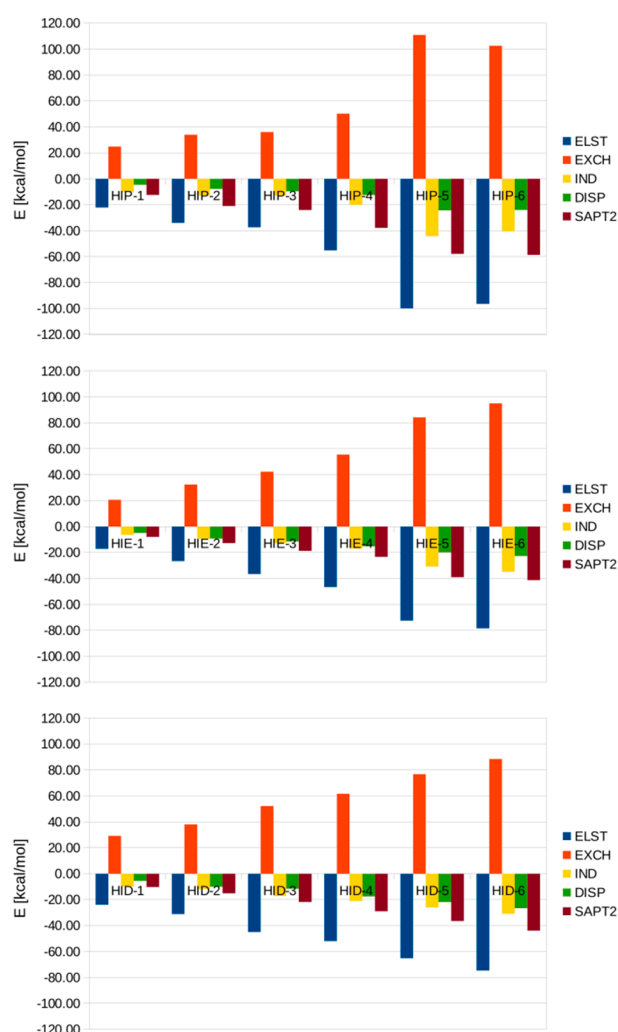


Figure 3. Interaction energy components at the SAPT2/aug-cc-pVDZ level. ELST—electrostatic interaction of frozen monomer electron densities, EXCH—Pauli repulsion of the frozen monomer densities, IND—the effect of mutual polarization and relaxation of monomer electron densities, DISP—dispersion contribution.

4. Conclusions

A detailed study of a histidine molecule microsolvated explicitly by up to six water molecules is here presented, and the effect of implicit PCM solvation with water as a solvent is also included. All three tautomers (neutral HID, HIE and protonated HIP) are considered in our study. The structural investigations show that the explicit solvation with a low number of water molecules leads to the formation of intramolecular short contacts within the histidine residue. Some of the intra- and intermolecular short contacts are of the hydrogen bond nature, which is also visible in the topological study within the AIM framework. We have also shown that the two imidazole nitrogen atoms, which are topologically different, are however quite similar from the point of view of the electronic (charge flow) properties. This symmetrization of the imidazole electronic properties, incurred by the PCM model, is followed by the symmetrization of the solvation shell: the imidazole nitrogen–water oxygen intermolecular distances become very similar for the ND ... O and NE ... O contacts. Symmetrization is also revealed in the convergence of the AIM charges of the two ring nitrogen atoms.

The PCM solvation model was not able to yield the same large degree of polarization of histidine as was obtained with the explicit microsolvation and the presence of hydrogen bonding. On the other hand, even with six water molecules, the histidine residue is still additionally polarized when the PCM

solvation is included. This suggests the use of a combination of explicit microsolvation and the PCM implicit solvent in further studies, and we consider it the most important conclusion of the current study.

The nature of the interactions within the microsolvated clusters was investigated using SAPT partitioning of the interaction energy. A high correlation was found between the results obtained with the aug-cc-pVDZ and aug-cc-pVTZ basis sets, which is important due to the quickly growing memory requirements of the SAPT scheme. We show that the aug-cc-pVDZ basis set can be used to model the larger systems adequately. The origin of the microsolvation cluster's stability was also investigated. While the largest absolute contributions to the interaction energy are formed by the electrostatics and Pauli repulsion, these two effects mostly cancel each other out. This makes polarization the most important factor for the actual interaction energy, especially with regards to the protonated HIP residue. The increased stability of the charged complexes of HIP-1 to HIP-6 was put down to the ability of the HIP cation to polarize the explicit solvent molecules included in the cluster.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-8994/12/7/1153/s1>: a monolithic PDF file containing structural data computed with a rich set of DFT functionals (Tables S1I–S4I), extended visualization of microsolvated structures (Figures S1I–S7I), AIM topological maps (Figures S8I–S10I) and scaled components of the interaction energy (Figure S11I).

Author Contributions: Conceptualization, A.J.; methodology, B.K., J.J.P. and A.J.; investigation, B.K., J.J.P. and A.J.; writing—original draft preparation, J.J.P. and A.J.; writing—review and editing, J.J.P. and A.J.; project administration, A.J. All authors have read and agreed to the published version of the manuscript.

Funding: A.J. and J.J.P. would like to thank the National Science Centre (Poland) for supporting this work under the grant no. UMO-2015/17/B/ST4/03568.

Acknowledgments: The Authors acknowledge the Wrocław Center for Networking and Supercomputing (WCSS) for generous grants of the computing time and technical support.

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

References

1. Braga, D.; Grepioni, F.; Desiraju, G.R. Crystal Engineering and Organometallic Architecture. *Chem. Rev.* **1998**, *98*, 1375–1405. [[CrossRef](#)] [[PubMed](#)]
2. Biedermann, F.; Schneider, H.-J. Experimental Binding Energies in Supramolecular Complexes. *Chem. Rev.* **2016**, *116*, 5216–5300. [[CrossRef](#)] [[PubMed](#)]
3. Du, X.; Li, Y.; Xia, Y.-L.; Ai, S.-M.; Liang, J.; Sang, P.; Ji, X.-L.; Liu, S.-Q. Insights into Protein–Ligand Interactions: Mechanisms, Models, and Methods. *Int. J. Mol. Sci.* **2016**, *17*, 144. [[CrossRef](#)] [[PubMed](#)]
4. Jeffrey, G.A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer: New York, NY, USA, 1991.
5. Wytenbach, T.; Paizs, B.; Barran, P.; Brechi, L.; Liu, D.; Suhai, S.; Wysocki, V.H.; Bowers, M.T. The Effect of the Initial Water of Hydration on the Energetics, Structures, and H/D Exchange Mechanism of a Family of Pentapeptides: An Experimental and Theoretical Study. *J. Am. Chem. Soc.* **2003**, *125*, 13768–13775. [[CrossRef](#)] [[PubMed](#)]
6. Margenau, H.; Kestner, N. *Theory of Inter-Molecular Forces, International Series of Monographs in Natural Philosophy*; Pergamon Press: Oxford, UK, 1969.
7. Chałasiński, G.; Gutowski, M. Weak interactions between small systems. Models for studying the nature of intermolecular forces and challenging problems for ab initio calculations. *Chem. Rev.* **1988**, *88*, 943–962. [[CrossRef](#)]
8. Hobza, P.; Zahradnik, R. Intermolecular interactions between medium-sized systems. Nonempirical and empirical calculations of interaction energies. Successes and failures. *Chem. Rev.* **1988**, *88*, 871–897. [[CrossRef](#)]
9. Kolář, M.H.; Hobza, P. Computer Modeling of Halogen Bonds and Other σ -Hole Interactions. *Chem. Rev.* **2016**, *116*, 5155–5187. [[CrossRef](#)]
10. Grabowski, S.J. New type of halogen bond: Multivalent halogen interacting with π and σ -electrons. *Molecules* **2017**, *22*, 2150. [[CrossRef](#)]
11. Jabłoński, M. Theoretical insight into the nature of the intermolecular charge-inverted hydrogen bond. *Comput. Theor. Chem.* **2012**, *998*, 39–45. [[CrossRef](#)]

12. Jabłoński, M. Charge-inverted hydrogen bond vs. other interactions possessing a hydridic hydrogen atom. *Chem. Phys.* **2014**, *433*, 76–84. [\[CrossRef\]](#)
13. Jensen, F. *Introduction to Computational Chemistry*, 2nd ed.; John Wiley & Sons Ltd.: Chichester, UK, 2007.
14. Barone, V.; Cossi, M. Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J. Phys. Chem. A* **1998**, *102*, 1995–2001. [\[CrossRef\]](#)
15. 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**, *2*, 799–805. [\[CrossRef\]](#)
16. Skyner, R.E.; McDonagh, J.L.; Groom, C.R.; van Mourik, T.; Mitchell, J.B.O. A review of methods for the calculation of solution free energies and the modelling of systems in solution. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6174–6191. [\[CrossRef\]](#) [\[PubMed\]](#)
17. López, J.C.; Sánchez, R.; Blanco, S.; Alonso, J.L. Microsolvation of 2-azetidinone: A model for the peptide group-water interactions. *Phys. Chem. Chem. Phys.* **2015**, *17*, 2054–2066. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Da Silva, C.O.; Mennucci, B.; Vreven, T. Combining microsolvation and polarizable continuum studies: New insights in the rotation mechanism of amides in water. *J. Phys. Chem. A* **2003**, *107*, 6630–6637. [\[CrossRef\]](#)
19. Li, S.; Hong, M. Protonation, Tautomerization, and Rotameric Structure of Histidine: A Comprehensive Study by Magic-Angle-Spinning Solid-State NMR. *J. Am. Chem. Soc.* **2011**, *133*, 1534–1544. [\[CrossRef\]](#)
20. Day, R.M.; Thalhauser, C.J.; Sudmeier, J.L.; Vincent, M.P.; Torchilin, E.V.; Sanford, D.G.; Bachovchin, C.W.; Bachovchin, W.W. Tautomerism, acid-base equilibria, and H bonding of the six histidines in subtilisin BPN' by NMR. *Protein Sci.* **2003**, *12*, 794–810. [\[CrossRef\]](#)
21. Krishna Deepak, R.N.V.; Sankaramakrishnan, R. N-H...N Hydrogen Bonds Involving Histidine Imidazole Nitrogen Atoms: A New Structural Role for Histidine Residues in Proteins. *Biochemistry* **2016**, *55*, 3774–3783. [\[CrossRef\]](#)
22. Olivieri, F.A.; Burastero, O.; Drusin, S.I.; Defelipe, L.A.; Wetzler, D.E.; Turjanski, A.; Marti, M. Conformational and Reaction Dynamic Coupling in Histidine Kinases: Insights from Hybrid QM/MM Simulations. *J. Chem. Inf. Model.* **2020**, *60*, 833–842. [\[CrossRef\]](#)
23. Kawamura, K.; Yamada, T.; Kurihara, K.; Tamada, T.; Kuroki, R.; Tanaka, I.; Takahashi, H.; Niimura, N. X-ray and neutron protein crystallographic analysis of the trypsin-BPTI complex. *Acta Cryst. D* **2011**, *67*, 140–148. [\[CrossRef\]](#)
24. Saxena, A.K.; Singh, T.P.; Peters, K.; Fittkau, S.; Visanji, M.; Wilson, K.S.; Betzel, C. Structure of a ternary complex of proteinase K, mercury, and a substrate-analogue hexa-peptide at 2.2 Å resolution. *Proteins* **1996**, *25*, 195–201. [\[CrossRef\]](#)
25. Panek, J.J.; Mazzarello, R.; Novič, M.; Jezierska-Mazzarello, A. Impact of mercury(II) on proteinase K catalytic center: Investigations via classical and Born-Oppenheimer molecular dynamics. *Mol. Divers.* **2011**, *15*, 215–226. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Kuczer, M.; Błaszak, M.; Czarniewska, E.; Rosiński, G.; Kowalik-Jankowska, T. Mono- and Polynuclear Copper(II) Complexes of Alloferons 1 with Point Mutations (H6A) and (H12A): Stability Structure and Cytotoxicity. *Inorg. Chem.* **2013**, *52*, 5951–5961. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Olson, J.S.; Mathews, A.J.; Rohlf, R.J.; Springer, B.A.; Egeberg, K.D.; Sligar, S.G.; Tame, J.; Renaud, J.P.; Nagai, K. The role of the distal histidine in myoglobin and haemoglobin. *Nature* **1988**, *336*, 265–266. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Vila, J.A.; Arnautova, Y.A.; Vorobjev, Y.; Scheraga, H.A. Assessing the fractions of tautomeric forms of the imidazole ring of histidine in proteins as a function of pH. *Proc. Natl Acad. Sci. USA* **2011**, *108*, 5602–5607. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Kim, M.O.; Nichols, S.E.; Wang, Y.; McCammon, A. Effects of histidine protonation and rotameric states on virtual screening of M. tuberculosis RmlC. *J. Comput. Aided Mol. Des.* **2013**, *27*, 235–246. [\[CrossRef\]](#)
30. Cysewski, P.; Szeffler, B. Environment influences on the aromatic character of nucleobases and amino acids. *J. Mol. Model.* **2010**, *16*, 1709–1720. [\[CrossRef\]](#)
31. Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871. [\[CrossRef\]](#)
32. Kohn, W.; Sham, L.J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133–A1138. [\[CrossRef\]](#)
33. Bader, R.F.W. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, UK, 1990.
34. Jezierski, B.; Moszyński, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *7*, 1887–1930. [\[CrossRef\]](#)

35. Jezierska, A.; Panek, J.J. Cooperativity of hydrogen bonding network in microsolvated biotin, the ligand of avidin class proteins. *J. Mol. Model.* **2019**, *25*, 361. [CrossRef] [PubMed]
36. Rai, A.K.; Fei, W.; Lu, Z.; Lin, Z. Effects of microsolvation and aqueous solvation on the tautomers of histidine: A computational study on energy, structure and IR spectrum. *Theor. Chem. Acc.* **2009**, *124*, 37–47. [CrossRef]
37. Lee, S.-S.; Kim, J.-Y.; Han, Y.; Shim, H.-J.; Lee, S. Thermodynamic and kinetic stability of zwitterionic histidine: Effects of gas phase hydration. *Chem. Phys. Lett.* **2015**, *637*, 42–50. [CrossRef]
38. Cambridge Crystallographic Data Centre. Available online: <https://www.ccdc.cam.ac.uk/> (accessed on 1 July 2020).
39. Skelton, B.W.; Piggott, M.J.; Walkey, M.C. *CCDC 1840707: Experimental Crystal Structure Determination*; The Cambridge Crystallographic Data Centre (CCDC): Cambridge, UK, 2018.
40. Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620. [CrossRef]
41. Ditchfield, R.; Hehre, W.J.; Pople, J.A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724–728. [CrossRef]
42. Schaftenaar, G.; Noordik, J.H. Molden: A pre- and post-processing program for molecular and electronic structures. *J. Comput.-Aided Mol. Design* **2000**, *14*, 123–134. Available online: <http://cheminf.cmbi.ru.nl/molden/> (accessed on 4 June 2020). [CrossRef]
43. Becke, A.D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652. [CrossRef]
44. Lee, C.; Yang, W.; Parr, R.G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789. [CrossRef]
45. Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868. [CrossRef]
46. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parameterization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **2010**, *132*, 154104. [CrossRef]
47. 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.
48. Hohenstein, E.G.; Sherrill, C.D. Density fitting of intramonomer correlation effects in symmetry-adapted perturbation theory. *J. Chem. Phys.* **2010**, *133*, 014101. [CrossRef] [PubMed]
49. Møller, C.; Plesset, M.S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622. [CrossRef]
50. Kendall, R.A.; Dunning, T.H., Jr.; Harrison, R.J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806. [CrossRef]
51. PSI4 Open-Source Quantum Chemistry. Available online: <http://www.psicode.org/psi4manual/master/dfmp2.html#> (accessed on 4 June 2020).
52. Boys, S.F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553–566. [CrossRef]
53. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. *Gaussian 16*; rev. B.01; Gaussian, Inc.: Wallingford, CT, USA, 2016.
54. AIMAll (Version 19.10.12) Software by T. A. Keith. Available online: aim.tkgristmill.com (accessed on 4 June 2020).
55. Parrish, R.M.; Burns, L.A.; Smith, D.G.A.; Simmonett, A.C.; DePrince, A.E.; Hohenstein, E.G.; Bozkaya, U.; Sokolov, A.Y.; Di Remigio, R.; Richard, R.M.; et al. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J. Chem. Theory Comput.* **2017**, *13*, 3185–3197. [CrossRef]
56. Humphrey, W.; Dalke, A.; Schulten, K. VMD—Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38. [CrossRef]
57. The GIMP Development Team, GIMP. 2019. Available online: <https://www.gimp.org> (accessed on 1 July 2020).
58. Koch, U.; Popelier, P.L.A. Characterization of C–H–O Hydrogen Bonds on the Basis of the Charge Density. *J. Phys. Chem.* **1995**, *99*, 9747–9754. [CrossRef]

59. Grimme, S.; Mück-Lichtenfeld, C.; Erker, G.; Kehr, G.; Wang, H.; Beckers, H.; Willner, H. When Do Interacting Atoms Form a Chemical Bond? Spectroscopic Measurements and Theoretical Analyses of Dideuteriophenanthrene. *Angew. Chem. Int. Ed.* **2009**, *48*, 2592–2595. [[CrossRef](#)]
60. Espinosa, E.; Molins, E.; Lecomte, C. Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chem. Phys. Lett.* **1998**, *285*, 170–173. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).