Article

Dissecting Bonding Interactions in Cysteine Dimers

Santiago Gómez [1] , **Sara Gómez** [2] , **Jorge David** [3] , **Doris Guerra** [1] , **Chiara Cappelli** [2,*]
and Albeiro Restrepo [1,*]

[1] Instituto de Química, Universidad de Antioquia UdeA, Calle 70 No. 52-21, Medellín 050010, Colombia
[2] Scuola Normale Superiore, Classe di Scienze, Piazza dei Cavalieri 7, 56126 Pisa, Italy
[3] Escuela de Ciencias y Humanidades, Departamento de Ciencias Básicas, Universidad Eafit, AA 3300, Medellín 050022, Colombia
* Correspondence: chiara.cappelli@sns.it (C.C.); albeiro.restrepo@udea.edu.co (A.R.)

Abstract: Neutral (*n*) and zwitterionic (*z*) forms of cysteine monomers are combined in this work to extensively explore the potential energy surfaces for the formation of cysteine dimers in aqueous environments represented by a continuum. A simulated annealing search followed by optimization and characterization of the candidate structures afforded a total of 746 structurally different dimers held together via 80 different types of intermolecular contacts in 2894 individual non-covalent interactions as concluded from Natural Bond Orbitals (NBO), Quantum Theory of Atoms in Molecules (QTAIM) and Non-Covalent Interactions (NCI) analyses. This large pool of interaction possibilities includes the traditional primary hydrogen bonds and salt bridges which actually dictate the structures of the dimers, as well as the less common secondary hydrogen bonds, exotic $X\cdots Y$ (X = C, N, O, S) contacts, and $H\cdots H$ dihydrogen bonds. These interactions are not homogeneous but have rather complex distributions of strengths, interfragment distances and overall stabilities. Judging by their Gibbs bonding energies, most of the structures located here are suitable for experimental detection at room conditions.

Keywords: cysteine dimers; NBO; NCI; QTAIM; stochastic optimization; hydrogen bonding; salt bridges; non-covalent interactions

1. Introduction

Cysteine, $HOOCCH(NH_2)CH_2SH$ is the only amino acid among the unique list of 20 found in proteins that possesses a thiol functional group [1]. This thiol group, which is comparatively a weaker Brønsted–Lowry acid than O–H in carboxylic acids, is extremely important in biochemistry. Among a large number of known functionalities, the S–H group is responsible for nucleophilic additions to $\alpha, \beta-$unsaturated carbonyl compounds via Michael reactions [2–5], serves as a deprotonation agent [6], and its strong nucleophilicity renders cysteine a key component of the active sites of several protease enzymes [7,8]. In addition to S–H, depending on the conditions, sulfur atoms in cysteine engage in S–S disulfide bonds, which are a central element determining secondary and tertiary structure in proteins [9–11] and are relevant in physiological redox activity. According to the SwissProt databank [12], six percent of all proteins contain at least one disulfide bridge, and the median number of disulfide bonds is two.

Protein$\cdots$protein interaction is one of the central problems in molecular biology. Unfortunately, with present days computational methods, a thorough understanding from a molecular perspective is unattainable because the number of explicit contacts grows exponentially with the size of the protein. For example, insulin is one of the smallest biologically active proteins containing a primary sequence of just 51 amino acids (six cysteines among them), for the dimer of this protein, not counting salt bridges and other intermolecular interactions, there are at least 6×10^4 possibilities for hydrogen bonding in the classical $X^{\delta-}\cdots^{+\delta}H-Y^{\delta-}$ description [13]. Typical proteins and other biomolecules contain in excess

of 1000 amino acids and it is not uncommon to find very large proteins such as titin, which depending on the splice isoform, contains between 27,000 to 35,000 amino acids [14]; thus, the number of specific amino acid$\cdots$amino acid contacts quickly becomes intractable.

In an effort to understand the intricacies of protein$\cdots$protein interactions, the astonishingly large number of specific contacts calls for the use of reduced molecular models, often as gas phase isolated dimers of individual amino acids [13]. In this context, we attempt a detailed study of the cysteine dimers. This is a very complicated issue in its own right: First, there are the two enantiomeric forms of the amino acid. Second, there is the possibility of neutral and zwitterionic forms. Third, there is a large number of conformers of the monomer in a small energy window amenable to form dimers, which, for just the neutral form, has been estimated via ab initio computations to be 42, 51 or 71 depending on the sophistication of the model chemistry [1,9,15]. Notice that six well defined conformers have been experimentally identified via IR and MW spectroscopies [16,17]. Fourth, cysteine contains seven hydrogen atoms and four electronegative atoms; thus, ignoring salt bridges, dispersive dihydrogen interactions and other exotic contacts [13], from the classical $X^{\delta-}\cdots^{+\delta}H-Y^{\delta-}$ perspective, a total of 28 individual primary plus secondary hydrogen bonds are possible for each dimer. The number of possibilities is reduced to 20, distributed as 12 primary and 8 secondary HBs if the two H–C_β bonds are grouped into just one type and if the two N–H bonds are considered as another type. Fifth, as seen for example in alanine [13], several dimers are attached by more than one contact. Sixth, S–H leads to considerably weaker interactions than O–H, then, the potential energy surface (PES) for the dimers is expected to be considerably richer in weakly bound pairs and thus high levels of electron correlation are needed to correctly describe the intermolecular interactions. Seventh, the environment sensibly impacts the ability of biomolecules to interact in biological settings; thus, using gas phase dimers as a reduced model for protein$\cdots$protein interactions does not seem enough and at least solvent effects must be included.

Cysteine has been thoroughly studied through experiments and computations. Besides the above mentioned publications dealing with the conformations of the monomer [1,9,15–17], Kaminski et al. [18] and Sadlej et al. [19] undertook somewhat exhaustive explorations of the conformational PES for the monomers to rationalize Raman Optical Activity (ROA) and Vibrational Circular Dichroism (VCD) spectra. For the dimers, early studies focused on gas phase and implicitly solvated models with limited explorations of the PES using a few hand constructed configurations [20], later studies considered both explicit water molecules and the neutral and zwitterionic forms [21,22]. There are reports dealing with the formation of the dimers, their stability and bonding (via density differences) when adsorbed in gold surfaces [23,24]. Group IA cations bonded to cysteine dimers have also been studied [25].
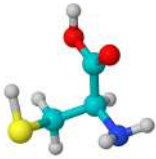
In view of the expected complexity arising from the multiple classical donor and acceptor hydrogen bonding sites of cysteine, which as a reduced molecular model has profound implications in the protein interaction problem, the brief summary of the scientific literature dealing with cysteine dimers just exposed reveals an unsatisfactory level of understanding not only of the potential energy surface but of the nature of the intermolecular bonding interactions for cysteine$\cdots$cysteine. The present work is an attempt to remedy this situation. To that end, we undertake systematic explorations of the neutral (*n*) and zwitterionic (*z*) pairs in $n\cdots n$, $n\cdots z$ and $z\cdots z$ combinations of low lying energy monomers via stochastic samplings of the corresponding PES, and dissected the nature of the interactions using formal quantum descriptors of bonding as provided by the Quantum Theory of Atoms in Molecules [26–29] (QTAIM), the Natural Bond Orbitals [30–33] (NBO) and the Non Covalent Interactions [34,35] (NCI), as discussed in the Section 2.

2. Computer Methods

Sampling the potential energy surfaces for all possible neutral (72 computed, 6 experimentally found) and zwitterionic forms (12 computed) is not only impossible but unnecessary under the premise that a few representative pairs would capture the vast majority of the specific contacts and thus would provide a sound picture applicable to all

cases. Accordingly, we took two of the experimentally detected neutral monomers [17] (n_1, n_2 in Table 1) and two of the computed lowest energy zwitterions [9] (z_1, z_2 in Table 1) and exhausted all $x(x-1)/2 + x = 10$ possible dimeric combinations for $x = 4$. Each pair was superimposed at the center of a cubic box of 512 Å^3 (8 Å side) and was allowed to evolve under simulated annealing conditions [36–38] as implemented in the ASCEC program [39]. Superimposing the interacting system at the center of the box gives the algorithm the worst possible starting point (we call this the big bang initial conditions) and guarantees that the located stationary points within the corresponding PES are free of any structural bias. ASCEC [40,41], after its Spanish acronym Annealing Simulado Con Energía Cuántica, randomly explores the quantum energy landscape for the dimeric interaction, subjects the generated structures to a modified Metropolis acceptance test, and delivers a set of candidate structures that undergo further optimization via gradient following techniques and characterization as true minima via harmonic vibrational analysis. Each one of the 10 possible dimeric combinations was treated to duplicate ASCEC runs. All ASCEC runs and geometry optimizations were carried out in an aqueous environment represented by a continuum according to the PCM (ASCEC) and IEFPCM (optimization) models. [42–44].

Table 1. Structures of the B3LYP–D3/6–311++G(d, p) monomers of neutral (n_1, n_2) and zwitterionic (z_1, z_2) L–cysteine. In each case, $\Delta\Delta G$ are the corresponding differences in Gibbs energies at room conditions with respect to n_1, z_1, the lowest energy monomers. Descriptors of intramolecular bonding derived from QTAIM and NBO are included along with the specific NBO orbitals responsible for the interactions. $\Delta\Delta G, E_{d \to a}^{(2)}$ in kcal mol^{-1}, all other descriptors in a.u. n_1 and n_2 have been experimentally detected [17].

Monomer →	n_1	n_2	z_1		z_2
Properties ↓					
$\Delta\Delta G$	0.0	2.2	0.0	0.0	0.3
A$\cdots$B Distance (Å)	1.86		2.63	1.91	1.94
$10^2 \times \rho(\mathbf{r}_c)$	4.07		1.37	3.29	3.11
$10^2 \times \nabla^2\rho(\mathbf{r}_c)$	10.87		4.63	11.78	11.29
$|\mathcal{V}(\mathbf{r}_c)|/\mathcal{G}(\mathbf{r}_c)$	1.12		0.82	0.93	0.92
$10^2 \times \mathcal{H}(\mathbf{r}_c)/\rho(\mathbf{r}_c)$	-9.18		12.73	5.62	7.04
$\phi_d \to \phi_a$	$n_N \to \sigma^*_{O-H}$		$n_S \to \sigma^*_{N-H}$	$n_O \to \sigma^*_{N-H}$	$n_O \to \sigma^*_{N-H}$
$E_{d \to a}^{(2)}$	14.9		2.3	8.3	7.3

Final equilibrium geometries and Gibbs energies for every located dimer computed with the Gaussian09 suite of programs [45] are reported here using the dispersion corrected B3LYP–D3/6–311++G(d, p) model chemistry. Binding using the Gibbs energies at room conditions (1 atm, 298.16 K) are calculated as the negative difference between the energy of the cluster and the energy of the fragments, $BE = -(E_{cluster} - \sum E_{fragments})$, in this way, positive binding energies indicate strongly bonded clusters.

Once the molecular wavefunctions and electron densities for the optimized geometries are recovered by the procedure just stated, we use them to gain insight into the nature of intermolecular bonding interactions using the tools provided by QTAIM, NBO, and NCI following strategies described elsewhere [46–53]. At this point, we state that we use those methods as well established analysis tools, the interested reader is directed to the specialized literature for detailed discussions of their merits and shortcomings and for a description of how the calculated descriptors are related to bonding [6,33,49]. In short, use the Multiwfn suite [54] to find the bond critical points (BCPs, $\mathbf{r}_c$) corresponding to

intermolecular interactions, analyze their properties, i.e., the electron density $\rho(\mathbf{r}_c)$, its Laplacian $\nabla^2\rho(\mathbf{r}_c)$, the total, kinetic, and potential energy densities $\mathcal{H}(\mathbf{r}_c) = \mathcal{G}(\mathbf{r}_c) + \mathcal{V}(\mathbf{r}_c)$, and the virial ratio $|\mathcal{V}(\mathbf{r}_c)|/\mathcal{G}(\mathbf{r}_c)$. With the same program, we calculate the Laplacian of the electron density, a scalar field that gives direct information about the most probable regions to find the electrons. Then, we use NBO6 [55] to pinpoint the specific orbitals involved in the intermolecular interactions associated with each BCP and estimate the strength of the interaction via second order perturbation energy for the interaction between the donor and acceptor orbitals, $E_{d\rightarrow a}^{(2)}$. The NCIPLOT program [56] was used to derive the non-covalent interaction surfaces. Jmol and VMD [57,58] were used to visualize the molecular structures, and their related surfaces and orbital interactions.

3. Results

Table 1 shows the structures of the n_1, n_2 neutral and z_1, z_2 zwitterionic forms chosen in this work to study the dimers of cysteine. The four non-covalent intramolecular interactions, as derived from QTAIM are displayed as dotted lines along with the involved NBOs. QTAIM and NBO descriptors are included as well. Only n_2 has a structure free from intramolecular hydrogen bonds while z_1 exhibits two intramolecular contacts. Except for n_1, all intramolecular interactions are characterized as weak, long range contacts because of the positive Laplacians, relatively small accumulation of electron densities at BCPs, virial ratios smaller than 1, and positive bond degree parameters. However, the $n_N \rightarrow \sigma_{O-H}^*$ interaction in n_1 is uncharacteristically strong, with values for the bonding descriptors that in every case surpass those of the archetypal hydrogen bond in the water dimer [30,59]. These intramolecular contacts are quite important because the formation of the dimers will usually involve investing energy to eliminate those interactions in favor of dimeric contacts. The electrostatic potentials in Figure S1 of the Supplementary Information show the blue and red regions which are more susceptible to the formation of intermolecular contacts according to the classic electrostatic view of hydrogen bonding.

3.1. Structures and Energies

Complex and rich potential energy surfaces are uncovered by our stochastic searches in every case. A total of 746 distinct well defined dimers were located in the 10 monomer + monomer possible combinations. Table 2 lists the number of structural isomers for each PES and also shows that the vicinities of the putative global minima are populated with other close energy dimers; thus, all structures accounting for populations larger than 1% are within 2.1 kcal/mol of the $n \cdots n$ lowest energy structure, and so on. This point is emphasized by the results shown in Tables S1–S3 in the Supplementary information and in Figure 1, which clearly show that there are no dominant isomers.

Table 2. Summary of structural and energetical properties of the cysteine dimers. ΔG range: Gibbs energy difference between the highest and lowest energy structure in kcal/mol. %x_i: isomer populations listed in Tables S1–S3 in the Supplementary Material.

Composition	Structures	ΔG Range	
		for %$x_i > 1$	**for All Structures**
$n \cdots n$	244	2.1	21.4
$n \cdots z$	316	1.1	25.5
$z \cdots z$	186	2.3	27.9

Although 746 is a very large number of structures and is considerably higher than the numbers reported in any of the previous studies, we recognize that given the complexity of our problem, no stochastic or analytic search algorithm is able to locate all possible geometries. A representative set including only those dimers with populations exceeding 5% within each PES is shown in Figure 1, along with the NBOs responsible for the inter-

molecular interactions. Cartesian coordinates for all 746 structures located in this work are provided in the Supplementary Material.
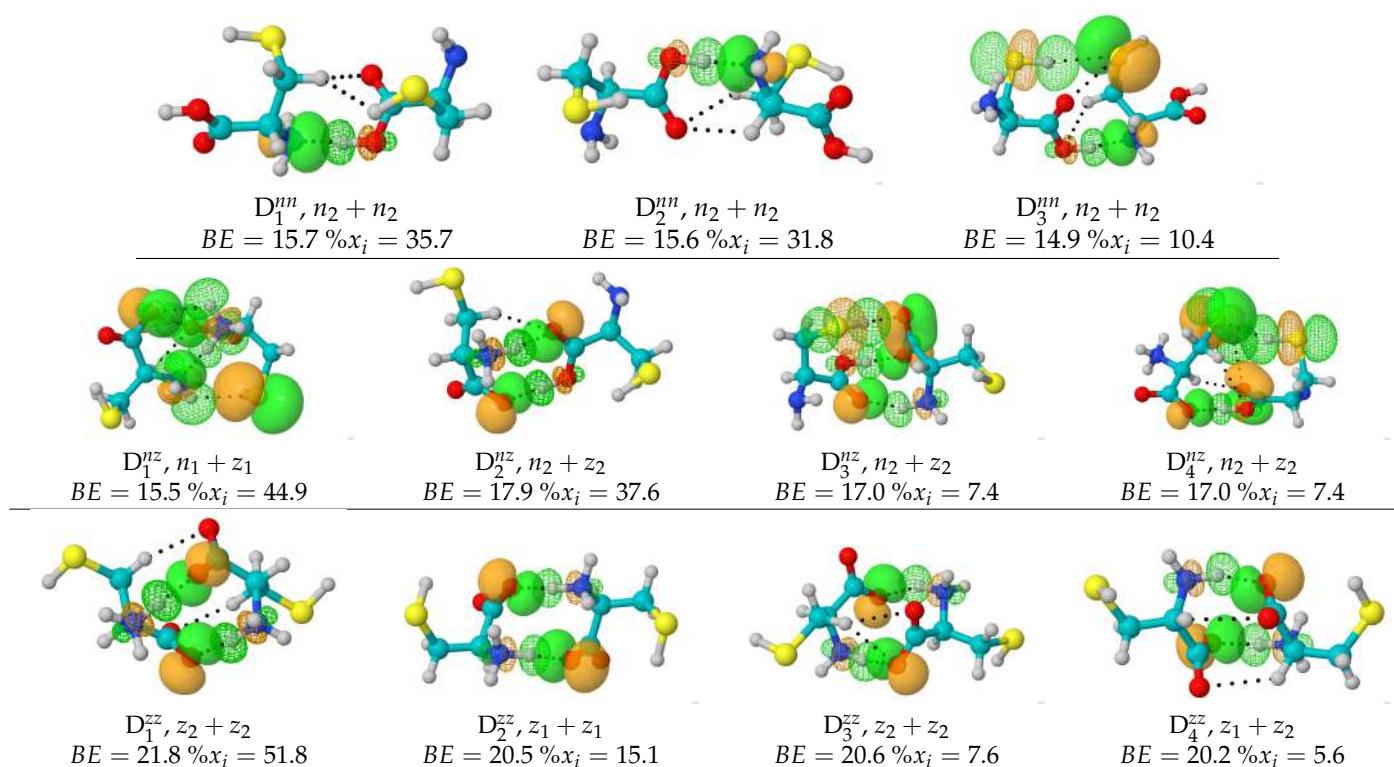

D_1^{nn}, $n_2 + n_2$
$BE = 15.7$ %$x_i = 35.7$

D_2^{nn}, $n_2 + n_2$
$BE = 15.6$ %$x_i = 31.8$

D_3^{nn}, $n_2 + n_2$
$BE = 14.9$ %$x_i = 10.4$

D_1^{nz}, $n_1 + z_1$
$BE = 15.5$ %$x_i = 44.9$

D_2^{nz}, $n_2 + z_2$
$BE = 17.9$ %$x_i = 37.6$

D_3^{nz}, $n_2 + z_2$
$BE = 17.0$ %$x_i = 7.4$

D_4^{nz}, $n_2 + z_2$
$BE = 17.0$ %$x_i = 7.4$

D_1^{zz}, $z_2 + z_2$
$BE = 21.8$ %$x_i = 51.8$

D_2^{zz}, $z_1 + z_1$
$BE = 20.5$ %$x_i = 15.1$

D_3^{zz}, $z_2 + z_2$
$BE = 20.6$ %$x_i = 7.6$

D_4^{zz}, $z_1 + z_2$
$BE = 20.2$ %$x_i = 5.6$

Figure 1. Lowest energy structures and the NBOs responsible for the strongest intermolecular interactions in the neutral $n \cdots n$ (top), neutral + zwitterionic $n \cdots z$ (middle) and zwitterionic $z \cdots z$ (bottom) B3LYP–D3/6–311++g(d, p) potential energy surfaces of the cysteine dimers under the continuum IEFPCM solvent model for water. Solid/meshed surfaces correspond to charge donor/acceptor orbitals, respectively. *BE*: binding energies in kcal/mol calculated using the Gibbs free energies at room conditions. See Table 1 for the structures of n_1, n_2, z_1, z_2, the isolated monomers. Only those structures with populations (%x_i) higher than 5% within each PES are included. Energetics for the entire set of 746 dimers is provided in Tables S1–S3 of the supplementary material.

On the basis of purely ZPE–corrected electronic energies (Tables S1–S3), all cysteine dimers are stable towards fragmentation into the corresponding monomers, however, Figure 2 shows that consideration of temperature and entropy leads to 235 clusters (80 $n \cdots n$, 154 $n \cdots z$ and 1 $z \cdots z$) having negative binding energies as calculated from the Gibbs energies; thus, those particular structures correspond to unstable dimers and are not amenable to experimental detection at room conditions in aqueous environments, a fact that is emphasized by their %$x_i \approx 0$ populations. Notice the contrast with the 416 $n \cdots n$ gas phase equilibrium structures reported for the Alanine dimers [13], which are all strongly bonded. Binding energies show a clear $BE_{n \cdots n} < BE_{n \cdots z} < BE_{z \cdots z}$ ordering; thus, there is a marked preference for charged cysteine dimers in aqueous environments.

Figure 2, showing distribution plots of the Gibbs binding energies leads to a few relevant observations: Dashed vertical lines indicate the expected values of the binding energies using the Boltzmann populations of the Gibbs energies within each PES as weighing factors. 14.3, 16.6 and 20.9 kcal/mol are obtained for $n \cdots n$, $n \cdots z$, $z \cdots z$, again showing a preference for charged dimers in aqueous environments. To put these binding energies in context, they are larger than the gas phase Gibbs binding energies of acetamide and acetic acid, which are 2.1 and 3.8 kcal/mol respectively, according to Copeland et al. [60] Notice that the same authors reported substantially higher binding energies when using only the ZPE-corrected electronic energies: 12.3 kcal/mol for acetamide and 14.7 kcal/mol for acetic acid. High ZPE-corrected binding energies have also been reported for the lowest

energy structures in similar systems: 16.6 kcal/mol for the dimers of formic acid according to Kalescky et al. [61] and 12.7 according to Farfán et al. [62], 19.0 kcal/mol for the dimers of carbonic acid [63], and 20.9 kcal/mol for the alanine dimers [13]. Tables S1–S3 in the Supplementary material show exactly the same trend for all cysteine dimers calculated here, that is, comparatively much higher binding energies are obtained when only ZPE-corrected are considered with expected values of 25.9, 28.9, 33.7 kcal/mol for the $n \cdots n$, $n \cdots z$, and $z \cdots z$ cases, respectively. Notice that these numbers are up to over 6 times larger than the 5.0 kcal/mol binding energy arising because of a single hydrogen bond in the archetypal water dimer [59]. Finally, notice that those structures being unstable towards fragmentation ($BE < 0$) have minimal populations and thus do not contribute to the expected value of the binding energy. The role of dispersive interactions is clearly seen in the fact that when the D3 correction is removed from B3LYP, all strongly bound isomers become unstable towards fragmentation (values within parentheses in Tables S1–S3). For the cysteine dimers with positive binding energies, Tables S1–S3 show that the structures with the largest populations are strongly bonded.
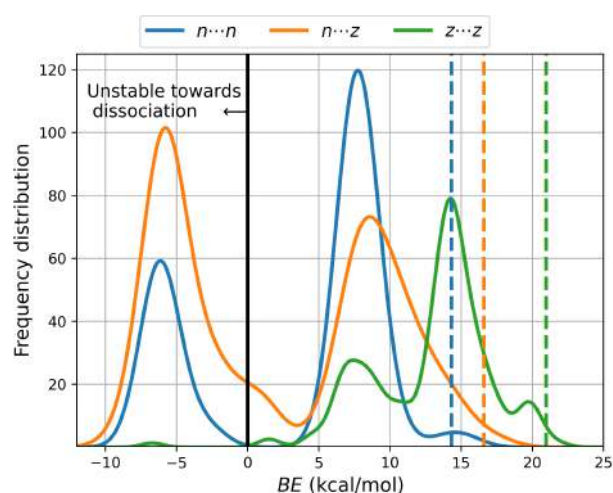

Figure 2. Distribution of binding energies of the cysteine dimers using the Gibbs energies. Dashed vertical lines mark the expected value for each potential energy surface using the Boltzmann populations at room conditions as weighing factors.

As general structural features of the cysteine dimers, we point out that in all cases where neutral monomers are involved, n_2 (no intramolecular HB, Table 1) leads to lower energy dimers. Additionally, except for D_1^{nz}, in all structures that contain n_1, the intramolecular HB in n_1 remains in the dimer. A surprising result is that contrary to the well known structures of the dimers of carboxylic acids, out of the 244 well characterized $n \cdots n$ local minima, only two (D_5^{nn}, $\%x_i = 1.7$ and D_7^{nn}, $\%x_i = 1.3$) exhibit the traditional eight atom, cyclic double C=O$\cdots$H–O stabilizing network. We attribute this to two factors: one, the influence of the solvent which favors other configurations, and two, the intramolecular hydrogen bond occupying the O–H bond in n_1 remains in all but one $n \cdots n$ dimer; thus, this bond is not available for intermolecular bonding (see the dissection of intermolecular bonding interactions below).

3.2. Bonding

The configurational space for the cysteine dimers is complex and rich. We located and characterized a total of 746 structures and there might as well be many more. This geometrical variety arises because of the large number of possible interactions discussed above. Our stochastic search and subsequent dissection of bonding interactions (see below) uncovers an astonishing total of 80 well characterized physically different types of direct intermolecular contacts listed in Tables 3 and 4. Gratifyingly, the found structures account for every single one of the 20 possible hydrogen bonds among the monomers as exposed in

the Introduction and also reveal additional salt bridges, dihydrogen bonds, and a number of exotic $X \cdots Y$ (X, Y = O, S, N, C) and $C \cdots H$–C, $C \cdots H$–S contacts. Notice that the $n \cdots n$ dimers exhibit a well balanced field of all non-charged interactions while the $z \cdots z$ dimers favor the salt bridges by a long shot (159 appearances) and, to a lesser extent, other interactions where only one of the fragments is charged. What should be clear is that the largest contributors to the stabilization of the dimers are $N \cdots H$–O interactions in the $n \cdots n$ dimers, $C=O^- \cdots H$–O in $n \cdots z$, and $C=O^- \cdots H$–N^+ salt bridges in $z \cdots z$. We think it is important to point out that, as a general rule, due to the comparatively larger interaction strength, it is the primary neutral, charged, or salt bridges forms of HBs that determine the molecular geometry of the dimers while secondary HBs and exotic contacts are a consequence of the structure (*vide infra*), however, the collective action of the multiple weak interactions on the stabilization energy of each cluster should not be ignored.

Our topological analysis of the electron densities of the 746 equilibrium structures located in this work affords a total of 2894 intermolecular contacts, which are collected into 80 different types in Tables 3 and 4. Without a single exception, positive Laplacians at bond critical points (see Figure S2 in the supplementary material) indicate that bonding in the $n \cdots n$, $n \cdots z$ and $z \cdots z$ cysteine dimers occurs via closed shell interactions, in the form of either ionic bonding or long range weak interactions. We dissect the nature of intermolecular interactions next.

Table 3. Properties of the 20 types of primary hydrogen bonds, 10 types of secondary hydrogen bonds, and 12 types of dihydrogen bonds stabilizing the cysteine dimers. $N_i^{A \cdots B}$ is the number of times that the interaction appears in the corresponding type of dimers. ϕ_d, ϕ_a are the charge donor and acceptor orbitals as identified from NBO. An example of a dimer containing each particular interaction is given in the rightmost column.

Label	Type	$n \cdots n$	$z \cdots z$	$n \cdots z$	$\phi_d \to \phi_a$	Example
			$N_i^{A \cdots B}$			
Primary hydrogen bonds						
1	$C=O \cdots H$–O	22	-	-	$n_O \to \sigma^*_{H-O}$	D_5^{nn}
2	$C=O^- \cdots H$–O	-	-	52	$n_O \to \sigma^*_{H-O}$	D_{38}^{nz}
3	$C=O \cdots H$–N	46	-	-	$n_O \to \sigma^*_{H-N}$	D_{119}^{nn}
4	$C=O \cdots H$–N^+	-	-	51	$n_O \to \sigma^*_{H-N}$	D_2^{nz}
5	$C=O^- \cdots H$–N	-	-	89	$n_O \to \sigma^*_{H-N}$	D_{55}^{nz}
6	$C=O^- \cdots H$–N^+	-	159	-	$n_O \to \sigma^*_{H-N}$	D_{16}^{zz}
7	H–O $\cdots H$–O	18	-	-	$n_O \to \sigma^*_{H-O}$	D_{172}^{nn}
8	H–O $\cdots H$–N	51	-	-	$n_O \to \sigma^*_{H-N}$	D_{137}^{nn}
9	H–O $\cdots H$–N^+	-	-	32	$n_O \to \sigma^*_{H-N}$	D_{207}^{nz}
10	$N \cdots H$–O	7	-	-	$n_N \to \sigma^*_{H-O}$	D_1^{nn}
11	$N \cdots H$–N	19	-	-	$n_N \to \sigma^*_{H-N}$	D_{174}^{nn}
12	$N \cdots H$–N^+	-	-	19	$n_N \to \sigma^*_{H-N}$	D_1^{nz}
13	$S \cdots H$–S	29	30	54	$n_S \to \sigma^*_{H-S}$	D_{61}^{nn}
14	$S \cdots H$–N	55	-	19	$n_S \to \sigma^*_{H-N}$	D_{51}^{nn}
15	$S \cdots H$–N^+	-	41	52	$n_S \to \sigma^*_{H-N}$	D_{175}^{zz}
16	$N \cdots H$–S	19	-	14	$n_N \to \sigma^*_{H-S}$	D_{108}^{nn}
17	$S \cdots H$–O	16	-	12	$n_S \to \sigma^*_{H-O}$	D_{107}^{nn}
18	H–O $\cdots H$–S	34	-	29	$n_O \to \sigma^*_{H-S}$	D_8^{nn}
19	$C=O \cdots H$–S	49	-	27	$n_O \to \sigma^*_{H-S}$	D_{18}^{nn}
20	$C=O^- \cdots H$–S	-	85	68	$n_O \to \sigma^*_{H-S}$	D_{142}^{zz}
Secondary hydrogen bonds						
21	$C=O \cdots H$–C_α	49	-	26	$n_O \to \sigma^*_{H-C}$	D_{46}^{nn}
22	$C=O \cdots H$–C_β	57	-	44	$n_O \to \sigma^*_{H-C}$	D_{37}^{nn}
23	$C=O^- \cdots H$–C_α	-	68	40	$n_O \to \sigma^*_{H-C}$	D_{145}^{zz}
24	$C=O^- \cdots H$–C_β	-	108	71	$n_O \to \sigma^*_{H-C}$	D_{144}^{zz}
25	H–O $\cdots H$–C_α	48	-	20	$n_O \to \sigma^*_{H-C}$	D_{144}^{nn}
26	H–O $\cdots H$–C_β	71	-	41	$n_O \to \sigma^*_{H-C}$	D_{156}^{nn}
27	$N \cdots H$–C_α	16	-	11	$n_N \to \sigma^*_{H-C}$	D_{188}^{nn}

Table 3. *Cont.*

Label	Type	$N_i^{A \cdots B}$			$\phi_d \rightarrow \phi_a$	Example
		$n \cdots n$	$z \cdots z$	$n \cdots z$		
28	$N \cdots H–C_\beta$	34	-	19	$n_N \rightarrow \sigma^*_{H–C}$	D^{nn}_{186}
29	$S \cdots H–C_\alpha$	52	24	57	$n_S \rightarrow \sigma^*_{H–C}$	D^{nn}_{39}
30	$S \cdots H–C_\beta$	59	37	81	$n_S \rightarrow \sigma^*_{H–C}$	D^{nn}_{44}
Dihydrogen contacts						
31	$C_\alpha–H \cdots H–N$	7	-	3	$\sigma_{C–H} \rightarrow \sigma^*_{H–N}$	D^{nn}_{135}
32	$C_\beta–H \cdots H–N$	13	-	3	$\sigma_{C–H} \rightarrow \sigma^*_{H–N}$	D^{nn}_{86}
33	$C_\beta–H \cdots H–N^+$	-	-	1	$\sigma_{C–H} \rightarrow \sigma^*_{H–N}$	D^{nz}_{139}
34	$N–H \cdots H–N$	3	-	-	$\sigma_{N–H} \rightarrow \sigma^*_{H–N}$	D^{nn}_{99}
35	$C_\alpha–H \cdots H–C_\alpha$	4	3	-	$\sigma_{C–H} \rightarrow \sigma^*_{H–C}$	D^{nn}_{152}
36	$C_\beta–H \cdots H–C_\beta$	42	15	23	$\sigma_{C–H} \rightarrow \sigma^*_{H–C}$	D^{nn}_{78}
37	$C_\alpha–H \cdots H–C_\beta$	27	20	22	$\sigma_{C–H} \rightarrow \sigma^*_{H–C}$	D^{nn}_{64}
38	$S–H \cdots H–S$	5	3	3	$\sigma_{S–H} \rightarrow \sigma^*_{H–S}$	D^{nn}_{62}
39	$S–H \cdots H–N$	5	-	3	$\sigma_{S–H} \rightarrow \sigma^*_{H–N}$	D^{nn}_{61}
40	$S–H \cdots H–N^+$	-	1	-	$\sigma_{S–H} \rightarrow \sigma^*_{H–N}$	D^{zz}_{53}
41	$S–H \cdots H–C_\alpha$	15	8	12	$\sigma_{S–H} \rightarrow \sigma^*_{H–C}$	D^{nn}_{159}
42	$S–H \cdots H–C_\beta$	23	14	21	$\sigma_{S–H} \rightarrow \sigma^*_{H–C}$	D^{nn}_{157}

Table 4. Properties of the "exotic" intermolecular contacts found in the cysteine dimers. $N_i^{A \cdots B}$ is the number of times that the interaction appears in the corresponding type of dimers. ϕ_d, ϕ_a are the charge donor and acceptor orbitals as identified from NBO. An example of a dimer containing each particular interaction is given in the rightmost column.

Label	Type	$N_i^{A \cdots B}$			$\phi_d \rightarrow \phi_a$	Example
		$n \cdots n$	$z \cdots z$	$n \cdots z$		
O $\cdots$ C contacts						
43	$C=O \cdots C=O$	6	-	-	$n_O \rightarrow \pi^*_{C=O}$	D^{nn}_{113}
44	$C=O \cdots C=O^-$	-	-	4	$n_O \rightarrow \pi^*_{C=O}$	D^{nz}_{6}
45	$C=O^- \cdots C=O$	-	-	5	$n_O \rightarrow \pi^*_{C=O}$	D^{nz}_{90}
46	$C=O^- \cdots C=O^-$	-	7	-	$n_O \rightarrow \pi^*_{C=O}$	D^{zz}_{14}
47	$C=O \cdots C_\alpha–C$	5	-	-	$\pi_{C=O} \rightarrow \sigma^*_{C–C}$	D^{nn}_{112}
48	$C=O^- \cdots C_\alpha–C$	-	-	1	$n_O \rightarrow \sigma^*_{C–C}$	D^{nz}_{125}
49	$C=O \cdots C_\beta–S$	2	-	2	$n_O \rightarrow \sigma^*_{C–S}$	D^{nn}_{25}
50	$C=O^- \cdots C_\beta–S$	-	2	2	$n_O \rightarrow \sigma^*_{C–S}$	D^{zz}_{132}
51	$H–O \cdots C=O$	4	-	-	$n_O \rightarrow \pi^*_{C=O}$	D^{nn}_{161}
52	$H–O \cdots C_\beta–S$	2	-	1	$n_O \rightarrow \sigma^*_{C–S}$	D^{nn}_{80}
O $\cdots$ O contacts						
53	$C=O \cdots O=C$	14	-	-	$\pi_{C=O} \rightarrow \pi^*_{O=C}$	D^{nn}_{102}
54	$C=O^- \cdots {}^-O=C$	-	17	-	$n_O \rightarrow \pi^*_{O=C}$	D^{zz}_{10}
55	$C=O^- \cdots O=C$	-	-	14	$n_O \rightarrow \pi^*_{O=C}$	D^{nz}_{23}
56	$C=O \cdots O–H$	26	-	-	$n_O \rightarrow \sigma^*_{O–H}$	D^{nn}_{230}
57	$C=O^- \cdots O–H$	-	-	13	$n_O \rightarrow \sigma^*_{O–H}$	D^{nz}_{217}
58	$H–O \cdots O–H$	11	-	-	$n_O \rightarrow \sigma^*_{O–H}$	D^{nn}_{45}
O $\cdots$ N contacts						
59	$N \cdots O=C$	2	-	-	$n_N \rightarrow \pi^*_{O=C}$	D^{nn}_{11}
60	$N \cdots O–H$	2	-	-	$n_N \rightarrow \sigma^*_{O–H}$	D^{nn}_{35}
61	$C=O^- \cdots {}^+N–H$	-	16	-	$n_O \rightarrow \sigma^*_{N–H}$	D^{zz}_{117}
62	$C=O \cdots {}^+N–H$	-	-	3	$n_O \rightarrow \sigma^*_{N–H}$	D^{nz}_{12}
63	$H–O \cdots {}^+N–H$	-	-	1	$n_O \rightarrow \sigma^*_{N–H}$	D^{nz}_{15}
N $\cdots$ C contacts						
64	$N \cdots C=O$	3	-	-	$n_N \rightarrow \pi^*_{C=O}$	D^{nn}_{130}
N $\cdots$ N contacts						
65	$N \cdots N–H$	2	-	-	$n_N \rightarrow \sigma^*_{N–H}$	D^{nn}_{173}
C $\cdots$ H contacts						
66	$H–C_\beta \cdots H–C_\alpha$	1	2	2	$\sigma_{H–C} \rightarrow \sigma^*_{H–C}$	D^{nn}_{70}
67	$H–C_\beta \cdots H–C_\beta$	2	2	2	$\sigma_{H–C} \rightarrow \sigma^*_{H–C}$	D^{nn}_{205}
68	$H–C_\beta \cdots H–S$	2	1	3	$\sigma_{H–C} \rightarrow \sigma^*_{H–S}$	D^{nn}_{105}
69	${}^-O=C \cdots H–C_\alpha$	-	1	-	$\pi_{O=C} \rightarrow \sigma^*_{H–C}$	D^{zz}_{32}
70	${}^-O=C \cdots H–C_\beta$	-	6	1	$\pi_{O=C} \rightarrow \sigma^*_{H–C}$	D^{zz}_{20}
S $\cdots$ S contacts						
71	$S \cdots S–H$	12	7	11	$n_S \rightarrow \sigma^*_{S–H}$	D^{nn}_{71}
S $\cdots$ C contacts						
72	$S \cdots C=O$	7	-	2	$n_S \rightarrow \pi^*_{C=O}$	D^{nn}_{135}
73	$S \cdots C=O^-$	-	12	2	$n_S \rightarrow \pi^*_{C=O}$	D^{zz}_{113}
74	$H–C_\alpha \cdots S–H$	-	1	-	$\sigma_{H–C} \rightarrow \sigma^*_{S–H}$	D^{zz}_{109}
75	$H–C_\beta \cdots S–H$	-	1	-	$\sigma_{H–C} \rightarrow \sigma^*_{S–H}$	D^{zz}_{110}

Table 4. *Cont.*

Label	Type	$n \cdots n$	$N_i^{A \cdots B}$ $z \cdots z$	$n \cdots z$	$\phi_d \to \phi_a$	Example
S$\cdots$O contacts						
76	C=O$\cdots$S–H	10	-	10	$n_O \to \sigma^*_{S-H}$	D^{nn}_{127}
77	C=O$^-\cdots$S–H	-	27	11	$\pi_{C=O} \to \sigma^*_{S-H}$	D^{zz}_{51}
78	S$\cdots$O–H	26	-	12	$n_S \to \sigma^*_{O-H}$	D^{nn}_{148}
S$\cdots$N contacts						
79	N$\cdots$S–H	9	-	3	$n_N \to \sigma^*_{S-H}$	D^{nn}_{173}
80	S$\cdots ^+$N–H	-	3	3	$n_S \to \sigma^*_{N-H}$	D^{zz}_{178}

3.2.1. Interaction Distances

Figure 3 shows the distribution of the distances associated with individual intermolecular contacts separated by interaction type, that is, primary and secondary hydrogen bonds, dihydrogen bonds, and exotic contacts for all dimers. Remarkably, the spectrum of A$\cdots$B distances for direct intermolecular contacts covers a wide range, from the very short (1.50 Å for a C=O$^- \cdots$H–O in D^{nz}_{38}) to the very large (4.19 Å for the exotic S$\cdots$S in D^{nn}_{77}), which sensibly departs in both directions from the reference 1.98 Å in the isolated gas phase water dimer. Notice that regardless of the constituting monomers, only primary hydrogen bonds fall below 1.98 Å. In a classical sense, a zwitterion may be conveniently seen as two remote charge islands within the same molecule, in this view, the effect of the charges in the structural complexity of the cysteine dimers is clear: on one hand, intermolecular distances are reduced for the dimers with more charge islands, i.e., $r^{nn}_{AB} > r^{nz}_{AB} > r^{zz}_{AB}$, on the other, the structural complexity is also sensibly reduced for the more charge-separated species because the $n \cdots n$ distributions have more peaks than $n \cdots z$ which in turn have more peaks than $z \cdots z$. In addition, it may be argued that among all the types of interactions stabilizing the cysteine dimers, salt bridges should be the strongest and thus the most important structural determining factor whenever they occur. Indeed, the lowest energy $z \cdots z$ dimers with populations larger than 5% shown in Figure 1, are actually stabilized by two salt bridges. Notice that the center of the peak for the distribution of C=O$^- \cdots$H-N$^+$ distances (1.66 Å, Figure 3C) is actually larger than 1.57 Å, the center of the peak for the distribution of C=O$^- \cdots$H-O interactions (Figure 3B), which are *a priori* not as strong as the salt bridges but which dictate the structures of the $n \cdots z$ dimers, the reason for this apparent contradiction is that formation of the two salt bridges confers structural rigidity to the clusters.

When immersed in a continuum aqueous environment, there is partial dissociation of the O–H bonds upon the formation of the dimers. Figure 4 shows the changes in the corresponding distances and Wiberg Bond Indices (WBI) compared against the reference monomer. Evidence for partial dissociation is provided by the peak centered at ≈0.58 WBI, which actually corresponds to O–H groups of the low energy, high population dimers where neutral monomers are involved.

3.2.2. Electron Densities at Bond Critical Points, $\rho(\mathbf{r}_c)$

The relationship between electron density at bond critical points and the nature of the interaction is clear: large accumulations of electron densities at BCPs indicate that the electrons are shared between two fragments or atoms, otherwise known as covalent bonding. Conversely, small electron densities at BCPs indicate that the electrons are displaced towards the nuclei, thus signaling either ionic bonding or long range interactions. Figure 5 shows the values for the calculated electron densities at the 2894 bond critical points associated to intermolecular interactions in the 746 cysteine dimers. Electron densities at those points cover the $[9.1 \times 10^{-4}, 7.6 \times 10^{-2}]$ a.u. interval. These values are sensibly smaller than the 0.24 and 0.35 a.u obtained for the covalent C–C and O–H bonds in D^{nn}_1. The smaller electron densities correspond to secondary HBs and exotic contacts while among primary HBs, those with the smallest densities involve the S–H group. Only some primary HBs and salt bridges have larger densities than the reference water dimer. The

distributions plotted in Figure 5 are wide; thus, there are many possibilities for the same type of interaction. Finally, as expected [49,64–66], there seems to be an inverse correlation between interaction distance and electron density at intermolecular BCPs.
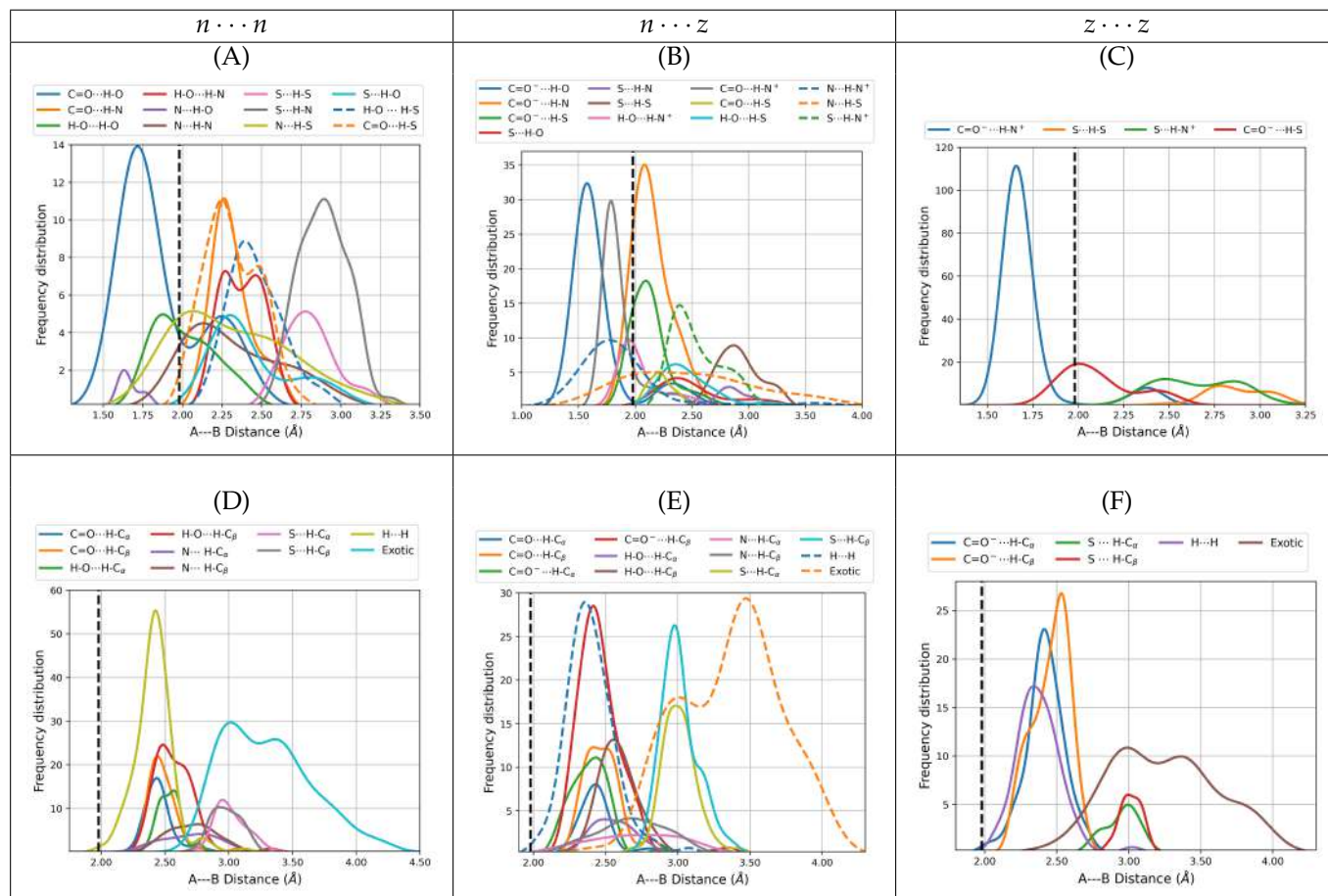

Figure 3. Distributions of the A$\cdots$B distances for intermolecular contacts for all dimers of cysteine found in this work. The distributions are fitted to the actual histograms, so the center of the peaks of the distributions are statistically relevant. The top row shows only primary hydrogen bonds including salt bridges. The bottom row shows secondary hydrogen bonds, dihydrogen bonds, and all exotic interactions. The left column is reserved for the $n\cdots n$ dimers (subfigures (**A,D**)), the middle column for $n\cdots z$ (subfigures (**B,E**)) and the right column for $z\cdots z$ (subfigures (**C,F**)). All distances taken from the B3LYP–D3/6–311++G(d, p) potential energy surfaces with water represented as a continuum solvent. The dashed vertical lines at 1.98 Å mark the reference H$\cdots$O distance in the gas phase water dimer.

Figure 4. Variation of the O–H bond properties as a consequence of intermolecular interactions. The reference values for the n_2 cysteine monomer are included as vertical dashed lines.

Figure 5. Electron densities at bond critical points for intermolecular contacts for all dimers of cysteine found in this work. The top row shows only primary hydrogen bonds including salt bridges. The bottom row shows secondary hydrogen bonds, dihydrogen bonds, and all exotic interactions. The left column is reserved for the $n \cdots n$ dimers (subfigures (**A**,**D**)), the middle column for $n \cdots z$ (subfigures (**B**,**E**)) and the right column for $z \cdots z$ (subfigures (**C**,**F**)). All values taken from the B3LYP–D3/6–311++G(d, p) potential energy surfaces with water represented as a continuum solvent. The dashed vertical lines mark the reference value for the gas phase water dimer.

3.2.3. Bond Degree Parameters $\mathcal{H}(\mathbf{r}_c)/\rho(\mathbf{r}_c)$

The bond degree parameter is related to chemical bonding as follows. Kinetic energy is everywhere positive and repulsive ($mv^2/2 = p^2/2m > 0$ in classical mechanics) while potential energy is everywhere negative and attractive. The total energy is the sum of kinetic and potential energies, $\mathcal{H} = \mathcal{G} + \mathcal{V}$; thus, its sign reveals the winner of the local kinetic vs potential energy tug of war and dictates the nature of the interaction. Indeed, positive total energies at BCPs are obtained when there is a local dominance of the repulsive kinetic energy, indicating local depletion of electrons in the internuclear region and displacement of the electron density associated with the particular bonding interactions towards the nuclei. Conversely, negative total energies are obtained when there is a local dominance of the attractive potential energy indicating that there is shared electron density concentrated in the internuclear region and signaling an increasingly covalent character of the interaction. An alternative rigorous physical meaning to energy densities is offered by a dimensional analysis: energy density has units of pressure ($E/V = F/A = P$); thus, local negative energy densities may be equated to negative quantum pressures which strongly attract electrons towards the BCP, indicating increasingly covalent interactions while local positive energy densities correspond to positive quantum pressures that push electrons away from the BCPs towards the nuclei, indicating anionic or long range interactions.

It is well known that the sign of $\nabla^2\rho(\mathbf{r}_c)$ is not a sufficient criterium to establish the nature of the interaction in every case [28,67–70], specifically, it is quite often the case that a particular interaction has both positive Laplacian and negative bond degree parameter at the same time. Thus, the bond degree parameter is used in conjunction with the Laplacian of the electron density at BCPs to remove any ambiguity according to Rozas et al. [68]: weak to medium strength hydrogen bonds have both $\nabla^2\rho(\mathbf{r}_c), \mathcal{H}(\mathbf{r}_c)/\rho(\mathbf{r}_c) > 0$, strong hydrogen bonds have $\nabla^2\rho(\mathbf{r}_c) > 0, \mathcal{H}(\mathbf{r}_c)/\rho(\mathbf{r}_c) < 0$ and very strong HBs have both $\nabla^2\rho(\mathbf{r}_c), \mathcal{H}(\mathbf{r}_c)/\rho(\mathbf{r}_c) < 0$. Figure 6 plots distributions of the bond degree parameters for all dimers found in this work. It is clear from the distributions of $\mathcal{H}(\mathbf{r}_c)/\rho(\mathbf{r}_c)$ that all intermolecular contacts found here cover a wide spectrum of possibilities with a substantial number of only primary hydrogen bonds or salt bridges having $\mathcal{H}(\mathbf{r}_c)/\rho(\mathbf{r}_c) < 0$ (Figure 6A–C), thus should be considered as strong contacts by the above criteria. The wide spectrum of bond degree parameters, the large number of structural possibilities and the strong character of the interactions have deep implications in the biological role of cysteine and of the aminoacids that make up proteins and biomolecules: similar results have been obtained for example in the interactions between the spike protein of SARS-COV-2 and the ACE2 receptors [48,71] and between the envelope protein of the Zika virus and the glycosaminoglycans that act as receptors [47]. In the case of SARS-COV-2, the formation of strong salt bridges and hydrogen bonds is one of the main factors of the pressure driving the evolution of the virus towards new variants. For the cysteine dimers, many of the primary hydrogen bonds with positive bond degree parameters are located to the left of the reference isolated gas phase water dimer, which confers them medium to strong character. All secondary and exotic contacts (Figure 6D–F) exhibit positive bond degree parameters and many areas actually to the right of the reference water dimer; thus, they are classified as weak. As a general rule, hydrogen bonds involving the carbonyl, carboxylate and amino groups as electron donors and the hydroxyl, amino and ammonium groups as electron acceptors, are the ones with highly negative $\mathcal{H}(\mathbf{r}_c)/\rho(\mathbf{r}_c)$ values. Some HBs involving the thiol group, either as donor or acceptor, have slightly negative bond degree parameters.

3.2.4. Virial Ratios, $|\mathcal{V}(\mathbf{r}_c)|/\mathcal{G}(\mathbf{r}_c)$

Analysis of the virial ratios at bond critical points serves as a more quantitative description of the nature of the interactions than the Laplacians of the electron density and the bond degree parameters. See the works of Grabowski [28] and of Rozas et al. [68] for a formal description of how the virial ratio is related to bonding. In short, local depletion of electron density (local dominance of the repulsive kinetic energy), which is indicative of ionic or long interactions have $0 < |\mathcal{V}(\mathbf{r}_c)|/\mathcal{G}(\mathbf{r}_c) < 1$, local concentration of electron density (local dominance of the attractive potential energy), indicative of covalent interactions have $|\mathcal{V}(\mathbf{r}_c)|/\mathcal{G}(\mathbf{r}_c) > 2$, and the $1 < |\mathcal{V}(\mathbf{r}_c)|/\mathcal{G}(\mathbf{r}_c) < 2$ interval describes interactions with mixed contributions.

Figure 7 shows the distribution of the virial ratios for all dimers found in this work, which cover the [0.61, 1.49] interval for primary hydrogen bonds and salt bridges (Figure 7A–C) and the [0.56, 0.92] interval for secondary HBs, exotic and dihydrogen contacts (Figure 7A–C). Notice that as in the analysis of the previous descriptors, the fitted distributions go a little beyond the actual limits. It is quite revealing that a large number of contacts, especially those involving the carbonyl group have virial ratios larger than 1, which confers them a high degree of covalency while not being formal bonds. Interestingly, these include the charged carboxylate which may naively be thought as being involved in highly ionic contacts. Virial ratios larger than 1 transcend the carbonyl group, which is indeed the case for the following HBs: N$\cdots$H–O, C=O$\cdots$H–O, S$\cdots$H–O, N$\cdots$H–S for $n\cdots n$ dimers, C=O$^-\cdots$H–O, N$\cdots$H–N$^+$, S$\cdots$H–N$^+$, C=O$\cdots$H–N$^+$ for $n\cdots z$ and for all salt bridges in $z\cdots z$. This high covalency of the *a priori* ionic contacts has been reported for other cases, including for example the microsolvation of charged species [49,53,72]. Most secondary HBs, H$\cdots$H, and exotic contacts have virials smaller than the water dimer reference.

Surprisingly, the thiol group in a number of cases is involved in stronger interactions than the H–O···H–O and N···H–O contacts.
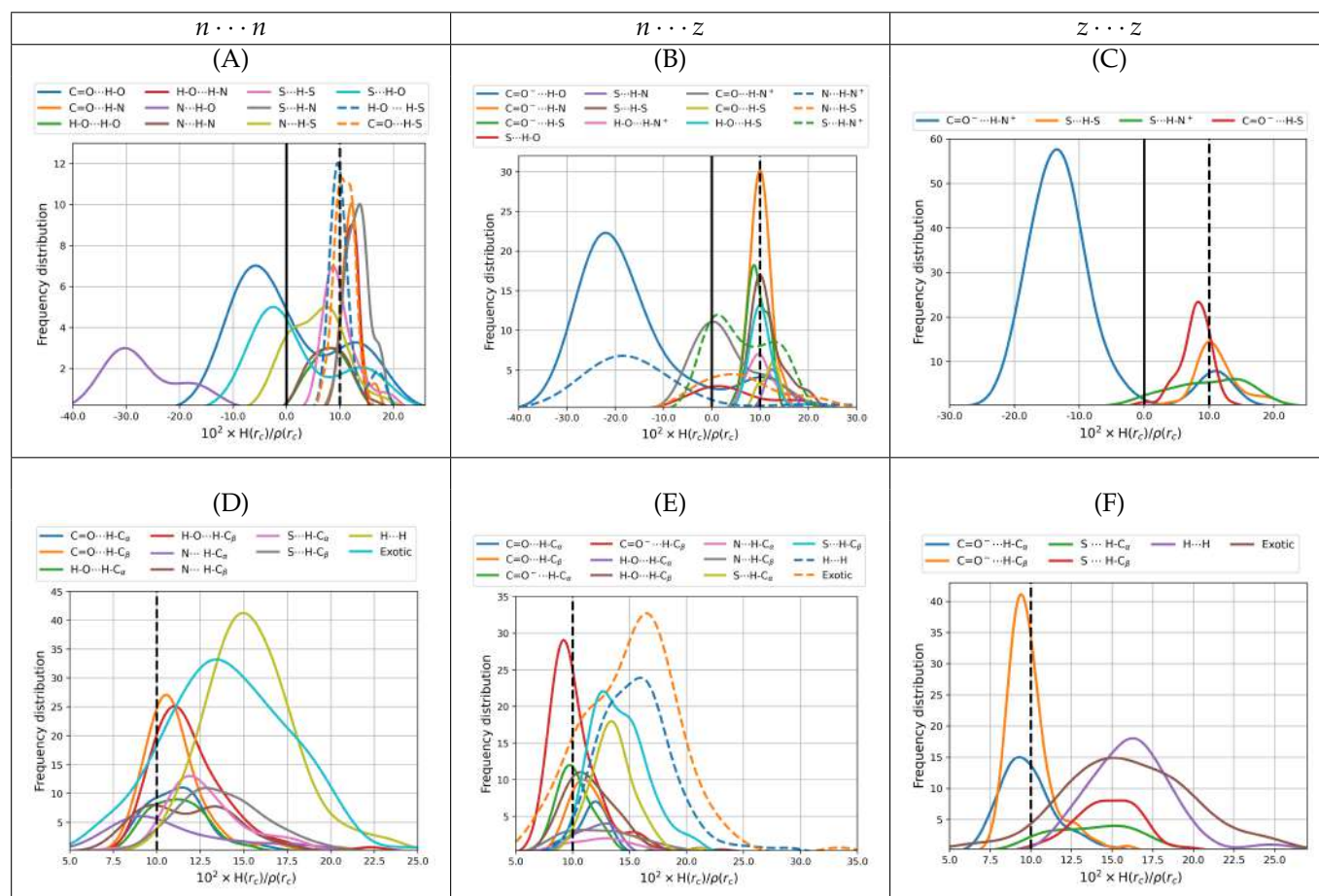

Figure 6. Bond degree parameters for intermolecular contacts for all dimers of cysteine found in this work. The top row shows only primary hydrogen bonds including salt bridges. The bottom row shows secondary hydrogen bonds, dihydrogen bonds, and all exotic interactions. The left column is reserved for the $n \cdots n$ dimers (subfigures (**A**,**D**)), the middle column for $n \cdots z$ (subfigures (**B**,**E**)) and the right column for $z \cdots z$ (subfigures (**C**,**F**)). All values taken from the B3LYP–D3/6–311++G(d, p) potential energy surfaces with water represented as a continuum solvent. Solid vertical lines mark the QTAIM boundaries separating locally stabilizing from the locally destabilizing interactions. Dashed vertical lines mark the reference value for the gas phase water dimer.

3.2.5. NBO and NCI Picture of Intermolecular Interactions

Intermolecular interactions have been successfully studied under the NBO formalism in a wide range of problems [30,33]. In the particular case of the cysteine dimers, we proceeded to identify the localized donor Lewis orbitals from which charge is transferred to acceptor orbitals according to the $\phi_d \rightarrow \phi_a$ scheme. Tables 3 and 4 list the involved orbitals for each one of the 80 types of interactions found in this work, Figure 1 provides the corresponding surfaces for those dimers with populations larger than 5%. Once identified, we quantified the strength of the orbital interaction by second order perturbation theory on the Fock matrix as given by $-E_{d \rightarrow a}^{(2)} = q_d |\langle \phi_d | \mathcal{F} | \phi_a \rangle|^2 / (E_a - E_d)$. With this procedure, the strength of the interaction is directly related to the magnitude of $E_{d \rightarrow a}^{(2)}$.
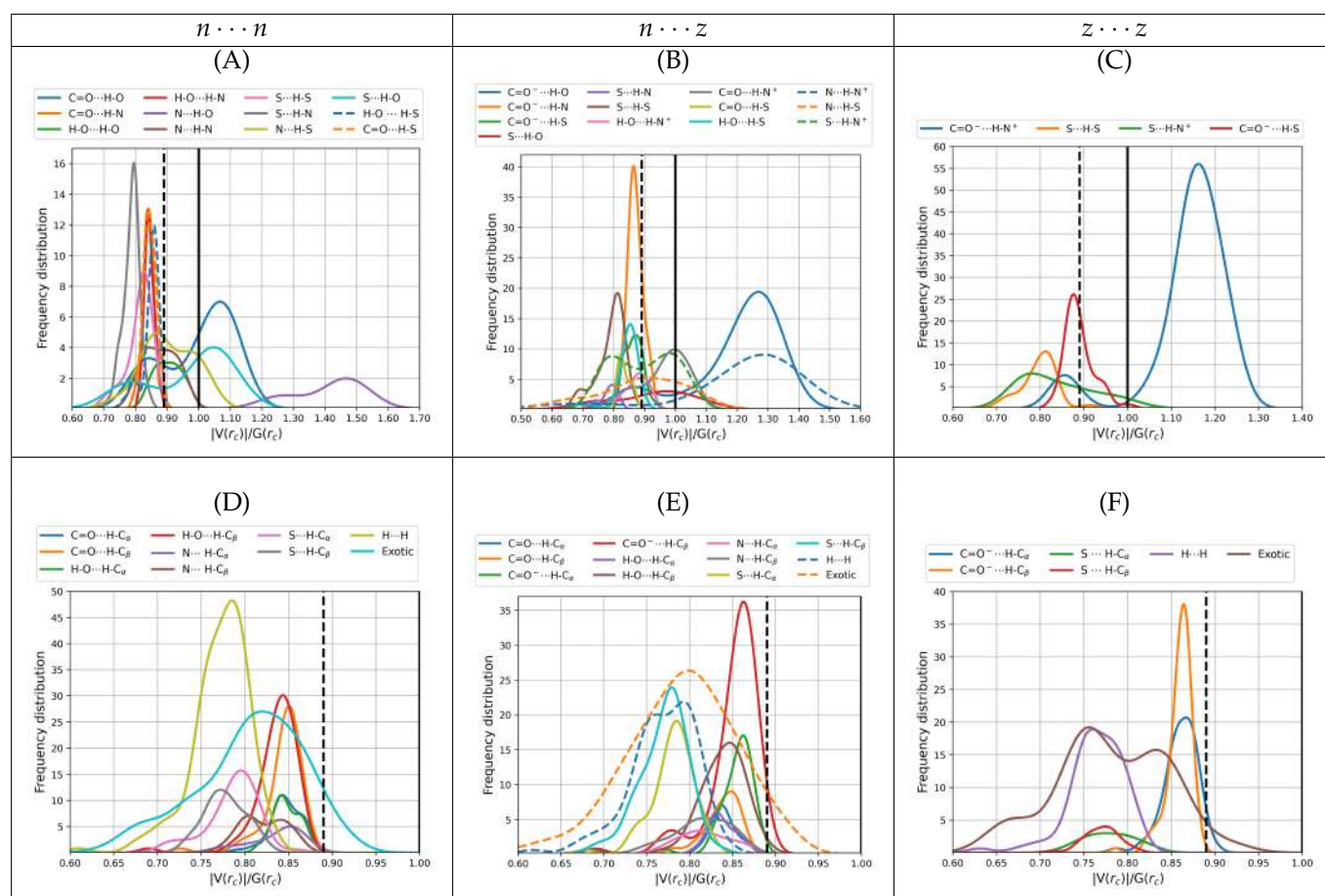
Figure 7. Virial ratios at bond critical points for intermolecular contacts for all dimers of cysteine found in this work. The top row shows only primary hydrogen bonds including salt bridges. The bottom row shows secondary hydrogen bonds, dihydrogen bonds, and all exotic interactions. The left column is reserved for the $n \cdots n$ dimers (subfigures (**A,D**)), the middle column for $n \cdots z$ (subfigures (**B,E**)) and the right column for $z \cdots z$ (subfigures (**C,F**)). All values taken from the B3LYP–D3/6–311++G(d, p) potential energy surfaces with water represented as a continuum solvent. Vertical solid lines mark the QTAIM boundaries separating long range from intermediate character interactions. Dashed vertical lines mark the reference value for the gas phase water dimer.

Donor $\rightarrow$ acceptor orbital interactions resulting in primary hydrogen bonds and salt bridges in $n \cdots n$, $n \cdots z$ and $z \cdots z$ cysteine dimers include everything from the very weak to the very strong, covering the wide [0.06, 50.30] kcal/mol interval (Figure 8A–C). Salt bridges exhibit an uncommonly complex distribution of energies with several shoulders. Interestingly, the strongest contacts are not salt bridges but rather N$\cdots$H–O primary hydrogen bonds, listed as interaction **10** in Table 3 and shown in Figure 8A. This interaction type, which is also the strongest intermolecular contact found in alanine dimers [13], arises from $n_N \rightarrow \sigma^*_{H-O}$ charge transfer in $n \cdots n$ dimers. Next in the strength hierarchy are the highly ionic C=O$^-$ $\cdots$ H–O and N$\cdots$H–N$^+$ contacts arising from $n_O \rightarrow \sigma^*_{H-O}$ and $n_N \rightarrow \sigma^*_{H-N}$ charge transfers in $n \cdots z$ dimers. These are listed as interactions **2, 12** with the corresponding distributions shown in Figure 8B. C=O$^-$ $\cdots$ H–N$^+$ salt bridges in $z \cdots z$ dimers come only third in the interaction energy scale. They arise from $n_O \rightarrow \sigma^*_{H-N}$ charge transfer, are listed as interaction **6** and the corresponding distributions are shown in Figure 8C. These sets of interactions are present on the structures with populations higher than 5%. In a manner consistent with the QTAIM descriptors analyzed above, a large number of primary HBs and salt bridges have interaction energies larger than 6.63 [30] kcal/mol, the orbital interaction energy for the reference water dimer, however, no secondary hydrogen bond, no exotic contact and no dihydrogen bond exceed the reference.
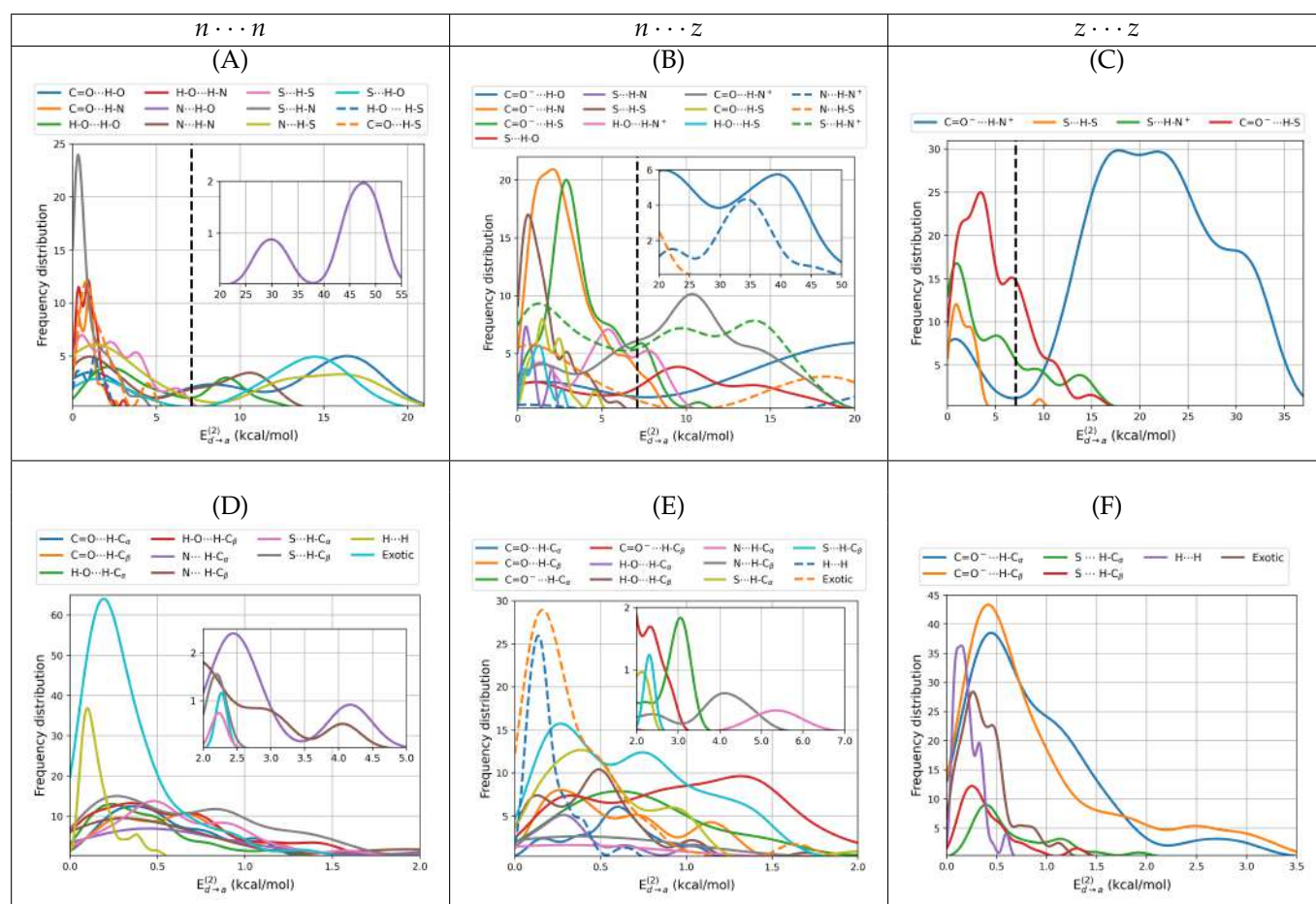
Figure 8. Donor $\cdots$ acceptor NBO energies for intermolecular contacts for all dimers of cysteine found in this work. The top row shows only primary hydrogen bonds including salt bridges. The bottom row shows secondary hydrogen bonds, dihydrogen bonds, and all exotic interactions. The left column is reserved for the $n \cdots n$ dimers (subfigures (**A**,**D**)), the middle column for $n \cdots z$ (subfigures (**B**,**E**)) and the right column for $z \cdots z$ (subfigures (**C**,**F**)). All values taken from the B3LYP–D3/6–311++G(d, p) potential energy surfaces with water represented as a continuum solvent. The dashed vertical lines mark the reference value for the gas phase water dimer.

As stated above, primary HBs and salt bridges determine the structure of the dimers. According to Table 3, they always arise from orbital interactions of the $n_X \rightarrow \sigma^*_{H-Y}$ type with X,Y = O, N, S. As also stated above, secondary HBs, dihydrogen bonds and exotic contacts are usually a consequence of the structure. Notwithstanding, the orbitals involved in the weaker interactions offer a quite interesting and rather uncommon picture. First, notice that all exotic O$\cdots$O contacts (**53–58** in Table 4) put the two negative ends of the fragments with various degrees of negative character in direct contact, with the most severe case being interaction **54** connecting two formal negative charges with no intermediaries. Second, notice that the 139 interactions grouped into **58, 61–63, 65, 71, 76, 78–80** may all be described by the general $n_X \rightarrow \sigma^*_{Y-H}$ charge transfer scheme with X, Y = O, N, S. These correspond to what David et al. [13] have called inverted hydrogen bonds because the lone pair on X donates electron charge to an antibonding σ^*_{Y-H} orbital which is inverted from the usual σ^*_{H-Y}, in other words, the charge donation occurs between two orbitals overlapping from one negative atom to another negative atom with no bridging proton. Anti electrostatic hydrogen bonds of the type described in this paragraph have been reported by Weinhold and Klein [73].

Figure 9 shows the obtained non covalent surfaces as well as the thoroughs of the reduced gradients for the largest binding energy dimers in the $n \cdots n$, $n \cdots z$, $z \cdots z$ po-

tential energy surfaces, we include the corresponding interacting NBOs to help visualize the source of the NCI surfaces. See Figures S3–S5 in the supplementary material for the corresponding surfaces in all dimers with populations larger than 5%. The standard NCI color code [34,35] assigns green surfaces to weakly bonding contacts and blue surfaces to strong interactions, for the thoroughs, negative values of $\text{sign}(\lambda_2)\rho$ reveal bonding interactions [34] which are weaker when $\text{sign}(\lambda_2)\rho \approx 0$. NCI reveals that the two salt bridges in the $z \cdots z$ dimers are actually very similar (in fact, they cannot be told apart in the thoroughs) and that in most cases, large stabilizing surfaces arising from individual contacts transferring tiny amounts of charge to the interstitial region have significant contributions to the overall binding of the dimers. Unexpectedly, these charge transfer contributions are major contributors to the charged cases. Notice that charge transfer to the interstitial region appears to be the norm when several molecular units are stabilized via non covalent interactions: these fluxional surfaces of charge have been found to be a major player in the molecular interpretation of hydrophobicity [46], in the initial recognition and attachment of viruses to cell receptors [47,48,71], in the microsolvation and encapsulation of charged and neutral species, in the microscopic structure of ionic liquids [51], etc.
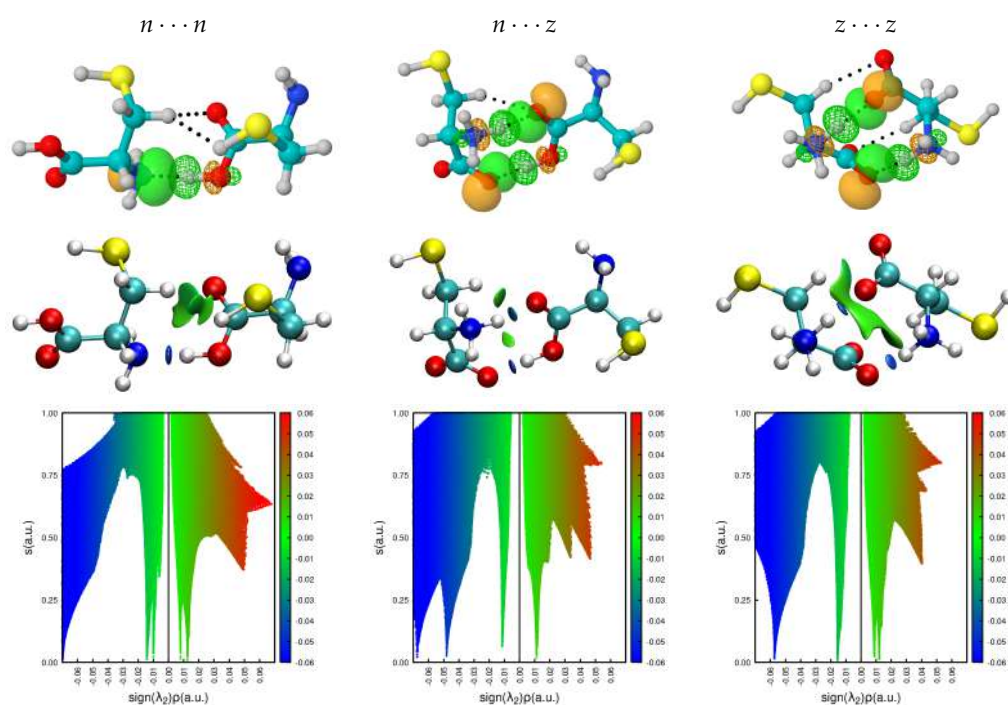

Figure 9. NCI surfaces and thoroughs describing the intermolecular contacts for the highest binding energy cysteine dimers. NBO pictures are also included to ease visualization of the interactions.

4. Discussion and Context

Accurate description and characterization of chemical bonding is a notoriously hard problem in chemistry, whose difficulty is magnified when dealing with weak intermolecular non covalent interactions. When studying molecules and their interactions, a large portion of the conceptual framework developed by experimentalists and theoreticians invokes a number of useful ideas that correspond to non observable quantities (partial atom charges, orbital interactions, virial ratios at BCPs, etc.); thus, there are no quantum mechanical operators whose expected values may be used to calculate them and therefore, approximate methods, however accurate, are used to determine these quantities. This approach has a fundamental problem: each quantity may be obtained by several methods and the results quite often vary among them. In this context, it is impressive and certainly reassuring that QTAIM, NBO and NCI, which are conceptually and methodologically substantially

different, afford a consistent, complementary picture of intermolecular bonding in the dimers of cysteine.

The large number of interactions, isomers, and types of binary contacts are intimately connected to the problem of molecular evolution and to the complexity of life observable on this planet. By virtue of the large number of accessible states, molecular systems where specific cysteine to cysteine contacts are observed are thermodynamically favored because of the ever increasing entropy of the universe, in other words, this type of systems will evolve towards equilibrium states with large structural diversity. Since the interactions dissected here are responsible for the molecular interactions between all aminoacid pairs, this argument of entropy driving molecular evolution readily applies to large proteins and biomolecules.

5. Summary and Conclusions

An intensive exploration of the potential energy surfaces for the interaction of neutral and charged cysteine monomers to form dimers in an aqueous environment represented by a continuum afforded a large number of isomers, amounting to 746 well characterized local minima. The isomers with the largest population are distributed within small energy differences of the putative global minima. Ten potential energy surfaces were explored in total for the $n \cdots n$, $n \cdots z$, and $z \cdots z$ combinations with two neutral (n) and two zwitterionic (z) forms. A number of strongly bound dimers were found, with interaction energies exceeding 20 kcal/mol in several cases and with interaction distances covering the very small to the very large in the [1.50, 4.19] Å interval. The nature of intermolecular bonding interactions was dissected using QTAIM, NCI, and NBO, three conceptually different methods, which for the present case afford consistent, complementary pictures. A total of 80 types of different intermolecular contacts were found in this complex and large universe of dimers. As a general rule, primary hydrogen bonds and salt bridges are the strongest of the interactions and determine the molecular geometry, conversely, secondary hydrogen bonds, exotic $X \cdots Y$ (X = C, N, O, S) and $H \cdots H$ dihydrogen contacts are weaker and most often a consequence of the structure. All interactions, even the highly ionic, may be described by the $\phi_d \rightarrow \phi_a$ orbital charge transfer scheme, leading to accumulation or depletion of electron density at the bond critical points as revealed by topological analysis of the electron densities. The large binding energies mentioned above are the result of unusually strong charge assisted hydrogen bonds and salt bridges. We found that the highly ionic salt bridges have large degrees of covalency; thus, a simplistic electrostatic attraction between positively and negatively charged fragments does not suffice for a proper account of bonding interactions whenever the zwitterions are involved. The weaker secondary hydrogen bonds, exotic $X \cdots Y$ and dihydrogen contacts in no few cases are stronger than, for example, the archetypal hydrogen bond in the water dimer. Moreover, independent of how strong or weak individual interactions are, their collective action cannot be ignored because they lead to the formation of large attractive non-covalent surfaces in the interstitial region between the fragments. A few antielectrostatic contacts [73] as well as a few inverted hydrogen bonds [13] in which the charge transfer occurs between the two negative ends of the fragments, were found; thus, they appear to be of common occurrence in nature.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/molecules27248665/s1, Figure S1: Electrostatic potential surfaces for the cysteine monomers.; Figure S2: Distributions of the Laplacians of the electron densities at bond critical points for all dimers.; Figures S3–S5: Additional NBO and NCI plots of descriptors of bonding interactions for the dimers with populations larger than 5%; Table S1: Binding energies and energy differences for neutral dimers.; Table S2: Binding energies and energy differences for mixed dimers. Table S3: Binding energies and energy differences for zwitterionic dimers. Cartesian coordinates for the entire set of 746 dimers are also provided.

References

1. Wilke, J.J.; Lind, M.C.; Schaefer, H.F., III; Császár, A.G.; Allen, W.D. Conformers of gaseous cysteine. *J. Chem. Theory Comput.* **2009**, *5*, 1511–1523. [CrossRef] [PubMed]

2. Giraldo, C.; Gómez, S.; Weinhold, F.; Restrepo, A. Insight into the Mechanism of the Michael Reaction. *ChemPhysChem* **2016**, *17*, 2022–2034. [CrossRef] [PubMed]

3. Cardona, W.; Guerra, D.; Restrepo, A. Reactivity of δ-substituted α, β-unsaturated cyclic lactones with antileishmanial activity. *Mol. Simul.* **2014**, *40*, 477–484. [CrossRef]

4. Clayden, J.; Greeves, N.; Warren, S. *Organic Chemistry*; Oxford University Press: Oxford, UK, 2012.

5. Santos, M.; Moreira, R. Michael acceptors as cysteine protease inhibitors. *Mini-Rev. Med. Chem.* **2007**, *7*, 1040–1050. [CrossRef]

6. Gómez, S.; Ramírez-Malule, H.; Cardona-G, W.; Osorio, E.; Restrepo, A. Double-ring epimerization in the biosynthesis of clavulanic acid. *J. Phys. Chem. A* **2020**, *124*, 9413–9426. [CrossRef] [PubMed]

7. Lu, X.; Galkin, A.; Herzberg, O.; Dunaway-Mariano, D. Arginine deiminase uses an active-site cysteine in nucleophilic catalysis of L-arginine hydrolysis. *J. Am. Chem. Soc.* **2004**, *126*, 5374–5375. [CrossRef] [PubMed]

8. Li, H.; Thomas, G.J., Jr. Studies of virus structure by Raman spectroscopy. Cysteine conformation and sulfhydryl interactions in proteins and viruses. 1. Correlation of the Raman sulfur-hydrogen band with hydrogen bonding and intramolecular geometry in model compounds. *J. Am. Chem. Soc.* **1991**, *113*, 456–462. [CrossRef]

9. Dobrowolski, J.C.; Rode, J.E.; Sadlej, J. Cysteine conformations revisited. *J. Mol. Struct. THEOCHEM* **2007**, *810*, 129–134. [CrossRef]

10. Wiedemann, C.; Kumar, A.; Lang, A.; Ohlenschläger, O. Cysteines and disulfide bonds as structure-forming units: insights from different domains of life and the potential for characterization by NMR. *Front. Chem.* **2020**, *8*, 280. [CrossRef]

11. Fraser, R.; MacRae, T.P.; Sparrow, L.G.; Parry, D. Disulphide bonding in α-keratin. *Int. J. Biol. Macromol.* **1988**, *10*, 106–112. [CrossRef]

12. Bairoch, A.; Apweiler, R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* **2000**, *28*, 45–48. [CrossRef]

13. David, J.; Gómez, S.; Guerra, D.; Guerra, D.; Restrepo, A. A Comprehensive Picture of the Structures, Energies, and Bonding in the Alanine Dimers. *ChemPhysChem* **2021**, *22*, 2401–2412. [CrossRef]

14. Tonino, P.; Kiss, B.; Strom, J.; Methawasin, M.; Smith, J.E.; Kolb, J.; Labeit, S.; Granzier, H. The giant protein titin regulates the length of the striated muscle thick filament. *Nat. Commun.* **2017**, *8*, 1–11. [CrossRef]

15. Gronert, S.; O'Hair, R.A. Ab initio studies of amino acid conformations. 1. The conformers of alanine, serine, and cysteine. *J. Am. Chem. Soc.* **1995**, *117*, 2071–2081. [CrossRef]

16. Dobrowolski, J.C.; Jamróz, M.H.; Kołos, R.; Rode, J.E.; Sadlej, J. Theoretical Prediction and the First IR Matrix Observation of Several l-Cysteine Molecule Conformers. *ChemPhysChem* **2007**, *8*, 1085–1094. [CrossRef]

17. Sanz, M.E.; Blanco, S.; López, J.C.; Alonso, J.L. Rotational probes of six conformers of neutral cysteine. *Angew. Chem. Int. Ed.* **2008**, *47*, 6216–6220. [CrossRef] [PubMed]

18. Kaminski, M.; Kudelski, A.; Pecul, M. Vibrational optical activity of cysteine in aqueous solution: A comparison of theoretical and experimental spectra. *J. Phys. Chem. B* **2012**, *116*, 4976–4990. [CrossRef]

19. Sadlej, J.; Dobrowolski, J.C.; Rode, J.E.; Jamróz, M.H. Density functional theory study on vibrational circular dichroism as a tool for analysis of intermolecular systems:(1: 1) cysteine- water complex conformations. *J. Phys. Chem. A* **2007**, *111*, 10703–10711. [CrossRef]

20. Liu, S.Z.; Wang, H.Q.; Zhou, Z.Y.; Dong, X.L.; Gong, X.L. Theoretical study of helical structure caused by chirality of cysteine dimer. *Int. J. Quant. Chem.* **2005**, *105*, 66–73. [CrossRef]

21. Tiwari, S.; Mishra, P. Vibrational spectra of cysteine zwitterion and mechanism of its formation: Bulk and specific solvent effects and geometry optimization in aqueous media. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2009**, *73*, 719–729. [CrossRef]

22. Bachrach, S.M.; Nguyen, T.T.; Demoin, D.W. Microsolvation of cysteine: A density functional theory study. *J. Phys. Chem. A* **2009**, *113*, 6172–6181. [CrossRef]

23. Chapman, C.R.; Ting, E.C.; Kereszti, A.; Paci, I. Self-assembly of cysteine dimers at the gold surface: A computational study of competing interactions. *J. Phys. Chem. C* **2013**, *117*, 19426–19435. [CrossRef]

24. Buimaga-Iarinca, L.; Calborean, A. Electronic structure of the ll-cysteine dimers adsorbed on Au (111): A density functional theory study. *Phys. Scr.* **2012**, *86*, 035707. [CrossRef]

25. Ieritano, C.; Carr, P.J.; Hasan, M.; Burt, M.; Marta, R.A.; Steinmetz, V.; Fillion, E.; McMahon, T.B.; Hopkins, W.S. The structures and properties of proton-and alkali-bound cysteine dimers. *Phys. Chem. Chem. Phys.* **2016**, *18*, 4704–4710. [CrossRef]

26. Bader, R. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, UK, 1990.

27. Popelier, P.L. *Atoms in Molecules: An Introduction*; Prentice Hall: London, UK, 2000.

28. Grabowski, S.J. What Is the Covalency of Hydrogen Bonding? *Chem. Rev.* **2011**, *111*, 2597–2625. [CrossRef] [PubMed]

29. Becke, A. *The Quantum Theory of Atoms in Molecules: From Solid State to DNA and Drug Design*; John Wiley & Sons: Hoboken, NJ, USA, 2007.

30. Reed, A.E.; Curtiss, L.A.; Weinhold, F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* **1988**, *88*, 899–926. [CrossRef]

31. Weinhold, F.; Landis, C.R. *Discovering Chemistry with Natural Bond Orbitals*; Wiley-VCH: Hoboken NJ, USA, 2012; 319p.

32. Glendening, E.D.; Landis, C.R.; Weinhold, F. Natural bond orbital methods. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 1–42. [CrossRef]

33. Weinhold, F.; Landis, C.; Glendening, E. What is NBO analysis and how is it useful? *Int. Rev. Phys. Chem.* **2016**, *35*, 399–440. [CrossRef]

34. Johnson, E.R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A.J.; Yang, W. Revealing Noncovalent Interactions. *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506. [CrossRef]

35. DiLabio, G.A.; Otero-de-la Roza, A., Noncovalent Interactions in Density Functional Theory. In *Reviews in Computational Chemistry*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2016; Chapter 1, pp. 1–97. [CrossRef]

36. Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092. [CrossRef]

37. Kirkpatrick, S.; Gelatt, C.D.; Vecchi, M.P. Optimization by Simulated Annealing. *Science* **1983**, *220*, 671–680. [CrossRef] [PubMed]

38. van Laarhoven, P.; Aarts, E. *Simulated Annealing: Theory and Applications*; Mathematics and Its Applications; Springer: Dordrecht, The Netherlands, 1987.

39. Pérez, J.; Restrepo, A. ASCEC V02: Annealing Simulado con Energía Cuántica. In *Property, Development, and Implementation: Grupo de Química–Física Teórica*; Instituto de Química, Universidad de Antioquia: Medellín, Colombia, 2008.

40. Pérez, J.F.; Hadad, C.Z.; Restrepo, A. Structural studies of the water tetramer. *Int. J. Quant. Chem.* **2008**, *108*, 1653–1659. [CrossRef]

41. Pérez, J.F.; Florez, E.; Hadad, C.Z.; Fuentealba, P.; Restrepo, A. Stochastic Search of the Quantum Conformational Space of Small Lithium and Bimetallic Lithium–Sodium Clusters. *J. Phys. Chem. A* **2008**, *112*, 5749–5755. [CrossRef] [PubMed]

42. Tomasi, J.; Persico, M. Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent. *Chem. Rev.* **1994**, *94*, 2027–2094. [CrossRef]

43. Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999–3094. [CrossRef]

44. Gonzalez, C.; Restrepo-Cossio, A.; Márquez, M.; Wiberg, K.B. Ab Initio Study of the Stability of the Ylide-like Intermediate Methyleneoxonium in the Reaction between Singlet Methylene and Water. *J. Am. Chem. Soc.* **1996**, *118*, 5408–5411. [CrossRef]

45. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A.; et al. *Gaussian 09 Revision E.01*; Gaussian Inc.: Wallingford, CT, USA, 2009.

46. Gómez, S.; Rojas-Valencia, N.; Gómez, S.A.; Cappelli, C.; Merino, G.; Restrepo, A. A molecular twist on hydrophobicity. *Chem. Sci.* **2021**, *12*, 9233–9245. [CrossRef]

47. Gómez, S.A.; Rojas-Valencia, N.; Gómez, S.; Lans, I.; Restrepo, A. Initial recognition and attachment of the Zika virus to host cells: A molecular dynamics and quantum interaction approach. *ChemBioChem* **2022**, *23*, e202200351. [CrossRef]

48. Gómez, S.A.; Rojas-Valencia, N.; Gómez, S.; Egidi, F.; Cappelli, C.; Restrepo, A. Binding of SARS-CoV-2 to Cell Receptors: A Tale of Molecular Evolution. *ChemBioChem* **2021**, *22*, 724–732. [CrossRef]

49. Rojas-Valencia, N.; Gómez, S.; Guerra, D.; Restrepo, A. A detailed look at the bonding interactions in the microsolvation of monoatomic cations. *Phys. Chem. Chem. Phys.* **2020**, *22*, 13049–13061. [CrossRef]

50. Moreno, N.; Hadad, C.Z.; Restrepo, A. Microsolvation of electrons by a handful of ammonia molecules. *J. Chem. Phys.* **2022**, *157*, 134301. [CrossRef]

51. Correa, E.; Montaño, D.; Restrepo, A. Cation···anion bonding interactions in 1–Ethyl–3–Methylimidazolium based ionic liquids. *Chem. Phys.* **2022**, *562*, 111648. [CrossRef]

52. Flórez, E.; Acelas, N.; Gómez, S.; Hadad, C.; Restrepo, A. To Be or Not To Be? that is the Entropic, Enthalpic, and Molecular Interaction Dilemma in the Formation of (Water) 20 Clusters and Methane Clathrate. *ChemPhysChem* **2022**, *23*, e202100716. [CrossRef]

53. Flórez, E.; Gómez, S.; Acelas, N.; Hadad, C.; Restrepo, A. Microsolvation versus encapsulation in mono, di, and trivalent cations. *ChemPhysChem* **2022**, e202200456. [CrossRef] [PubMed]

54. Lu, T.; Chen, F. Multiwfn: A multifunctional wavefunction analyzer. *J. Comput. Chem.* **2012**, *33*, 580–592. [CrossRef] [PubMed]

55. Glendening, E.D.; Badenhoop, J.K.; Reed, A.E.; Carpenter, J.E.; Bohmann, J.A.; Morales, C.M.; Landis, C.R.; Weinhold, F. *NBO 6.0*; Theoretical Chemistry Institute, University of Wisconsin: Madison, WI, USA, 2013.

56. Contreras-García, J.; Johnson, E.R.; Keinan, S.; Chaudret, R.; Piquemal, J.P.; Beratan, D.N.; Yang, W. NCIPLOT: A Program for Plotting Noncovalent Interaction Regions. *J. Chem. Theory Comput.* **2011**, *7*, 625–632. [CrossRef] [PubMed]

57. Herraez, A. Biomolecules in the computer: Jmol to the rescue. *Biochem. Mol. Biol. Educ.* **2006**, *34*, 255–261. [CrossRef] [PubMed]

58. Humphrey, W.; Dalke, A.; Schulten, K. VMD—Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [CrossRef]

59. Reed, A.E.; Weinhold, F. Natural bond orbital analysis of near-Hartree–Fock water dimer. *J. Phys. Chem.* **1983**, *78*, 4066–4073. [CrossRef]

60. Copeland, C.; Menon, O.; Majumdar, D.; Roszak, S.; Leszczynski, J. Understanding the influence of low-frequency vibrations on the hydrogen bonds of acetic acid and acetamide dimers. *Phys. Chem. Chem. Phys.* **2017**, *19*, 24866–24878. [CrossRef]

61. Kalescky, R.; Kraka, E.; Cremer, D. Accurate determination of the binding energy of the formic acid dimer: The importance of geometry relaxation. *J. Chem. Phys.* **2014**, *140*, 084315. [CrossRef] [PubMed]

62. Farfán, P.; Echeverri, A.; Diaz, E.; Tapia, J.D.; Gómez, S.; Restrepo, A. Dimers of formic acid: Structures, stability, and double proton transfer. *J. Chem. Phys.* **2017**, *147*, 044312. [CrossRef] [PubMed]

63. Murillo, J.; David, J.; Restrepo, A. Insights into the structure and stability of the carbonic acid dimer. *Phys. Chem. Chem. Phys.* **2010**, *12*, 10963–10970. [CrossRef] [PubMed]

64. Knop, O.; Rankin, K.N.; Boyd, R.J. Coming to Grips with N–H···N Bonds. 1. Distance Relationships and Electron Density at the Bond Critical Point. *J. Phys. Chem. A* **2001**, *105*, 6552–6566. [CrossRef]

65. Knop, O.; Rankin, K.N.; Boyd, R.J. Coming to Grips with N–H···N Bonds. 2. Homocorrelations between Parameters Deriving from the Electron Density at the Bond Critical Point. *J. Phys. Chem. A* **2003**, *107*, 272–284. [CrossRef]

66. Ramírez, F.; Hadad, C.; Guerra, D.; David, J.; Restrepo, A. Structural studies of the water pentamer. *Chem. Phys. Lett.* **2011**, *507*, 229–233. [CrossRef]

67. Cremer, D.; Kraka, E. A description of the chemical bond in terms of local properties of electron density and energy. *Croat. Chem. Acta* **1984**, *57*, 1259–1281.

68. Rozas, I.; Alkorta, I.; Elguero, J. Behavior of ylides containing N, O, and C atoms as hydrogen bond acceptors. *J. Am. Chem. Soc.* **2000**, *122*, 11154–11161. [CrossRef]

69. Jenkins, S.; Morrison, I. The chemical character of the intermolecular bonds of seven phases of ice as revealed by ab initio calculation of electron densities. *Chem. Phys. Lett.* **2000**, *317*, 97–102. [CrossRef]

70. Arnold, W.D.; Oldfield, E. The chemical nature of hydrogen bonding in proteins via NMR: J-couplings, chemical shifts, and AIM theory. *J. Am. Chem. Soc.* **2000**, *122*, 12835–12841. [CrossRef]

71. Gómez, S.A.; Rojas-Valencia, N.; Gómez, S.; Cappelli, C.; Restrepo, A. The Role of Spike Protein Mutations in the Infectious Power of SARS-COV-2 Variants: A Molecular Interaction Perspective. *ChemBioChem* **2022**, *23*, e202100393. [CrossRef]

72. Flórez, E.; Acelas, N.; Ramírez, F.; Hadad, C.; Restrepo, A. Microsolvation of F$^-$. *Phys. Chem. Chem. Phys.* **2018**, *20*, 8909–8916. [CrossRef]

73. Weinhold, F.; Klein, R.A. Anti-Electrostatic Hydrogen Bonds. *Angew. Chem. Int. Ed.* **2014**, *53*, 11214–11217. [CrossRef]