

Article

Probing Non-Covalent Interactions through Molecular Balances: A REG-IQA Study

Fabio Falcioni , Sophie Bennett, Pallas Stroer-Jarvis and Paul L. A. Popelier * 

Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, UK; fabio.falcioni@manchester.ac.uk (F.F.); s.i.bennett@soton.ac.uk (S.B.)

* Correspondence: pla@manchester.ac.uk; Tel.: +44-161-3064511

Abstract: The interaction energies of two series of molecular balances (1-X with X = H, Me, OMe, NMe₂ and 2-Y with Y = H, CN, NO₂, OMe, NMe₂) designed to probe carbonyl...carbonyl interactions were analysed at the B3LYP/6-311++G(d,p)-D3 level of theory using the energy partitioning method of Interacting Quantum Atoms/Fragments (IQA/IQF). The partitioned energies are analysed by the Relative Energy Gradient (REG) method, which calculates the correlation between these energies and the total energy of a system, thereby explaining the role atoms have in the energetic behaviour of the total system. The traditional “back-of-the-envelope” open and closed conformations of molecular balances do not correspond to those of the lowest energy. Hence, more care needs to be taken when considering which geometries to use for comparison with the experiment. The REG-IQA method shows that the 1-H and 1-OMe balances behave differently to the 1-Me and 1-NMe₂ balances because the latter show more prominent electrostatics between carbonyl groups and undergoes a larger dihedral rotation due to the bulkiness of the functional groups. For the 2-Y balance, REG-IQA shows the same behaviour across the series as the 1-H and 1-OMe balances. From an atomistic point of view, the formation of the closed conformer is favoured by polarisation and charge-transfer effects on the amide bond across all balances and is counterbalanced by a de-pyramidalisation of the amide nitrogen. Moreover, focusing on the oxygen of the amide carbonyl and the α -carbon of the remaining carbonyl group, electrostatics have a major role in the formation of the closed conformer, which goes against the well-known $n-\pi^*$ interaction orbital overlap concept. However, REG-IQF shows that exchange–correlation energies overtake electrostatics for all the 2-Y balances when working with fragments around the carbonyl groups, while they act on par with electrostatics for the 1-OMe and 1-NMe₂. REG-IQF also shows that exchange–correlation energies in the 2-Y balance are correlated to the inductive electron-donating and -withdrawing trends on aromatic groups. We demonstrate that methods such as REG-IQA/IQF can help with the fine-tuning of molecular balances prior to the experiment and that the energies that govern the probed interactions are highly dependent on the atoms and functional groups involved.

Keywords: Relative Energy Gradient (REG); molecular balances; quantum chemical topology (QCT); QTAIM; IQA; DFT; conformational analysis



Citation: Falcioni, F.; Bennett, S.; Stroer-Jarvis, P.; Popelier, P.L.A. Probing Non-Covalent Interactions through Molecular Balances: A REG-IQA Study. *Molecules* **2024**, *29*, 1043. <https://doi.org/10.3390/molecules29051043>

Academic Editors: Donald G. Truhlar, M. Natália D.S. Cordeiro, Chiara Cappelli, Antonio Fernández-Ramos and Bryan M. Wong

Received: 1 February 2024

Revised: 19 February 2024

Accepted: 21 February 2024

Published: 28 February 2024



Copyright: © 2024 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Non-covalent interactions are defined as weak interactions between atoms or molecules that do not involve the sharing of electrons. More precisely, the purely quantum mechanical exchange effect between electrons of non-covalently interacting constituents is practically not important. The best-known types of non-covalent interactions are hydrogen bonds and van der Waals contacts. Although comparatively weak compared to interactions of covalent (i.e., bonds) or ionic nature, many-body effects involved in non-covalent interactions can lead to subtle but complex phenomena in biological, organic and inorganic systems. Protein folding, enzyme–substrate recognition, host–guest binding, the formation of lipid bilayers,

solubility, boiling points and melting points are some of the phenomena governed by the presence of non-covalent interactions.

Essentially, non-covalent interactions form between sub-units (for example, atoms, functional groups or amino acids) of a larger system, and their cumulative effect becomes significant, even more than their covalent counterpart when considering the system as a whole [1–5]. While their definition is straightforward, the debate continues on how these interactions act, more specifically on how to tailor them in the engineering of enzymes and materials. One promising methodological route of growing popularity in the experimental community is the synthesis and engineering of molecular balances.

Molecular balances are carefully engineered organic molecules that present a unique conformational equilibrium. Specifically, they are designed to rotate about a single bond in such a way that, in theory, the atoms at one side of the main dihedral angle can either interact with the other side of the dihedral or simply interact with a solvent or a more distant part of the balance. The former state is usually defined as the *closed* conformer (CC), while the latter is defined as the *open* conformer (OC). Figure 1 shows these states.

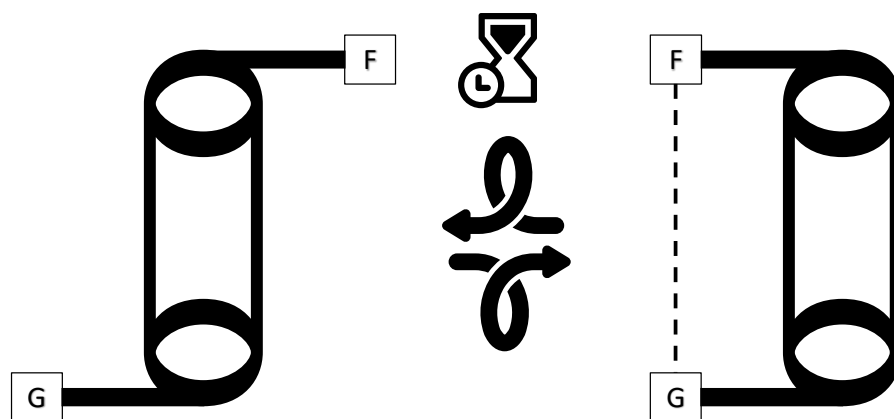


Figure 1. Schematic representation of the conformational equilibrium of a typical molecular balance. On the left, the molecular balance is in its *open* conformer state, with functional groups *F* and *G* not interacting. On the right, the balance ends up in its *closed* conformation state where functional groups *F* and *G* are interacting after a series of rotations marked by the curved arrows. The timescale, indicated by the hourglass, of the conformational equilibrium is slow enough such that two separate peaks are observed in an NMR spectrum.

The primary objective of a molecular balance is to isolate a non-covalent interaction such that it can be properly investigated. This objective is achieved by exploiting the slow conformational equilibrium between the open and closed conformers, using NMR techniques in a solvated environment. Through NMR, it is possible to obtain an energetic quantity, specifically a free energy variation, which is usually related to a type of non-covalent interaction.

In 1994, Wilcox et al. designed [6] the first molecular balances, which involved aryl–aryl interactions. Subsequently, other research groups either modified the Wilcox balance or engineered completely different molecules with shapes and degrees of freedom suitable for specific interactions. For example, Cozzi et al. developed [7,8] sterically strained aromatic systems in order to probe parallel and off-set stacking interactions. Shimizu et al. also investigated various aromatic interactions but then used N-arylamide balances with extended aromatic face-to-face surfaces [9–16]. Formamide balances with a simple structure were developed more recently by Cockroft et al. These balances are capable of probing conformational effects such as through-space substituent effects, solvent effects and carbonyl–carbonyl interactions [17–24]. A review [25] by Cockroft et al. and one [26] by Motherwell et al. both provide a deeper understanding of molecular balances.

Although the molecular balances literature was originally experimental, it has recently seen an increased number of articles connecting experiment and computation. Energy

decomposition methods are well-suited for understanding non-covalent interactions because they recover meaningful and intuitive energy contributions from the total energy of the wavefunction. Natural Bond Orbital analysis (NBO) and Symmetry-Adapted Perturbation Theory (SAPT) are two well-established methods that are generally used, as well as their variants [27–32]. However, these methods suffer from some shortcomings. For instance, SAPT introduces an unphysical reference state on which the analysis then depends. Secondly, the NBO paradigm suffers from the arbitrary and cumbersome nature of the procedure that builds the initial natural atomic orbitals on which the results then depend [33]. We, therefore, believe that partitioning in real space [34] paves the way for understanding the true role of atoms in the formation of both covalent and non-covalent interactions. We also believe that there is a need for atomistic detail in order to help the design of molecular balances. In fact, it is vital to conduct a thorough computational analysis of all potential non-covalent interactions in a selected molecular balance before quantifying the strength of a specific non-covalent interaction experimentally.

In this study, we employ the newly developed Relative Energy Gradient (REG) method [35] and couple it with the energy partitioning scheme coined Interacting Quantum Atoms (IQA) [36,37]. The latter is a generalisation of the Quantum Theory of Atoms in Molecules (QTAIM) [38]. IQA is perfectly suited both for understanding non-covalent interactions in molecular balances and guiding their design. This capacity is achieved by IQA's recovery of chemically intuitive energetic terms, which are related to familiar concepts such as sterics, covalency, (hyper)conjugation, ionisation and polarisation [39–43]. Indeed, REG-IQA has attracted a growing number of case studies ranging from small molecules to enzymatic active sites. This expanding body of work includes the following non-exhaustive list of case studies: the anomeric effect, the halogen bond, the two torsional energy barriers in biphenyl, the fluorine *gauche* effect, cycloaddition reactions, XH- π interactions, as well as reactions in larger systems such as enzymes. Moreover, REG has been recently shown to work with multipolar electrostatics and with dispersion corrections [44,45].

We base much of the current work on a recent paper [19] by Cockroft et al. In that study, the authors synthesised various types of molecular balances in order to probe carbonyl...carbonyl interactions. They strove to reconcile [46] the long debate on the roots of the so-called $n\text{-}\pi^*$ interaction discovered by Raines et al. [47–52]. Two opposing views exist on the nature of carbonyl–carbonyl interactions. Some researchers argue that they are orbital–orbital interactions, in which the lone pair electrons of the donor carbonyl's oxygen interact with the antibonding π^* orbital of the acceptor carbonyl. Others argue [53,54] that they are primarily electrostatic interactions.

We analyse two types of molecular balances with different functional groups, as shown in Figure 2. The first balance corresponds to the recently engineered formamide balance, while the second one is the *N*-formylproline balance. The first type of balance will be referred to as the 1-*X* balance, where *X* refers to any of the four functional groups studied (Figure 2, top). The second type of balance will be referred to as the 2-*Y* balance, where *Y* refers to any of the five investigated functional groups (Figure 2, bottom).

We also benefit from the well-curated data in the work [19] of Cockroft and co-workers, such as the $\alpha\text{-}\beta$ Hunter solvation model energies and the X-ray structures of the formamide balances, for comparison and enhancement of our discussion [55]. The application of REG-IQA to these balances has two main objectives: (i) analysis of the role of all the pairwise interactions in the conformational equilibrium of molecular balances and (ii) giving a different perspective to the orbital-based definition of $n\text{-}\pi^*$. The latter is a consequence of the orbital-invariant nature of the IQA approach. Moreover, we employ REG-IQF to analyse the interaction between fragments of the molecular balances, showing that the combination of information at both atom and fragment levels reveals the origin behind their interactions.

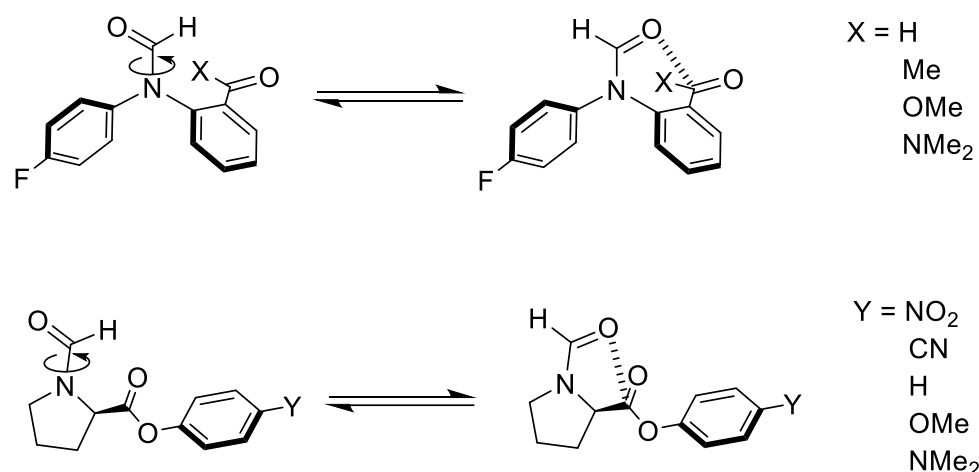


Figure 2. (top) Formamide balances by Cockroft et al. [19] and (bottom) *N*-formylproline balances by Raines et al. [50]. The open conformers for both balances are shown on the left-hand side of the equilibrium arrows, while the closed conformers are on the right-hand side.

2. Results and Discussion

This section consists of two main parts. The first part regards the conformational landscape of the 1-X balances, where we show that the back-of-the-envelope open and closed conformers are not the true lowest energy minima and that the conformational landscape of these balances is much more complicated than it seems. Back-of-the-envelope open and closed conformers arise from the rotation of only a single flexible bond, particularly the C-N bonds involving the 1-X and 2-Y balances. The second part focuses on the REG-IQA and REG-IQF analysis of the 1-X and 2-Y series of balances.

2.1. Conformational Landscape of 1-X

We first perform calculations to understand how experiments can be linked to computation. In the literature on molecular balances, an experimental free energy barrier is usually related to two geometries: the open and closed conformers, as depicted in Figure 2. This free energy barrier is obtained from two NMR chemical shifts, one for the open and one for the closed conformer, through $\Delta G_{exp} = -RT \ln K = -RT \ln [\text{closed}] / [\text{open}]$. However, NMR is performed in a solvated environment. Hence, we believe that the back-of-the-envelope conformational exchange that is usually shown does not correspond to the real conformational exchange occurring in solution. This is why we perform a series of optimisations on the 1-X balances, starting from the relaxed open and closed conformers, and manually search for other possible conformations and local energy minima. Specifically, we compare the variation of energies between the lowest energy closed conformer (LECC) and various conformations of the open conformer. We note that manual search is simple in this case as the number of rotatable bonds is very low compared to that found in enzymes where this kind of non-covalent interaction occurs. We freeze the main C-N amide bond, either in the open or closed state conformers, and then start rotating the other dihedrals. We consider four rotatable bonds for the main backbone of the balance in the 1-X balances in our analysis. The rotatable bonds correspond to the C-N amide bond, the other two C-N bonds, each involving one aromatic group, and fourthly, the C-C bond between the XC=O fragment (containing the functional group) and the aromatic ring bonded to it. We do not explicitly consider the rotatable bonds within the X functional group (specifically in the 1-Me, 1-OMe and 1-NMe₂ balances), but we let those relax with the geometry optimisation. The optimised structures of the local energy minima found are shown in Section S1 of the Supplementary Materials.

Figure 3 shows the conformations of the LECC for the 1-X molecular balances. In this case, the LECC correspond to the back-of-the-envelope conformation (BOTEC) shown in

Figure 2 across all the 1-X balances. For the 1-H and 1-Me balances, the X-ray geometry is also overlayed with the corresponding LECC. This comparison shows that the DFT calculation is in good agreement with the experiment geometrically.

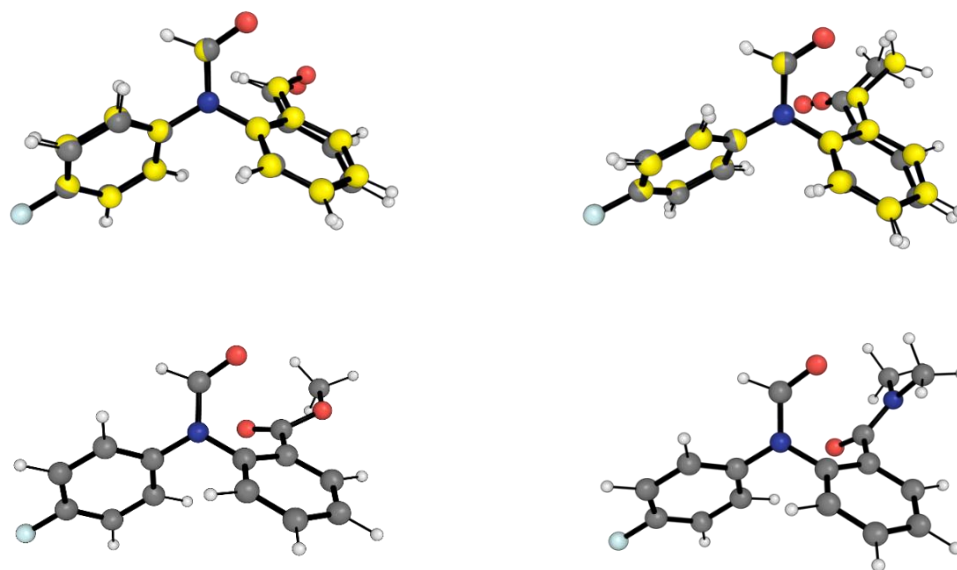


Figure 3. Optimised geometries of closed conformers for the four 1-X balances. Gray carbon geometries correspond to the B3LYP/6-311++G(d,p) local energy minimum structures for the BOTECs, which also correspond to the LECC geometries. Yellow carbon geometries correspond to the X-ray structures (only for 1-H and 1-Me), as obtained from the Cambridge Crystallographic Database as in reference [19]. The red, blue, white and pale blue balls are respectively oxygens, nitrogens, hydrogens and fluorines. The optimised geometries and the energies for the 1-X balances are also shown in Figures S1–S4.

Figure 4 shows the conformation for the lowest energy open conformers (LEOC) for the 1-X molecular balances overlayed with the BOTEC that is expected from common intuition, as presented in Figure 2. Here, we see clear geometrical differences between the BOTEC and the LEOC. This corroborates our hypothesis that the BOTEC, although being a local energy minimum, is not the global minimum for any of the 1-X balances. From a simple geometrical and visual analysis, it is more favourable for the X carbonyl group to rotate away from the amide carbonyl group.

Cockroft et al. employed Hunter solvation models in various studies [17,19,21]. Briefly, this methodology consists of an iterative fit to experimentally obtain ΔG values in different solvents. This gives an approximated experimental energy difference, denoted in their work as ΔE_{exp} , between the closed and open conformers of a specific molecular balance. While this is useful to assess experimental free energy barriers and the dependency of the molecular balance conformational equilibrium across different solvents, it can also be used to benchmark energy barriers for computational techniques.

Table 1 shows the energy differences of open and closed conformers that can be related to the NMR conformational equilibrium through ΔE_{exp} . Specifically, the correlation between ΔE_{exp} and $\Delta E_{\text{LECC-BOTEC}}$ is only $R^2 = 0.78$ compared to $R^2 = 0.84$ between ΔE_{exp} and $\Delta E_{\text{LECC-LEOC}}$. Hence, the conformational equilibrium occurs between the LECC and the LEOC. Secondly, the BOTEC is not the geometry on which further analysis should focus. The agreement with the experiment is also demonstrated by computing the root mean squared error (RMSE) between the experimental energy difference and the computed energy differences, which results in a clearly better agreement for the $\Delta E_{\text{LECC-LEOC}}$.

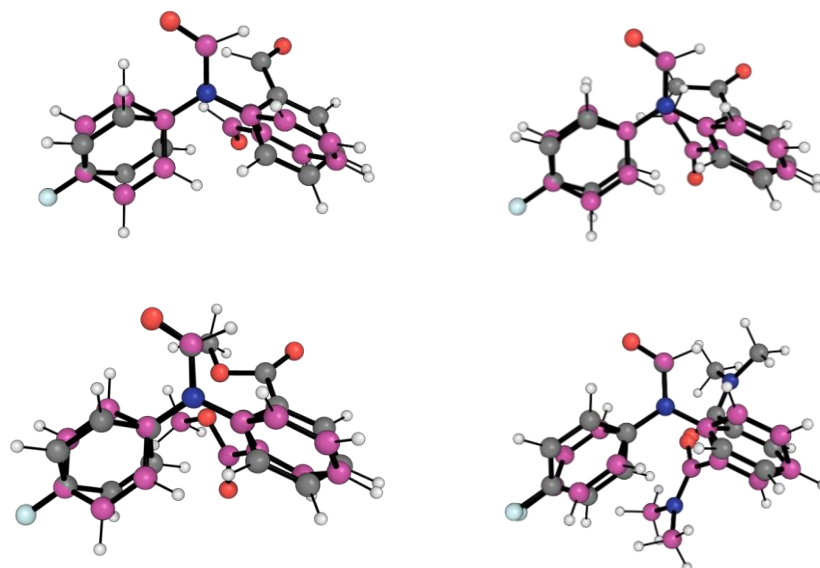


Figure 4. Optimised geometries of open conformers for the four 1-X balances. Gray carbon geometries correspond to the B3LYP/6-311++G(d,p) + D3 and Becke–Johnson damping local energy minima structures for the BOTECs. Magenta carbon geometries correspond to the B3LYP/6-311++G(d,p) + D3 and Becke–Johnson damping local energy minima structures for the LEOC. The red, blue, white and pale blue balls are respectively oxygens, nitrogens, hydrogens and fluorines. The optimised geometries and the energies for the 1-X balances are also shown in Figures S1–S4.

Table 1. Energy differences for the 1-X balances. ΔE_{exp} is the energy variation for the conformational equilibrium obtained with the Hunter solvation model fitted through NMR free energy differences obtained in different solvents, obtained from Ref. [19] The quantity $\Delta E_{\text{LECC-BOTEC}}$ is the energy difference between the LECC and the “open” BOTEC computed at B3LYP/6-311++G(d,p) level of theory, with D3 dispersion correction and Becke–Johnson damping. $\Delta E_{\text{LECC-LEOC}}$ is the energy difference between the LECC and the LEOC (again at B3LYP/6-311++G(d,p)). Geometries and their corresponding energies are shown in Figures S1–S4.

Balance	ΔE_{exp}	$\Delta E_{\text{LECC-BOTEC}}$	$\Delta E_{\text{LECC-LEOC}}$
1-H	3.1	0.3	1.2
1-Me	−4.5	−6.1	−5.7
1-OMe	−0.8	−1.4	0.5
1-NMe ₂	−5.3	−12.6	−4.4
RMSE (kJ/mol)		4.0	1.4
R ²		0.78	0.84

Note that we do not perform the same analysis for the 2-Y balance because experimental data to perform a Hunter solvation analysis are not available to our knowledge.

To understand the forces that cause a carbonyl...carbonyl interaction, we analyse the PES of a series of molecular balances with different electron-donating and electron-withdrawing groups. We start at the LECC and move towards the optimised transition state (TS) for each molecule. We are aware that the PES between the LECC and the LEOC is not a simple one-step conformational equilibrium and that analysing each possible 1D PES corresponding to one of the four dihedral rotations analysed would lead to an overload of information. Hence, we focus only on the energy barrier leading directly to the closed conformer, where the carbonyl...carbonyl interaction is mostly probed.

To make sure that this PES is properly modelled, we perform a relaxed scan where, at each step, three to four dihedrals of the 1-X balance and three dihedrals of the 2-Y balance are rotated over a specified number of degrees. Thus, the control coordinate of the REG-IQA analysis is a collection of dihedral rotations and concomitant geometry relaxations, that

is, a collective variable. In other words, we are projecting the higher dimensional space in which all four dihedral rotations belong to a 1D PES space. Section S2 of the Supplementary Materials provides more detail on the PES studied for each balance.

2.2. REG-IQA Analysis

The first step of a typical REG-IQA analysis is to assess which are the most relevant interactions for the energy barrier under study. Thanks to the IQA partitioning, we can also connect its energy terms to meaningful chemical contributions such as covalency (i.e., making and breaking of bonds) for V_{xc} terms, electrostatics/ionisation/polarisation for V_{cl} , and sterics for E_{intra} . The interactions that correspond to the highest positive REG values constructively work with the PES, while the opposite holds for the highest negative values. In the case of the PES from TS to LECC, positive values correspond to interactions that help the formation of the LECC, while negative values go against it.

We take the 1-H balance as an example of a fully atomistic REG-IQA analysis and then explore the similarities across the other 1-X balances. Table 2 shows the REG-IQA analysis of the 1-X balance, while the corresponding atomic labels are shown in Figure 5. The most relevant interaction that helps the formation of the LECC is the electrostatic $V_{cl}(N1,C4)$ term, which appears as the most positive REG value for all four balances. The behaviour of this energy contribution is related to the amide bond shortening in going from the TS to the LECC. Specifically, there is a strong charge-transfer effect [$REG-V_{ct}(N1,C4) = 2.7$] and a noticeable polarisation [$REG-V_{pl}(N1,C4) = 2.0$] between the amidic C and N (Table S3 of the Supplementary Materials). Furthermore, the nitrogen atom transitions from a stable trigonal pyramidal in the TS to a locally flat geometry in the LECC, which explains the strong steric effect [$REG-E_{intra}(N1) = -2.3$] that is going against the formation of the LECC. Similar conclusions have been obtained for N-substituted aziridines [56]. Note that these findings are also in agreement with more established work by Kemnitz and Loewen on C–N rotation barriers [57]. All this shows that the rotation along the amide bond is essential for forming the LECC. Surprisingly, it turns out that the other carbonyl group in the molecule is less important for determining the energy barrier than expected based on chemical intuition. Indeed, the first two relevant interactions that involve the two carbonyl groups have small REG values: the positive $REG-V_{cl}(O5,C22) = 0.7$ and the negative $REG-V_{cl}(C4,C22) = -0.7$. These energy contributions are of an electrostatic nature, and they counterbalance each other. All other most positive or most negative REG values are either related to 1–2 bond interactions or intra-atomic steric effects.

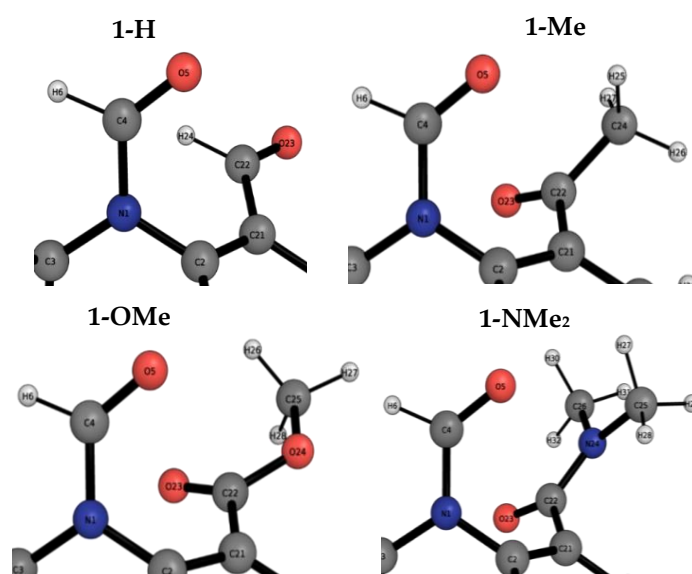


Figure 5. The carbonyl...carbonyl moiety in the closed conformer of the 1-X molecular balance series. Labels of atoms are shown to aid the REG-IQA analysis.

Table 2. REG-IQA tables for the 10 (5 positive and 5 negative) most important REG values for all the 1-X series of balances. The REG-IQA analysis is run on the PES going from the LECC to a TS structure. Positive REG values help the formation of the LECC, while negative values work against it.

1-H			1-Me		
TERM	REG	R	TERM	REG	R
Eintra(n1)	−2.3	−0.96	V _{cl} (c4,c22)	−2.5	−0.98
V _{cl} (n1,o5)	−1.4	−0.99	V _{cl} (n1,o23)	−2.5	−1.00
Eintra(c4)	−1.1	−0.98	V _{cl} (o5,o23)	−2.3	−0.98
V _{xc} (c4,o5)	−0.8	−0.98	Eintra(n1)	−2.0	−1.00
V _{cl} (c4,c22)	−0.7	−0.92	V _{cl} (n1,o5)	−1.2	−0.97
Eintra(o5)	0.5	0.97	V _{cl} (n1,c2)	1	0.99
V _{cl} (o5,c22)	0.7	0.82	V _{cl} (n1,c22)	1.8	1
V _{cl} (n1,c2)	0.8	0.88	V _{cl} (o5,c22)	2.1	0.97
V _{xc} (n1,c4)	1	0.96	V _{cl} (c4,o23)	3	0.99
V _{cl} (n1,c4)	4.7	0.99	V _{cl} (n1,c4)	3.6	0.97

1-OMe			1-NMe ₂		
TERM	REG	R	TERM	REG	R
Eintra(n1)	−2.1	−0.98	V _{cl} (c4,c22)	−3.2	−0.99
V _{cl} (c4,c22)	−1.2	−0.98	V _{cl} (o5,n24)	−2.6	−0.99
V _{cl} (n1,o5)	−1.2	−0.99	V _{cl} (o5,o23)	−2.5	−0.99
Eintra(c4)	−1.1	−0.99	V _{cl} (n1,o23)	−1.9	−0.98
V _{cl} (o5,o23)	−1.0	−0.97	Eintra(n1)	−1.6	−0.95
V _{xc} (n1,c4)	0.7	0.95	V _{cl} (h13,o23)	1	0.91
V _{cl} (n1,c2)	0.8	0.92	V _{cl} (c4,n24)	1.7	0.98
V _{cl} (o5,c22)	1	0.89	V _{cl} (c4,o23)	3	0.99
V _{cl} (c4,o23)	1.5	0.99	V _{cl} (o5,c22)	3.3	0.99
V _{cl} (n1,c4)	4.1	1	V _{cl} (n1,c4)	3.9	0.97

We can deduce the same REG-IQA analysis for the 1-OMe system, but the 1-Me and 1-NMe₂ systems behave differently. For these two systems, electrostatics play a major role in forming the LECC. In both systems, the V_{cl}(N1,C4) interaction is still the most important, but interactions involving the two carbonyl groups, C4=O5 and C22=O23, become more significant. For example, the V_{cl}(O5,C22) and V_{cl}(C4,O23) interactions have strong positive REG values, while V_{cl}(C4,C22) and V_{cl}(O5,O23) interactions have important negative REG values, which is not the case for the 1-H and 1-OMe balances where steric effects (E_{intra} of C4 and N1) are predominant compared to electrostatics. We note that V_{cl}(N1,O23) also shows a significant negative REG value, highlighting the direct involvement of the amidic N atom with the opposite carbonyl group. The atomistic REG-IQA analysis effectively differentiates the 1-H/1-OMe set and the 1-Me/1-NMe₂ Set. The reason for this difference between the 1-H/1-OMe set and the 1-Me/1-NMe₂ set can be related to the ‘bulkiness’ of the functional groups. Indeed, in the 1-Me/1-NMe₂ set, the carbonyl connected to the X functional group undergoes a rotation (around CC, going from the LECC to the TS) with respect to the aromatic ring of, respectively, 112.7° (=131.3–18.6°, see Figure S6 (Me)) and 62.7° (=122.7–60.0°, see Figure S8 (NMe₂)), while little to no rotation, respectively, is observed for 1-OMe (14.3° = 147.8–133.5°, see Figure S7) and 1-H (where the CC angle does not exist). Note that the 1-NMe₂ also experiences a very large rotation of both the N-C bond involving the aromatic ring bearing the carbonyl-X group (80.5°) and the N-C bond with the aromatic ring bearing the fluorine atom (85.5°). These rotations are depicted in Figures S5–S8. Overall, the above similarities observed between the 1-Me/1-NMe₂ and 1-H/1-OMe are in strong agreement with the experimental energy variations shown in Table 1, where the ΔE_{LECC-LEOC} are negative for the 1-Me/1-NMe₂ and positive for 1-H/1-OMe. Thus, the 1-Me/1-NMe₂ are better molecular balances at probing the carbonyl–carbonyl interactions than the 1-H/1-OMe. In a way, we expect that the bulkier the X group is, the better the balance will be at probing a carbonyl...carbonyl interaction.

For the 2-Y series of molecular balances, we observe an almost identical REG-IQA analysis across all the functional groups. This could mean that this side of the PES is very similar across balances and that the Y functional group has no direct effect on this barrier. Figure 6 depicts the 2-Y balances functional groups and labels, while Table 3 shows the corresponding REG-IQA analyses for the most relevant interactions.

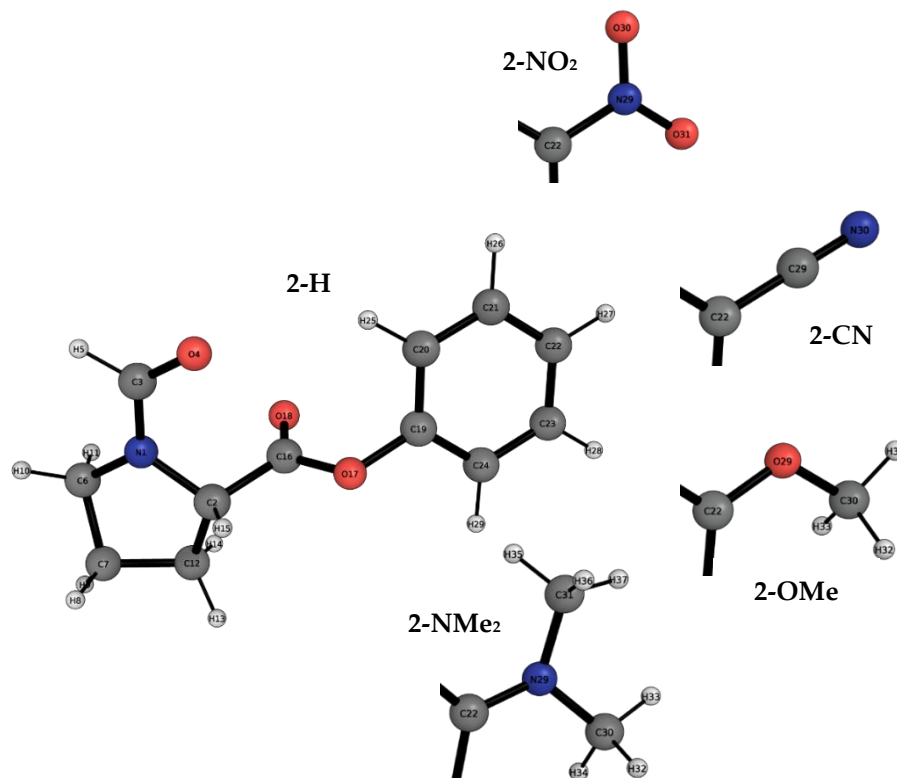


Figure 6. The atomic labelling of the whole 2-H molecular balance in the closed conformer. This labelling is common to all the balances of the 2-Y series except for the variation in any of the four remaining functional groups. Hence, only the changing functional groups are shown for the 2-NO₂, 2-CN, 2-OMe and 2-NMe₂ balances. Atomic labelling is shown to aid the REG-IQA analysis.

Here, the overall outcome of the REG-IQA analysis makes the amidic C-N interaction, $V_{cl}(N1,C3)$, the most relevant interaction for the formation of the LECC. For each of the five balances, this electrostatic interaction is partially counterbalanced by the unfavourable sterics of the N atom, which also undergoes a de-pyramidalisation (i.e., flattening) in going from the TS structure to the LECC. The rest of the most relevant REG values ranking displays a very similar trend to the 1-X balances where some electrostatic interactions between the two carbonyl groups [$V_{cl}(O4,C16)$] work with and against [$V_{cl}(C3,C16)$] the formation of the LECC. Surprisingly, the contributions of the oxygen (O17) attached to the aromatic ring are more important than the oxygen (O18) of the carbonyl group (Table S4). Taking the 2-H balance as an example, the $REG-V_{cl}(O4,C17) = -1.6$, while $REG-V_{cl}(O4,C18) = -1.3$, and $REG-V_{cl}(O3,C17) = 1.3$, while $REG-V_{cl}(O3,C18) = 0.9$. While these changes are subtle, they suggest an intuition for what is effectively probed by the conformational equilibrium of the 2-Y balances. In this case, all the molecular balances of the 2-Y series show similar behaviour to the 1-OMe and 1-H balance, which we earlier concluded to be less efficient at probing the carbonyl...carbonyl interaction.

From the fully atomistic analysis of both 1-X and 2-Y systems, we can determine that the formation of the LECC, specifically the last energy barrier of the overall energy landscape (which leads to its formation), is mainly driven by electrostatic effects. These effects do involve atoms in the carbonyl...carbonyl moiety, but these atoms are not the most important factors in determining the height of the energy barrier.

Table 3. REG-IQA tables for the 10 (5 positive and 5 negative) most important REG values for all the 2-Y series of balances. The REG-IQA analysis is run on the PES going from the LECC to a transition-state-like structure. Positive REG values help the formation of the LECC, while negative values work against it.

2-H					
TERM	REG	R			
Eintra(n1)	−4.1	−0.99			
V _{cl} (c3,c16)	−2.2	−0.98			
Eintra(c3)	−2.2	−1.00			
V _{cl} (n1,o4)	−1.9	−0.99			
V _{cl} (o4,o17)	−1.6	−1.00			
V _{cl} (c3,o17)	1.3	1			
V _{cl} (n1,c6)	1.4	0.97			
V _{cl} (n1,c16)	1.8	0.99			
V _{cl} (o4,c16)	2.5	1			
V _{cl} (n1,c3)	6.9	0.99			
2-NO ₂			2-CN		
TERM	REG	R	TERM	REG	R
Eintra(n1)	−3.8	−0.99	Eintra(n1)	−3.8	−0.98
V _{cl} (c3,c16)	−2.3	−0.98	V _{cl} (c3,c16)	−2.3	−0.99
Eintra(c3)	−2.0	−1.00	Eintra(c3)	−2.0	−1.00
V _{cl} (n1,o4)	−1.8	−0.99	V _{cl} (n1,o4)	−1.8	−0.99
V _{cl} (o4,o17)	−1.7	−1.00	V _{cl} (o4,o17)	−1.6	−1.00
V _{cl} (n1,c6)	1.3	0.97	V _{cl} (c3,o17)	1.2	1
V _{cl} (c3,o17)	1.3	1	V _{cl} (n1,c6)	1.4	0.97
V _{cl} (n1,c16)	1.7	0.99	V _{cl} (n1,c16)	1.7	0.99
V _{cl} (o4,c16)	2.7	1	V _{cl} (o4,c16)	2.6	1
V _{cl} (n1,c3)	6.4	0.99	V _{cl} (n1,c3)	6.4	0.99
2-OMe			2-NMe ₂		
TERM	REG	R	TERM	REG	R
Eintra(n1)	−4.1	−0.99	Eintra(n1)	−4.2	−0.99
V _{cl} (c3,c16)	−2.3	−0.98	Eintra(c3)	−2.3	−1.00
Eintra(c3)	−2.2	−1.00	V _{cl} (c3,c16)	−2.1	−0.98
V _{cl} (n1,o4)	−1.9	−0.99	V _{cl} (n1,o4)	−1.9	−0.99
V _{cl} (o4,o17)	−1.6	−1.00	V _{cl} (o4,o17)	−1.5	−1.00
V _{cl} (c3,o17)	1.3	1	V _{cl} (n1,c2)	1.3	0.97
V _{cl} (n1,c6)	1.4	0.97	V _{cl} (n1,c6)	1.5	0.97
V _{cl} (n1,c16)	1.8	0.99	V _{cl} (n1,c16)	1.9	0.99
V _{cl} (o4,c16)	2.6	1	V _{cl} (o4,c16)	2.4	0.99
V _{cl} (n1,c3)	6.8	0.99	V _{cl} (n1,c3)	7.1	0.99

An interesting aspect of analysing molecular balances with similar backbones and different functional groups is to understand the effect of electron-donating versus electron-withdrawing groups on the formation of the LECC. Earlier, we observed a different behaviour between 1-H/1-OMe groups and 1-Me/1-NMe₂ groups. However, this observation is not related to the functional groups' electronegativity behaviour but more to their size.

To uncover more subtle effects, it is necessary to investigate the patterns of both groups of interactions or specific interactions within the systems. A direct comparison of REG values across different systems is not straightforward because their values are linearly dependent on the total energy of the respective system. Instead, it is more informative to analyse the *rank* of a REG value, which represents its position in the REG table when sorted from most positive to most negative values. A higher rank indicates a more significant interaction, as previously demonstrated. If an interaction's rank varies *across* different systems, it implies that its impact on the overall energy barrier is either weaker or stronger in comparison to other systems.

One possible way to examine the rank of a REG value is to simply calculate its ratio with the overall highest REG value (positive or negative). Specifically, the highest positive REG-IQA (REG_{max}) value is divided by every REG value corresponding to an i -th energy term (REG_i) as $REG\ Ratio_i = \frac{REG_{max}}{REG_i}$. This is performed for every i -th energy term within each REG-IQA analysis (i.e., for every system). This measure quantifies how many times an interaction is less or more relevant than the overall most relevant interaction for the energy barrier. Let us, for example, compare two REG Ratios: if one REG Ratio exceeds the other, then the interaction with the higher REG Ratio influences the energy barrier the least. To give an example of this assertion, we imagine systems *A* and *B*. If in system *A*, the i -th energy term has a REG Ratio of 3, then this term is 3 times less effective than the strongest interaction. If in system *B*, the same i -th energy term has a ratio of 4, then this term is 4 times less effective than the strongest interaction. Thus, in system *B*, the same interaction is less effective overall than it is in system *A*.

In contrast, if the REG Ratio is lower, then the interaction will influence the energy barrier more. A REG Ratio close to one simply means that the interaction is as prominent as the strongest interaction for that barrier. A negative REG Ratio shows that the interaction is working against the barrier with a certain amount of intensity with respect to the strongest positive interaction.

We start by analysing the V_{xc} and V_{cl} between two fragments. For example, one fragment (*C*) covers the three atoms in the amide–carbonyl group, while the other fragment (*D*) covers the three atoms of the second carbonyl group plus the *X* or *Y* functional groups of the respective balances. These two groups of atoms considered are depicted in Figure 7 as *E* and *F*, alongside other groups of atoms, including single atoms (*A* and *B*). These groups were chosen arbitrarily, although based on the nearby environment of the carbonyl groups. Given that the purpose of these synthesised molecular balances is to probe carbonyl...carbonyl interactions, we want to explore the influence of the atoms around these carbonyl groups. Results of this analysis are shown in Table 4 for the 1-*X* series of balances and in Table 5 for the 2-*Y* series.

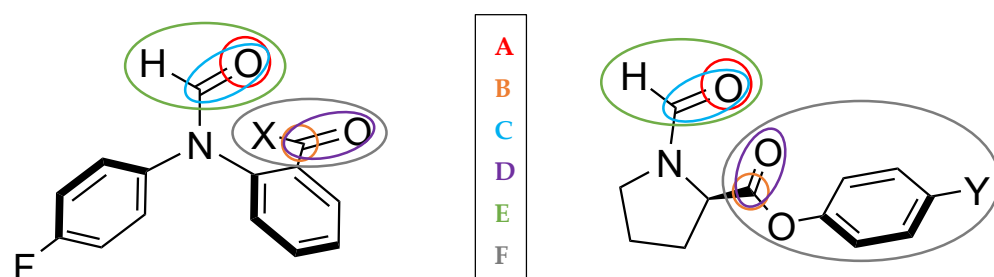


Figure 7. The 1-*X* and 2-*Y* balances in the closed conformer state. Groups of atoms for the REG-IQF analysis are highlighted with different colours.

For the 1-*H*, 1-*OMe* and 1-*NMe₂* balances, V_{xc} increases in relevance (i.e., REG Ratios go from higher to lower) in going from the traditional $V_{xc}(A,B)$ interaction to considering the whole carbonyl...carbonyl-*X* groups [$V_{xc}(E,F)$]. Note that the former corresponds to the $V_{xc}(O5,C22)$ interaction associated with the $n-\pi^*$ interaction, which sees the oxygen of one carbonyl group interacting with the carbon of the other carbonyl group. Hence, the carbonyl...carbonyl interaction is probed by all of the atoms near the carbonyl groups, as expected. However, the opposite occurs for the 1-*Me* balance, where consideration of the *Me* group in the $V_{xc}(E,F)$ term decreases the relevance of the interaction. This could mean that the *Me* group hyperconjugation effect is actually working against the formation of the carbonyl...carbonyl interaction. Looking next at the electrostatic V_{cl} interactions, the opposite occurs to what was seen for V_{xc} . $V_{cl}(A,B)$, which corresponds to the $V_{cl}(O5,C22)$ interaction, makes a very important contribution to the overall formation of the LECC. However, when considering the interactions going from the (*A*,*B*) group to the (*E*,*F*) groups, where the *X* functional group is considered, the relevance of electrostatics is attenuated in

all the balances. In other words, going top to bottom in Tables 4 and 5 (for each balance), which corresponds to adding more atoms to the groups and specifically X, electrostatics are less and less relevant. Specifically, the 1-H balance goes from a REG Ratio of 7 to a REG Ratio of 220, meaning that overall electrostatics for the formation of the LECC are not relevant for this balance.

Table 4. Table of the V_{xc} and V_{cl} REG-IQF and REG Ratios for the 1-X series of balances. REG-IQF values are obtained by summing over all the interatomic energies between the atoms of one group and the atoms of the other group. The REG Ratios are obtained as a proportion between a REG-IQF value between two groups and the highest positive REG-IQA value of the fully atomistic REG-IQA analysis.

Term	1-H		1-Me		1-OMe		1-NMe ₂	
	REG	REG Ratio	REG	REG Ratio	REG	REG Ratio	REG	REG Ratio
$V_{xc}(A,B)$	0.05	99	0.21	17	0.06	66	0.05	82
$V_{xc}(C,D)$	0.07	64	0.33	11	0.19	21	0.10	37
$V_{xc}(E,F)$	0.10	46	0.06	59	0.18	23	0.27	14
$V_{cl}(A,B)$	0.67	7	2.07	2	1.05	4	3.28	1
$V_{cl}(C,D)$	−0.05	−91	0.28	13	0.28	15	0.55	7
$V_{cl}(E,F)$	0.02	220	0.30	12	0.21	20	0.32	12

Table 5. Table of the V_{xc} and V_{cl} REG-IQF and REG Ratios for the 2-Y series of balances. REG-IQF values are obtained by summing over all the interatomic energies between the atoms of one group and the atoms of the other group. REG Ratios are obtained as a proportion between a REG-IQF value between two groups and the highest positive REG-IQA value of the fully atomistic REG-IQA analysis.

Term	2-NO ₂		2-CN		2-H		2-OMe		2-NMe ₂	
	REG	REG Ratio	REG	REG Ratio	REG	REG Ratio	REG	REG Ratio	REG	REG Ratio
$V_{xc}(A,B)$	0.22	30	0.21	31	0.19	36	0.18	37	0.17	42
$V_{xc}(C,D)$	0.19	34	0.18	35	0.12	55	0.13	53	0.09	83
$V_{xc}(E,F)$	0.31	21	0.27	24	0.22	31	0.17	41	0.09	76
$V_{cl}(A,B)$	2.67	2	2.64	2	2.55	3	2.57	3	2.42	3
$V_{cl}(C,D)$	0.00	−2256	0.03	226	−0.08	−87	−0.08	−86	−0.11	−62
$V_{cl}(E,F)$	−0.12	−54	−0.10	−62	−0.07	−105	−0.18	−38	−0.14	−50

We see a similar trend for the 1-OMe and 1-NMe₂ balances, but notably, the electrostatic REG Ratios are on par with the exchange–correlation terms in this case. The REG Ratio for $V_{xc}(E,F)$ is very similar to the REG Ratio for $V_{cl}(E,F)$, meaning that both electrostatics and exchange–correlation have a similar contribution to the energy barrier. Finally, a different trend is observed for the 1-Me balance, where the $V_{xc}(E,F)$ REG Ratio is much higher than that for $V_{cl}(E,F)$, demonstrating that electrostatics are the most contributing energy component in this balance. We also note that the V_{cl} interactions between (the intermedially sized) groups (C,D), which correspond to the carbonyl groups C=O atoms, have similar values to the larger (E,F) groups. However, the opposite is true for the V_{xc} interactions, where (C,D) display values closer to the atomic (A,B) group.

In the 2-Y balances, the Y functional groups are not directly bonded with the carbonyl group, but they act on it through an aromatic ring. Thus, we already expect slightly different behaviour compared to the 1-X balances. For the V_{xc} terms, there is one relevant trend for all the A, B, C, D, E and F groups of atoms. The REG Ratios follow a trend similar to the intuition offered by organic chemistry for inductive effects on aromatic groups, where more electron-withdrawing groups activate the aromatic ring, and electron-donating groups deactivate it [58]. Indeed, when going from the most electron-withdrawing Y=NO₂ group to the most electron-donating Y=NMe₂ group, the REG Ratio of the carbonyl...carbonyl inter-

action for all the groups of atoms considered goes from a lower value (i.e., more prominent interaction) to a higher value (i.e., less prominent interaction). This shows that REG-IQA values fully recover chemical intuition and that through-space exchange–correlation energies are key for the carbonyl...carbonyl interaction. In contrast, electrostatics are the most relevant across all the 2-Y balances if considering only the $V_{cl}(A,B)$ interaction, which shows low positive REG Ratios but shifts to negative REG Ratios for $V_{cl}(C,D)$ and $V_{cl}(E,F)$, which involve larger groups. The sign change of REG Ratios essentially means that electrostatics are working against the formation of the LECC.

We now point out a subtle but very important outcome of the overall analysis. For the carbonyl...carbonyl interaction in both balances, we argue that electrostatics are much more relevant than exchange–correlation effects when we consider the oxygen of one amide carbonyl interacting with the carbon of the other carbonyl (i.e., the *A* and *B* groups). This assertion goes against the idea that carbonyl...carbonyl interactions are actually $n-\pi^*$ interactions (i.e., electron exchange between orbitals). However, when considering the effect of involving more atoms, such as in the interaction of *C*, *D*, *E* and *F* groups, the picture shifts for most of the balances. Indeed, we observe that (i) the electrostatics work on a par with exchange–correlation for some functional groups, for example, 1-OMe and 1-NMe₂, or that (ii) simply exchange–correlation takes the full lead, for example, in the whole 2-Y series.

In this analysis, we recognise the continuous presence of both $n-\pi^*$ interactions and electrostatic interactions in these types of systems. However, it is an oversimplification to attribute the formation of carbonyl–carbonyl interactions solely to these forces. In these molecular structures, each contributes in varying degrees, either stabilising or destabilising the overall system. This assertion holds true in biological systems as well, which are much more complex. By combining DFT with the detailed atomic insights provided by REG-IQA, we have developed a method to precisely pinpoint the most influential atoms or groups of atoms within these interactions. This refined approach allows for targeted modifications to the molecules, enhancing our understanding and manipulation of these interactions before embarking on the experiment.

3. Methods

3.1. Interacting Quantum Atoms (IQA)

The IQA energy decomposition is a real-space scheme that, in principle, employs the one-particle and two-particle [59] density matrices partitioned over quantum topological atoms. The total energy of a system (E_{tot}) with *N* atoms is obtained as a sum of all the intra-atomic and interatomic contributions of an atom *A* as

$$E_{tot} = \sum_A^N \left(E_{intra}^A + \frac{1}{2} \sum_{B \neq A}^N V_{inter}^{AB} \right) \quad (1)$$

For each atom *A*, there is an intra-atomic energy (E_{intra}) defined as

$$E_{intra}^A = T^A + V_{ee}^{AA} + V_{en}^{AA} \quad (2)$$

where T^A , V_{ee}^{AA} and V_{en}^{AA} correspond, respectively, to the kinetic energy, the electron–electron repulsion, and the electron–nucleus attraction energies for atom *A*. The interatomic energy is defined as

$$V_{inter}^{AB} = V_{nn}^{AB} + V_{en}^{AB} + V_{ne}^{AB} + V_{ee}^{AB} \quad (3)$$

where V_{nn}^{AB} , V_{en}^{AB} , V_{ne}^{AB} are, respectively, the nucleus–nucleus repulsive potential energy, the electron–nucleus attraction energy between the electrons of atom *A* and nucleus *B*, and the electron–nucleus attraction energy between the electrons of atom *B* and nucleus *A*. A key energetic component of this interatomic partitioning is the electron–electron repulsion between two atoms, which can be defined as

$$V_{ee}^{AB} = V_{coul}^{AB} + V_x^{AB} + V_c^{AB} \quad (4)$$

where V_{coul}^{AB} represents the classical Coulombic energy (between electrons only), V_x^{AB} the exchange energy, and V_c^{AB} the correlation energy between the electrons of atoms A and B. The V_{ee}^{AB} term is obtained through a six-dimensional integral over atoms A and B, which is computationally expensive. Note that separate contributions for exchange and correlation energies are typically obtained when working with post Hartree–Fock methods, where dynamic correlation is explicitly accounted for [59]. A more common approach is to combine exchange and correlation energies to obtain a parametrised functional form, as in Density Functional Theory (DFT). Both post-Hartree–Fock and DFT methods [60,61] have been made compatible with the IQA partitioning. However, the computational expense [62] of the post-Hartree–Fock IQA partitioning is very high for even medium-sized systems, which is why it is common to proceed with DFT as we do here. After insertion of Equation (4) into Equation (3) and using the fact that $V_{xc}^{AB} = V_x^{AB} + V_c^{AB}$ the V_{inter}^{AB} energy then becomes:

$$V_{inter}^{AB} = \left(V_{nn}^{AB} + V_{en}^{AB} + V_{ne}^{AB} + V_{coul}^{AB} \right) + V_{xc}^{AB} = V_{cl}^{AB} + V_{xc}^{AB} \quad (5)$$

where V_{cl}^{AB} is the classical electrostatic term, and V_{xc}^{AB} is a combination of exchange and correlation energies. The former term can be related to ionisation and polarisation effects, while the latter term is mainly connected to the concept of covalency.

In DFT, both static and dynamic correlations are accounted for through an empirically parametrised functional. However, these functionals are usually parametrised for short-range correlation. When using DFT, long-range correlation can be approximated by means of a London dispersion correction, such as D3. Given the pairwise interaction and not-many-body, additive nature of the D3 correction, we can combine the interatomic IQA energy with the interatomic empirical dispersion energy (V_{D3}^{AB}):

$$V_{inter}^{AB} = V_{cl}^{AB} + V_{xc}^{AB} + V_{D3}^{AB} \quad (6)$$

This makes DFT-D3 compatible with the IQA partitioning. Thus, the total energy of the wavefunction can be correctly recovered when employing DFT-D3 calculations and IQA partitioning. In this study, we employ the D3 dispersion correction and compute the V_{D3}^{AB} contribution for every pairwise interaction.

Additionally, we can partition the electrostatic energy V_{cl}^{AB} into a monopolar charge-transfer contribution V_{ct}^{AB} and a polarisation contribution V_{pl}^{AB}

$$V_{cl}^{AB} = V_{ct}^{AB} + V_{pl}^{AB} \quad (7)$$

$$V_{ct}^{AB} = \frac{q_A q_B}{r_{AB}} \quad (8)$$

$$V_{pl}^{AB} = V_{cl}^{AB} - V_{ct}^{AB} \quad (9)$$

Note that the V_{cl}^{AB} can also be Taylor-expanded in order to obtain higher-order multipole moments in a faster but approximate way [44,63].

Topological atoms are quantum objects that occupy a specific volume (of electron density) in real 3D space. These atoms do not overlap, nor do they leave gaps between them. As a result, it becomes very easy to lump adjacent atoms into groups or fragments, which can be of arbitrary size. This space-exhausting partitioning scheme also means that the IQA partitioned energies can be simply added without the need to correct for overlap or energy not assigned to an atom. Thus, it is also easy to sum atoms within a functional group to understand the role of the functional group within the molecule. When multiple IQA energies are added to make up a (system's) fragment consisting of atoms, it is referred to as Interacting Quantum Fragments (IQF). For example, consider two groups of atoms: G_1 with n atoms and G_2 with m atoms, where the atoms in group G_1 are not in group

G₂ (and vice versa). Then, the exchange–correlation and electrostatic interaction energies between the groups are simply

$$V_{xc}^{G_1G_2} = \sum_A^n \sum_{B \neq A}^m V_{xc}^{AB} \quad (10)$$

$$V_{cl}^{G_1G_2} = \sum_A^n \sum_{B \neq A}^m V_{cl}^{AB} \quad (11)$$

The process of fragmentation can provide valuable insights into the overall impact of one group on another, as it takes into account any potential cancellation between atomic energies. As a result, IQF may well offer a more comprehensive understanding of the interactions within the system, albeit with a certain level of arbitrariness, compared to the atomistic perspective due to how atoms are grouped.

3.2. Relative Energy Gradient (REG)

When considering a system of N atoms, the IQA partitioning leads to a total of N intra-atomic terms and $\frac{N(N-1)}{2}$ interatomic terms for each type of interatomic energy. Thus, the total number of interactions grows quadratically (N^2), which makes a manual analysis of every possible interaction in a system burdensome. Even more challenging is detecting which interactions, by atom (or group) and by type, cause a chemical phenomenon such as a reaction, a conformational change, or the formation of a dimeric complex from its monomeric constituents. REG solves this problem by automating the analysis of all intra- and interatomic energies in a system. For that purpose, REG operates with a control coordinate s , which governs the geometric change behind the chemical phenomenon. For example, for hydrogen bond formation, a possible control coordinate is the hydrogen bond length. As will become clear later, the REG analysis of the molecular balances will consist of a collection of dihedral rotations and relaxations (i.e., a collective variable). It should also be clear that REG always needs a set of molecular geometries representing the chemical change under study.

There is a set of M molecular geometries and N^2 individual energy terms obtained after the IQA energy partitioning. Over the energy segment covered by M data points, we then quantify the correlation between the total energy of a system (E_{tot}) and each i -th partitioned energy term (E_i), using ordinary least-squares regression,

$$m_{REG,i} = \frac{\left(\mathbf{E}_{tot}^{transl}\right)^T \cdot \mathbf{E}_i^{transl}}{\left(\mathbf{E}_{tot}^{transl}\right)^T \cdot \mathbf{E}_{tot}^{transl}} \quad (12)$$

where

$$\mathbf{E}_{tot}^{transl} = [E_{tot}(s_1) - \bar{E}_{tot}, E_{tot}(s_2) - \bar{E}_{tot}, \dots, E_{tot}(s_M) - \bar{E}_{tot}]$$

$$\mathbf{E}_i^{transl} = [E_i(s_1) - \bar{E}_i, E_i(s_2) - \bar{E}_i, \dots, E_i(s_M) - \bar{E}_i]$$

Here, T denotes the vector transpose, while “*transl*” stands for translation, which refers to subtracting the relevant mean. This mean, whether of the individual (atomic) or the total energy terms, is obtained by averaging over the M molecular geometries. This translation enables the comparison of least-squares fits across different energy profiles. The coefficient $m_{REG,i}$ is the slope of the least-squares line, also simply called the *REG value*. This value is dimensionless. Note that the REG method is only valid when there is strong linearity between the total energy and the energetically partitioned term. Whether this is the case can be easily verified by inspecting the coefficient of determination, denoted as R^2 . Typically, only energy terms whose regression has an R^2 coefficient above 0.8 are considered. Equivalently, the correlation can be quantified (and thus confirmed) through calculation of the Pearson correlation coefficient R , which can be positive or negative.

Once the ordinary least-squares regression is computed, a set of REG values is obtained for a system along a control coordinate. The latter governs the progression of a molecular

balance between its open and closed states. The REG values are ranked from largest to smallest, that is, from most positive to most negative. A highly positive REG value expresses that the corresponding energy contribution is most relevant. Indeed, it dominates in constructing the total potential energy surface (PES) from its energy contribution. In other words, it works in favour of the total PES. In contrast, a negative REG value, high in absolute terms, demonstrates that the corresponding energy contribution is also important for the total potential energy, but this time, it works against it. More details of the REG method have been published [35] before.

4. Computational Details

Geometry optimisations and frequency calculations for all 1-X and 2-Y balances were performed using the program GAUSSIAN16 at the B3LYP/6-311++G(d,p) level of theory and the D3 dispersion correction with Becke–Johnson damping [64–66]. Some tests were also run with other functionals (Table S1). Frequency calculations of the optimised geometries yielded no imaginary frequencies, thereby confirming their minimum status. Calculations of transition states were performed using the Berny optimisation algorithm starting from a plausible transition state geometry [67]. Transition state frequency calculations yielded only one imaginary frequency corresponding to the expected motion from a closed to an open conformation. The wavefunction obtained for each point of the PESs was fed to the quantum chemical topology program AIMAll, version 19.10.12 [68]. The REG-IQA analysis is performed with the in-house program [69] REG.py. Images were generated with the open-source PyMol package (Version 2.5) [70].

5. Conclusions

Non-covalent interactions play a critical role in determining the shape and function of proteins and physical properties in materials. In recent decades, scientists have been trying to understand the forces that drive these interactions. One way to do this is to study small organic molecules, called molecular balances, which can resemble subunits of larger biological systems.

In this study, we used DFT and a real-space topological partitioning of the electron density to study two molecular balances designed to probe a carbonyl–carbonyl interaction. These balances were designed by Cockcroft et al. (formamide balance) and Raines et al. (N-formylproline balance). We pair the QTAIM-based energy decomposition IQA with the REG method, which helps in understanding which atoms or fragments contribute to the formation of the closed conformers of molecular balances.

We first performed a simple conformational search for the 1-X balances, which showed that the conformational equilibrium usually depicted in experiments is more complicated than these back-of-the-envelope drawings. This is because the conformational equilibrium involves the rotation of all the rotatable bonds in the balance, not just the rotation of the amide bond.

The fully atomistic REG-IQA analysis shows a clear separation between the 1-H and 1-OMe balances from the 1-Me and 1-NMe₂ balances, demonstrating that the latter two are probing the carbonyl...carbonyl interaction better than the former two because of the bulky X functional groups and stronger electrostatic effects. Thus, we believe that even more sterically hindering X groups will be helpful in isolating the carbonyl...carbonyl contributions. On the other hand, the 2-Y balances all behave similarly in the REG-IQA analysis, with electrostatics between the two carbonyl groups playing the major role.

However, we also find that effects not directly related to carbonyl...carbonyl interactions are the most relevant. For example, the most prominent REG-IQA values out of all interactions in the 1-X and 2-Y series are (i) the steric effect on the amide nitrogen related to its pyramidalisation and (ii) the charge-transfer and polarisation of the C-N bond related to the shortening of the latter. These findings show that the conformational equilibrium probed might not be the intuitively expected single bond rotation from an open to a closed conformation but that more complicated phenomena are involved.

Finally, we looked at meaningful contributions of different atoms and fragments around the carbonyl moieties for both series of balances. The REG-IQF analysis and comparison show that the well-known $n\text{-}\pi^*$ interaction, which is defined as the lone pair of the oxygen on one carbonyl donating to the antibonding π^* orbital of the carbon of the opposite carbonyl, is mainly electrostatically driven. Extending the interaction to consider atoms around the $\text{C}=\text{O} \cdots \text{C}=\text{O}$ moieties shows that electrostatics are still the most relevant but only for the 1-Me balance. For the 1-OMe and 1-NMe₂ balances, electrostatics work on par with exchange–correlation. For the 1-H and all 2-Y series, exchange–correlation energies between the largest fragments take the lead, while electrostatics work against the formation of the closed conformer.

Overall, our results suggest that the original orbital-based definition of $n\text{-}\pi^*$ interactions, which basically involves only two atoms, is not correct because many more atoms provide a contribution to the formation of carbonyl $\cdots$ carbonyl interactions. We also showed that a combination of quantum mechanical methods, such as DFT and REG-IQA, are needed to understand and design new balances prior to experiments.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/molecules29051043/s1>.

Author Contributions: Conceptualization, F.F. and P.L.A.P.; Software, F.F.; Validation, F.F. and P.S.-J.; Formal analysis, F.F. and P.S.-J.; Investigation, S.B.; Resources, P.L.A.P.; Data curation, S.B. and P.S.-J.; Writing—original draft, F.F., S.B. and P.S.-J.; Writing—review & editing, P.L.A.P.; Supervision, P.L.A.P.; Funding acquisition, P.L.A.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the UKRI Frontier Research grant number EP/XO24393/1.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The geometries, IQA calculations and REG-IQA results for all the 1-X and 2-Y series of balances are stored at Zenodo (<https://zenodo.org/records/10033247>), (accessed on 31 January 2024).

Acknowledgments: F.F. thanks the EPSRC and AstraZeneca for funding. P.L.A.P. is grateful to the European Research Council (ERC) for the award of an Advanced Grant underwritten by the UKRI-funded Frontier Research grant EP/XO24393/1. The authors acknowledge the assistance given by Research IT and the use of the Computational Shared Facility at The University of Manchester.

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

References

1. Kollman, P.A. Noncovalent interactions. *Acc. Chem. Res.* **1977**, *10*, 365–371. [[CrossRef](#)]
2. Mueller-Dethlefs, K.; Hobza, P. Noncovalent Interactions: A challenge for experiment and theory. *Chem. Rev.* **2000**, *100*, 143–167. [[CrossRef](#)]
3. London, F. The general theory of molecular forces. *Trans. Faraday Soc.* **1937**, *33*, 8b–26. [[CrossRef](#)]
4. Dzyaloshinskii, I.E.; Lifshitz, E.M.; Pitaevskii, L.P. The general theory of van der Waals forces. *Adv. Phys.* **1961**, *10*, 165–209. [[CrossRef](#)]
5. Hobza, P.; Müller-Dethlefs, K. *Non-Covalent Interactions: Theory and Experiment*; Royal Society of Chemistry: London, UK, 2010; Volume 2.
6. Paliwal, S.; Geib, S.; Wilcox, C.S. Molecular torsion balance for weak molecular recognition forces. Effects of “tilted-T” edge-to-face aromatic interactions on conformational selection and solid-state structure. *J. Am. Chem. Soc.* **1994**, *116*, 4497–4498. [[CrossRef](#)]
7. Cozzi, F.; Cinquini, M.; Annunziata, R.; Dwyer, T.; Siegel, J.S. Polar π interactions between stacked aryls in 1,8-diarylnaphthalenes. *J. Am. Chem. Soc.* **1992**, *114*, 5729–5733. [[CrossRef](#)]
8. Cozzi, F.; Siegel, J.S. Interaction between stacked aryl groups in 1,8-diarylnaphthalenes: Dominance of polar/ π over charge-transfer effects. *Pure Appl. Chem.* **1995**, *67*, 683–689. [[CrossRef](#)]
9. Lin, B.; Karki, I.; Pellechia, P.J.; Shimizu, K.D. Electrostatically-gated molecular rotors. *Chem. Commun.* **2022**, *58*, 5869–5872. [[CrossRef](#)] [[PubMed](#)]
10. Vik, E.C.; Li, P.; Madukwe, D.O.; Karki, I.; Tibbetts, G.S.; Shimizu, K.D. Analysis of the Orbital and Electrostatic Contributions to the Lone Pair–Aromatic Interaction Using Molecular Rotors. *Org. Lett.* **2021**, *23*, 8179–8182. [[CrossRef](#)]

11. Li, P.; Vik, E.C.; Shimizu, K.D. N-Arylimide Molecular Balances: A Comprehensive Platform for Studying Aromatic Interactions in Solution. *Accs. Chem. Res.* **2020**, *53*, 2705–2714. [[CrossRef](#)]
12. Vik, E.C.; Li, P.; Pellechia, P.J.; Shimizu, K.D. Transition-state stabilization by $n \rightarrow \pi^*$ interactions measured using molecular rotors. *J. Am. Chem. Soc.* **2019**, *141*, 16579–16583. [[CrossRef](#)]
13. Li, P.; Vik, E.C.; Maier, J.M.; Karki, I.; Strickland, S.M.; Umana, J.M.; Smith, M.D.; Pellechia, P.J.; Shimizu, K.D. Electrostatically driven CO— π aromatic interactions. *J. Am. Chem. Soc.* **2019**, *141*, 12513–12517. [[CrossRef](#)] [[PubMed](#)]
14. Carroll, W.R.; Pellechia, P.; Shimizu, K.D. A Rigid Molecular Balance for Measuring Face-to-Face Arene—Arene Interactions. *Org. Lett.* **2008**, *10*, 3547–3550. [[CrossRef](#)]
15. Hwang, J.; Li, P.; Carroll, W.R.; Smith, M.D.; Pellechia, P.J.; Shimizu, K.D. Additivity of Substituent Effects in Aromatic Stacking Interactions. *J. Am. Chem. Soc.* **2014**, *136*, 14060–14067. [[CrossRef](#)]
16. Hwang, J.; Dial, B.E.; Li, P.; Kozik, M.E.; Smith, M.D.; Shimizu, K.D. How important are dispersion interactions to the strength of aromatic stacking interactions in solution? *Chem. Sci.* **2015**, *6*, 4358–4364. [[CrossRef](#)]
17. Meredith, N.Y.; Borsley, S.; Smolyar, I.V.; Nichol, G.S.; Baker, C.M.; Ling, K.B.; Cockroft, S.L. Dissecting Solvent Effects on Hydrogen Bonding. *Angew. Chem.* **2022**, *61*, e202206604. [[CrossRef](#)] [[PubMed](#)]
18. Elmi, A.; Cockroft, S.L. Quantifying interactions and solvent effects using molecular balances and model complexes. *Acc. Chem. Res.* **2020**, *54*, 92–103. [[CrossRef](#)]
19. Muchowska, K.B.; Pascoe, D.J.; Borsley, S.; Smolyar, I.V.; Mati, I.K.; Adam, C.; Nichol, G.S.; Ling, K.B.; Cockroft, S.L. Reconciling electrostatic and $n \rightarrow \pi^*$ orbital contributions in carbonyl interactions. *Angew. Chem.* **2020**, *132*, 14710–14716. [[CrossRef](#)]
20. Yang, L.; Adam, C.; Nichol, G.S.; Cockroft, S.L. How much do van der Waals dispersion forces contribute to molecular recognition in solution? *Nat. Chem.* **2013**, *5*, 1006–1010. [[CrossRef](#)]
21. Mati, I.K.; Adam, C.; Cockroft, S.L. Seeing through solvent effects using molecular balances. *Chem. Sci.* **2013**, *4*, 3965–3972. [[CrossRef](#)]
22. Burns, R.J.; Mati, I.K.; Muchowska, K.B.; Adam, C.; Cockroft, S.L. Quantifying Through-Space Substituent Effects. *Angew. Chem. Int. Ed.* **2020**, *132*, 16860–16867. [[CrossRef](#)]
23. Pascoe, D.J.; Ling, K.B.; Cockroft, S.L. The Origin of Chalcogen-Bonding Interactions. *J. Am. Chem. Soc.* **2017**, *139*, 15160–15167. [[CrossRef](#)]
24. Yang, L.; Adam, C.; Cockroft, S.L. Quantifying Solvophobic Effects in Nonpolar Cohesive Interactions. *J. Am. Chem. Soc.* **2015**, *137*, 10084–10087. [[CrossRef](#)] [[PubMed](#)]
25. Mati, I.K.; Cockroft, S.L. Molecular balances for quantifying non-covalent interactions. *Chem. Soc. Rev.* **2010**, *39*, 4195–4205. [[CrossRef](#)] [[PubMed](#)]
26. Aliev, A.E.; Motherwell, M.B. Some Recent Advances in the Design and Use of Molecular Balances for the Experimental Quantification of Intramolecular Noncovalent Interactions of π Systems. *Chem. Eur. J.* **2019**, *25*, 10516–10530. [[CrossRef](#)] [[PubMed](#)]
27. Patkowski, K. Recent developments in symmetry-adapted perturbation theory. *WIREs Comp. Mol. Sci.* **2020**, *10*, e1452. [[CrossRef](#)]
28. Szalewicz, K. Symmetry-adapted perturbation theory of intermolecular forces. *WIREs Comp. Mol. Sci.* **2012**, *2*, 254–272. [[CrossRef](#)]
29. Glendening, E.D.; Landis, C.R.; Weinhold, F. Natural bond orbital methods. *WIREs Comp. Molec. Sci.* **2012**, *2*, 1–42. [[CrossRef](#)]
30. Weinhold, F.; Landis, C. Natural Bond Orbitals and Extensions of Localized Bonding Concepts. *Chem. Educ. Res. Pract. Eur.* **2001**, *2*, 91–104. [[CrossRef](#)]
31. 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](#)]
32. Yang, L.; Brazier, J.B.; Hubbard, T.A.; Rogers, D.M.; Cockroft, S.L. Can Dispersion Forces Govern Aromatic Stacking in an Organic Solvent? *Angew. Chem. Int. Ed.* **2015**, *55*, 912–916. [[CrossRef](#)]
33. Francisco, E.; Costales, A.; Menéndez-Herrero, M.; Martín Pendás, Á. Lewis Structures from Open Quantum Systems Natural Orbitals: Real Space Adaptive Natural Density Partitioning. *J. Phys. Chem. A* **2021**, *125*, 4013–4025. [[CrossRef](#)]
34. Popelier, P.L.A. Non-covalent interactions from a Quantum Chemical Topology perspective. *J. Mol. Model.* **2022**, *28*, 276. [[CrossRef](#)]
35. Thacker, J.C.R.; Popelier, P.L.A. The ANANKE Relative Energy Gradient (REG) Method to Automate IQA Analysis over Configurational Change. *Theor. Chem. Acc.* **2017**, *136*, 86. [[CrossRef](#)]
36. Guevara-Vela, J.M.; Francisco, E.; Rocha-Rinza, T.; Martín Pendás, A. Interacting Quantum Atoms—A Review. *Molecules* **2020**, *25*, 4028. [[CrossRef](#)] [[PubMed](#)]
37. Blanco, M.A.; Martín Pendás, A.; Francisco, E. Interacting Quantum Atoms: A Correlated Energy Decomposition Scheme Based on the Quantum Theory of Atoms in Molecules. *J. Chem. Theor. Comp.* **2005**, *1*, 1096–1109. [[CrossRef](#)] [[PubMed](#)]
38. Bader, R.F.W. Atoms in Molecules. *Acc. Chem. Res.* **1985**, *18*, 9–15. [[CrossRef](#)]
39. Wilson, A.L.; Popelier, P.L.A. Exponential Relationships capturing Atomistic Short-range Repulsion from the Interacting Quantum Atoms (IQA) Method. *J. Phys. Chem. A* **2016**, *120*, 9647–9659. [[CrossRef](#)] [[PubMed](#)]
40. Gallegos, M.; Costales, A.; Martín Pendás, Á. Energetic Descriptors of Steric Hindrance in Real Space: An Improved IQA Picture. *ChemPhysChem* **2021**, *22*, 775–787. [[CrossRef](#)] [[PubMed](#)]
41. Garcia-Revilla, M.; Francisco, E.; Popelier, P.L.A.; Martín Pendás, Á. Domain-averaged Exchange Correlation Energies as a Physical Underpinning for Chemical Graphs. *ChemPhysChem* **2013**, *14*, 1211–1218. [[CrossRef](#)] [[PubMed](#)]

42. Martín Pendás, Á.; Francisco, E.; Blanco, M.A.; Gatti, C. Bond Paths as Privileged Exchange Channels. *Chem.—Eur. J.* **2007**, *13*, 9362–9371. [CrossRef]
43. Symons, B.C.B.; Williamson, D.J.; Brooks, C.M.; Wilson, A.L.; Popelier, P.L.A. Does the Intra-Atomic Deformation Energy of Interacting Quantum Atoms Represent Steric Energy? *Chem. Open* **2019**, *8*, 560–570.
44. Falcioni, F.; Symons, B.C.B.; Popelier, P.L.A. REG-MULTI: Lifting the veil on electrostatic interactions. *Adv. Quant. Chem.* **2023**, *88*, 305–328.
45. Triestram, L.; Falcioni, F.; Popelier, P.L.A. Interacting Quantum Atoms and Multipolar Electrostatic Study of $\text{XH}\cdots\pi$ Interactions. *ACS Omega* **2023**, *8*, 34844–34851. [CrossRef]
46. Fischer, F.R.; Schweizer, W.B.; Diederich, F. Molecular Torsion Balances: Evidence for Favorable Orthogonal Dipolar Interactions Between Organic Fluorine and Amide Groups. *Angew. Chem. Int. Ed.* **2007**, *46*, 8270–8273. [CrossRef]
47. Newberry, R.W.; VanVeller, B.; Guzei, I.A.; Raines, R.T. $n \rightarrow \pi^*$ interactions of amides and thioamides: Implications for protein stability. *J. Am. Chem. Soc.* **2013**, *135*, 7843–7846. [CrossRef]
48. Choudhary, A.; Kamer, K.J.; Raines, R.T. An $n \rightarrow \pi^*$ interaction in aspirin: Implications for structure and reactivity. *J. Org. Chem.* **2011**, *76*, 7933–7937. [CrossRef] [PubMed]
49. Bartlett, G.J.; Choudhary, A.; Raines, R.T.; Woolfson, D.N. $n \rightarrow \pi^*$ interactions in proteins. *Nat. Chem. Biol.* **2010**, *6*, 615–620. [CrossRef] [PubMed]
50. Hodges, J.A.; Raines, R.T. Energetics of an $n \rightarrow \pi^*$ interaction that impacts protein structure. *Org. Lett.* **2006**, *8*, 4695–4697. [CrossRef]
51. Hinderaker, M.P.; Raines, R.T. An electronic effect on protein structure. *Protein Sci.* **2003**, *12*, 1188–1194. [CrossRef]
52. Newberry, R.W.; Raines, R.T. The $n \rightarrow \pi^*$ Interaction. *Acc. Chem. Res.* **2017**, *50*, 1838–1846. [CrossRef] [PubMed]
53. Paulini, R.; Mueller, K.; Diederich, F. Orthogonal Multipolar Interactions in Structural Chemistry and Biology. *Angew. Chem. Int. Ed.* **2005**, *44*, 1788–1805. [CrossRef] [PubMed]
54. Fischer, R.A.; Schulte, M.M.; Weiss, J.; Zsolnai, L.; Jacobi, A.; Huttner, G.; Frenking, G.; Boehme, C.; Vyboishchikov, S.F. Transition metal coordinated Al(X)L-2 and Ga(X)L-2 fragments. *J. Amer. Chem. Soc.* **1998**, *120*, 1237–1248. [CrossRef]
55. Hunter, C.A. Quantifying Intermolecular Interactions: Guidelines for the Molecular Recognition Toolbox. *Angew. Chem.* **2004**, *43*, 5310–5324.
56. Alkorta, I.; Elguero, J.; Popelier, P.L. A relative energy gradient (REG) study of the nitrogen inversion in N-substituted aziridines. *Chem. Phys. Lett.* **2020**, *758*, 137927. [CrossRef]
57. Kemnitz, C.R.; Loewen, M.J. “Amide resonance” correlates with a breadth of C–N rotation barriers. *J. Am. Chem. Soc.* **2007**, *129*, 2521–2528. [CrossRef]
58. Clayden, J.; Greeves, N.; Warren, S. *Organic Chemistry*; Oxford University Press: Oxford, UK, 2012.
59. Silva, A.F.; Duarte, L.J.; Popelier, P.L.A. Contributions of IQA electron correlation in understanding the chemical bond and non-covalent interactions. *Struct. Chem.* **2020**, *31*, 507–519. [CrossRef]
60. Maxwell, P.; Martín Pendás, A.; Popelier, P.L.A. Extension of the Interacting Quantum Atoms (IQA) Approach to B3LYP Level Density Functional Theory. *PhysChemChemPhys* **2016**, *18*, 20986–21000. [CrossRef]
61. Francisco, E.; Casals-Sainz, J.L.; Rocha-Rinza, T.; Martín Pendás, A. Partitioning the DFT Exchange-Correlation Energy in Line with the Interacting Quantum Atoms Approach. *Theor. Chem. Accs.* **2016**, *135*, 170. [CrossRef]
62. Vincent, M.A.; Popelier, P.L.A. IQA analysis of the two-particle density matrix: Chemical insight and computational efficiency. *Theor. Chem. Acc.* **2023**, *142*, 119. [CrossRef]
63. Stone, A.J. *The Theory of Intermolecular Forces*, 2nd ed.; Clarendon Press: Oxford, UK, 2013; Volume 32, p. 264.
64. Frisch, M.J.; Scuseria, H.B.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; Li, X.; Caricato, M.; et al. GAUSSIAN16. 2016. Available online: <https://gaussian.com/gaussian16/> (accessed on 31 January 2024).
65. Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465. [CrossRef] [PubMed]
66. Grimme, S. Density functional theory with London dispersion corrections. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 211–228. [CrossRef]
67. Schlegel, H.B. Optimization of Equilibrium Geometries and Transition Structures. *J. Comp. Chem.* **1982**, *3*, 214–218. [CrossRef]
68. Keith, T.A.; AIMAll. *TK Gristmill Software*; AIMAll: Overland Park, KS, USA, 2019.
69. Falcioni, F.; Duarte, L.J.; Popelier, P.L.A. REG.py (Version 0.1). Available online: <https://github.com/popelier-group/REG> (accessed on 31 January 2024).
70. Schrödinger, L.L.C. *The PyMOL Molecular Graphics System, Version 2.5*; Schrödinger, LLC: New York, NY, USA, 2022.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.