MDPI

Article

Cross-Talk between Overlap Interactions in Biomolecules: A Case Study of the *β*-Turn Motif

Jayashree Nagesh

Solid State and Structural Chemistry Unit, Indian Institute of Science Bangalore, Bengaluru 560012, Karnataka, India; jayashreen@iisc.ac.in

Abstract: Noncovalent interactions play a pivotal role in regulating protein conformation, stability and dynamics. Among the quantum mechanical (QM) overlap-based noncovalent interactions, $n \rightarrow \pi^*$ is the best understood with studies ranging from small molecules to *β*-turns of model proteins such as GB1. However, these investigations do not explore the interplay between multiple overlap interactions in contributing to local structure and stability. In this work, we identify and characterize all noncovalent overlap interactions in the *β*-turn, an important secondary structural element that facilitates the folding of a polypeptide chain. Invoking a QM framework of natural bond orbitals, we demonstrate the role of several additional interactions such as $n \rightarrow \sigma^*$ and $\pi \rightarrow \pi^*$ that are energetically comparable to or larger than $n \rightarrow \pi^*$. We find that these interactions are sensitive to changes in the side chain of the residues in the *β*-turn of GB1, suggesting that the $n \rightarrow \pi^*$ may not be the only component in dictating *β*-turn conformation and stability. Furthermore, a database search of $n \rightarrow \sigma^*$ and $\pi \rightarrow \pi^*$ in the PDB reveals that they are prevalent in most proteins and have significant interaction energies ($\sim$1 kcal/mol). This indicates that all overlap interactions must be taken into account to obtain a comprehensive picture of their contributions to protein structure and energetics. Lastly, based on the extent of QM overlaps and interaction energies, we propose geometric criteria using which these additional interactions can be efficiently tracked in broad database searches.

Keywords: beta-turns; noncovalent interactions; n-pi*; protein secondary structure; protein conformation; protein stability; NBO deletion analysis; bioinformatics

1. Introduction

Noncovalent interactions (NCIs) have been a central topic of interest in chemistry ever since the conception of chemical bonding ideas [1]. There is a recent surge of interest in the role of such interactions in biology with increase in the ability to handle large amounts of data involving molecular structures as well as advancements in the identification and classification of different interactions. The last two decades have brought forth the importance of NCIs such as π-π, ion-π, and lone-pair-π, to add to the existing electrostatic, hydrogen bonding and hydrophobic effects in protein stability [2,3], enzyme catalysis [4,5], molecular recognition [6,7] and phase separation [8]. These interactions are gaining importance not only as factors in modulating structure and stability, but also in facilitating kinetic events in molecular recognition and enzyme catalysis.

The importance of quantum mechanical (QM) orbital overlap based interactions in small molecules (<100 atoms) is well-established. Such interactions are typically detected based on the contact distance between a pair of atoms being smaller than the sum of their van der Waals radii [9], and could be short-range filled-filled repulsive type or long-range attractive donor-acceptor type. Several recent reports have shown that donor-acceptor groups can participate in more than one type of interaction depending on their relative orientations. For instance, the C=O group is known to engage in hydrogen bonding (C=O···H) [10], carbonyl-chalcogen (C=O···X where X can be Sulphur, Selenium or Tellurium) [11], and nucleophile-carbonyl (Burgi-Dunitz trajectory type) interactions [12,13].

The last group of aforementioned interactions, involving $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ overlaps [14] have been found in dimers of acetone and formaldehyde. Additionally, the C=O group orbitals can also overlap with σ orbitals [14] in its proximity. Other examples of the co-existence of several long-range interactions include $n \rightarrow \pi^*$ and hydrogen bonds [15,16], CO-CO and hydrogen bonds [17], and cation-π, anion-π and π-π [18]. Hence it is clear that the structure and electronic stability of a molecule is intimately governed by the interplay of multiple overlap interactions in addition to classical interactions such as salt bridges.

In comparison to small molecules, there are far fewer investigations of overlap interactions in biomolecules, and studies have primarily focused on classical electrostatic and hydrogen bonding interactions [19,20]. Some of the earliest overlap-based studies were centered around using the natural bond orbital (NBO) approach to understand hydrogen bond networks (see for example [21,22]) that are known to stabilise the native state of proteins [23,24]. Over the years, the characterisation of multiple (for example in hydrogen bond networks in DNA) [21,22] and multi-valent (for example in π-π networks in liquid-liquid phase separation) [8] interactions have revealed the complexity in understanding the basis of biomolecular structure.

The $n \rightarrow \pi^*$ interaction is the next best studied overlap type of interaction in biomolecules after hydrogen bonding [25,26]. It involves the sharing of electrons between a filled non-bonding orbital of a heteratom (like O, N, S) and an empty π^* orbital nearby. In β-turns of proteins, this is an inter-residue interaction comprising of $O_i(\text{n}) \rightarrow (\text{C=O})_{i+1}(\pi^*)$ or $O_{i+1}(\text{n}) \rightarrow (\text{C=O})_{i+2}(\pi^*)$. Recent experiments have established the presence of this interaction spectroscopically [27–29] in small molecules, and through melting temperature and free energy measurements in proteins [30]. In the context of proteins, Khatri et al. [30] examined the influence of $n \rightarrow \pi^*$ interaction on protein stability by engineering the sequence of a β-turn, thereby affecting the strength of the interaction itself. It was seen that changing the side chain identity of the residues in the β-turn resulted in a good correlation between the computationally estimated $n \rightarrow \pi^*$ interaction energy and the melting temperature of the proteins, suggesting that one could control protein stability by systematically altering the $n \rightarrow \pi^*$ interaction of the β-turn. Although the $n \rightarrow \pi^*$-melting temperature correlation is encouraging, the role of local electronic factors, such as other interactions, the solvent and the unfolded state remains unclear.

Here we identify additional QM overlap based interactions such as $n \rightarrow \sigma^*_{CH}$, $n \rightarrow \sigma^*_{CC}$ and $\pi \rightarrow \pi^*$ in a β-turn of the model protein GB1 (PDB:2QMT). We find that upon side chain substitution, these interactions show compensatory effects in terms of their electronic stabilization of the local β-turn conformation. We explore the prevalence of these overlap interactions in β-turns of several protein structures in the protein data bank (PDB) and find that they are prominent with interaction strengths of $\sim$1 kcal/mol. Taken together, our results suggest that a holistic view of all interactions is necessary to understand the influence of NCIs on biomolecular structure and function.

2. Results

2.1. Identifying Overlap-Based Interactions in β-Turns of Proteins

The β-turn is an important secondary structure motif that permits a change in the direction of a peptide chain. A β-turn is characterised by a chain of four amino acids, labeled as i, $i + 1$, $i + 2$ and $i + 3$ (Figure 1). It is an ideal model system for studying key interactions responsible for directing and stabilising the folding of the peptide chain. Recently, Khatri et al. [30] examined the effect of structural changes in the loop L1 β-turn of a model protein GB1 (Figure S2) on the strength of the $n \rightarrow \pi^*$ interaction and in turn on protein stability. Four β-turn variants of GB1 were synthesised by substituting Lys10 at the $i + 1$ position of L1 by D-valine("v") and Thr11 at the $i + 2$ position by the L-amino acids alanine, serine, valine and threonine to obtain "vA", "vS", "vV" and "vT" variants respectively. The use of D-valine fixes the type of the β-turn to be II', and systematic engineering of the $i + 2$ position creates differences in the backbone angles $\phi(C - N - C_\alpha - C)$ and $\psi(N - C_\alpha - C - N)$, that in turn tunes the $i + 1$–$i + 2$ inter-residue

$n \to \pi^*$ interaction. Using the NBO approach on the crystal structures of these β-turn variants (6L9B (vA), 6L91 (vT), 6L9D (vS) and 6LJI (vV); also see SI of Ref. [30]) it was found that the energy of the $i + 1$–$i + 2 \, n \to \pi^*$ interaction in these β-turns is in the range 0.20–0.46 kcal/mol. Here, we first explore the presence of other long-range interactions in β-turns that may be comparable in energy to $n \to \pi^*$ and hence likely play as much of an important role as the $n \to \pi^*$. In order to make systematic comparisons with previous studies, we initially focus on the $i + 1$–$i + 2$ residue pair.
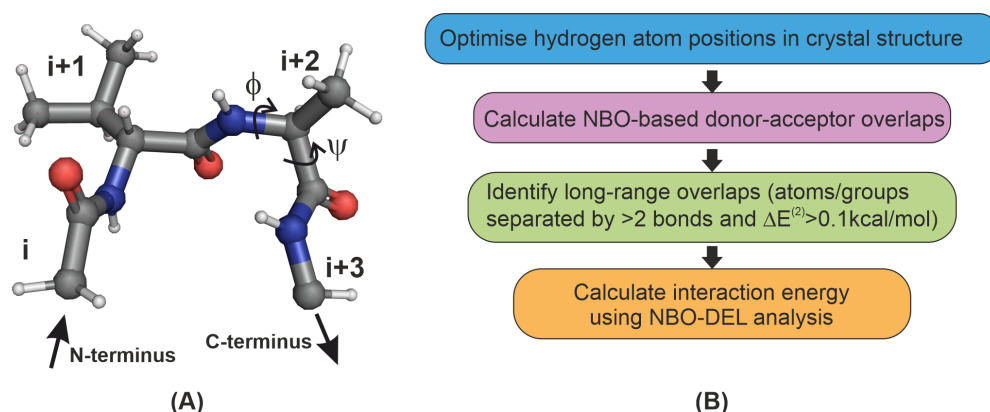

Figure 1. (**A**) Typical tetrapeptide showing the $i + 1$ and $i + 2$ residues at a β-turn and the backbone (ϕ, ψ) angles used in this work. (**B**) Flowchart of the steps used in identifying donor-acceptor interactions. See Section 4 for a description of the second order perturbation energy $\Delta E^{(2)}$.

The process of identifying noncovalent interactions is shown in Figure 1B. All calculations in this paper were carried out with wB97XD [31]/6-31+G(d,p) model chemistry using the Gaussian suite of programs [32] in conjunction with the NBO 6.0 package [33]. The tetrapeptide alone was used in the quantum chemical calculations with the i and $i + 3$ residues capped off with hydrogen atoms (see SI for coordinates). Only the positions of the hydrogen atoms were optimized, and the rest of the atomic positions were kept the same as the crystal structure. A polarizable continuum model [34] (Section 4.2) was used to implicitly account for the water solvent. To identify inter-residue interactions, (a) overlaps within a peptide bond or those involving a common atom were excluded, and (b) overlaps involving atoms or groups separated by at least two covalent bonds were considered; this ensures that we do not include overlaps due to rotation about a covalent bond. The criterion of $\Delta E^{(2)} > 0.1$ kcal/mol (see Section 4) was chosen to screen for the interactions that are of comparable or higher strength than the $n \to \pi^*$. This exhaustive procedure of sifting through all possible overlaps and interaction energies followed by specific screening criteria revealed numerous and previously neglected interactions apart from the $n \to \pi^*$ (Figure 2 and Table 1).

Table 1. Description of donor and acceptor orbitals of $i + 1$-$i + 2$ overlap interactions found in a β-turn.

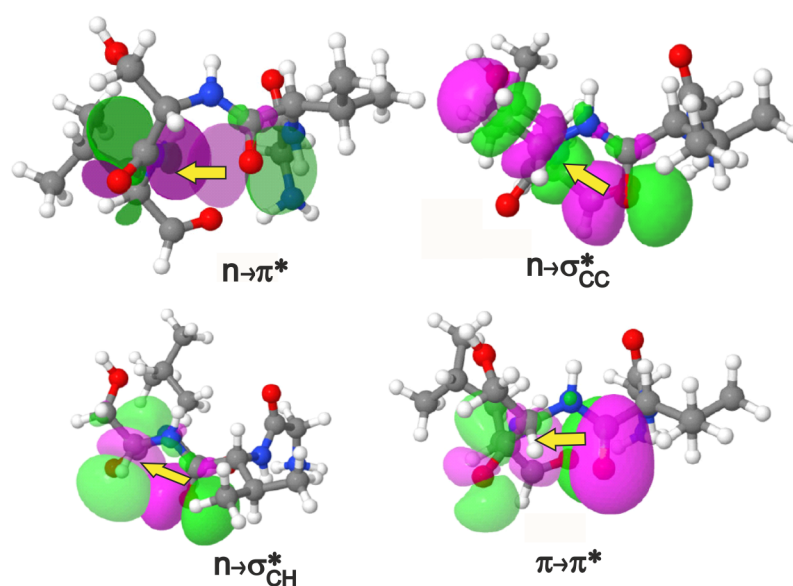| Interaction	Donor ($i + 1$)		Acceptor ($i + 2$)		Covalent Bond
	Atom(s)	Orbital	Atom(s)	Orbital	Separation
$n \to \pi^*$	O	n	C=O	π^*	4
$n \to \sigma^*_{CC}$	O	n	C_α-C_β	σ^*	3
$n \to \sigma^*_{CH}$	O	n	C_α-H_α	σ^*	3
$\pi \to \pi^*$	C=O	π	C=O	π^*	3

Figure 2. The $n \rightarrow \pi^*$, $n \rightarrow \sigma^*$ (2 types) and $\pi \rightarrow \pi^*$ interactions shown in terms of donor-acceptor orbital overlaps in a view where the $i + 1$–$i + 2$ residues project out of the plane of the paper towards the reader. An isovalue of $0.02 \, \text{e}/\text{Å}^3$ was used in the plots. The $i + 1$–$i + 2$ pairs shown above are valine-serine for the depiction of $n \rightarrow \pi^*$, $n \rightarrow \sigma^*_{CH}$ and $\pi \rightarrow \pi^*$, and valine-threonine for $n \rightarrow \sigma^*_{CC}$. The magenta/green stand for opposite phases of the orbital wavefunction. If we choose to assign positive values to the magenta lobe of the orbital in that region of 3D space, then the green lobe represents negative values of the same orbital and vice versa.

2.2. Analysis of NCIs in β-Turns in Terms of Structure and Interaction Energy

Once all the spatially overlapping interactions were identified, we subjected them to a NBO-deletion analysis (see Section 4) to estimate their stabilising effect on the electronic energy of a given geometry of the tetrapeptide. The deletion energy is a useful indicator for gauging the effect of a particular interaction to the electronic energy since it is based on the extent of overlap between the donor and acceptor orbitals in this interaction (see Section 4). Figure 3 shows a comparison amongst the energetic contributions of all the four overlap interactions in each of the β-turn variants. Upon change of side chain identity, there are changes to the individual interaction energies, but the overall energy stays about the same (1–1.2 kcal/mol). This is an intriguing observation because although a mutation is expected to bring about local changes in the electronic interactions, the total interaction energy locally seems to maintained approximately the same. This constancy of the overall energy comes from a compensatory effect of the interactions. In particular, it is known from previous studies [30] that the vV variant does not engage in an $n \rightarrow \pi^*$ interaction. The identification of additional interactions reveals that vV is instead stabilised by $n \rightarrow \sigma^*_{CH}$ and $n \rightarrow \sigma^*_{CC}$. The $\pi \rightarrow \pi^*$ interaction mirrors the trends shown by $n \rightarrow \pi^*$ as expected, since both involve interaction between the carbonyl groups of adjacent residues. Upon variation of the side chain, we see that the $n \rightarrow \sigma^*_{CC}$ is the least sensitive in its strength, indicating that it might not be important in relative stabilization studies of these β-turns. However this lack of sensitivity may not be a universal trend, and remains to be confirmed by examination of β-turns in other proteins (see below).

In order to understand the above results, we examine the structural variations induced by the side chain substitution of the loop L1 variants of GB1 protein. Figure 4 and Table 2 show how various internal coordinates, including the backbone angles of ϕ and ψ are affected by changing the $i + 2$ residue. The internal coordinates identified here come from simple considerations. In the $n \rightarrow \sigma^*_{CC}$ interaction, the O^{i+1}, C_α^{i+2} and C_β^{i+2} are involved, hence the angle τ is a prudent choice. Similarly, in $n \rightarrow \sigma^*_{CH}$ and $\pi \rightarrow \pi^*$ interactions, the dihedral angles ζ and Δ are obvious choices. However, these dihedral angles are not sufficient to characterise these interactions since they cannot describe the lateral overlap

of the orbitals. Hence we also require the use of distance criteria Ω and λ to quantify the interaction strengths.
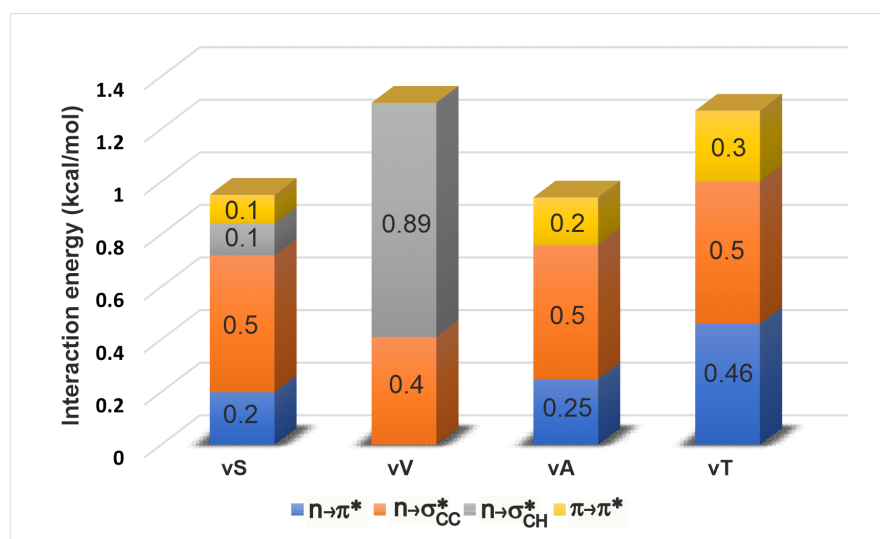

Figure 3. Energies of comparable long-range interactions obtained using NBO deletion analysis. It is seen that upon change in the amino acid (and therefore the side chain) in the $i + 2$ position, each type of interaction increases or decreases to maintain an overall contribution of ≈ 1 kcal/mol to the lowering of peptide electronic energy.
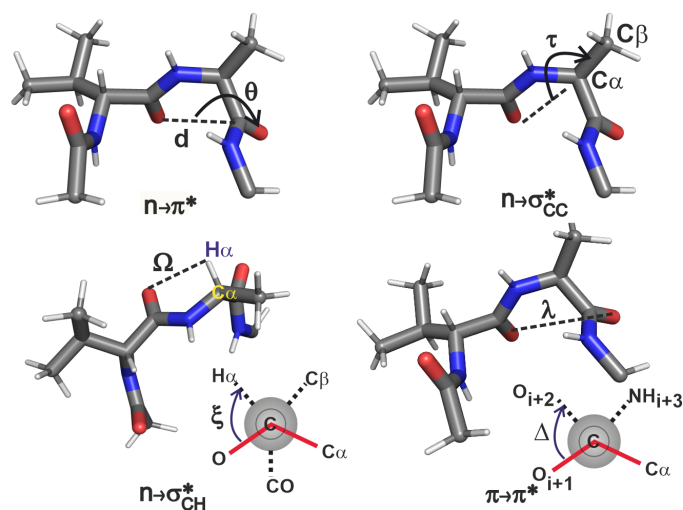

Figure 4. Internal coordinate definitions for analysing each interaction; the peptide shown for each interaction is only illustrative; all the four interactions are analysed for each peptide. The ξ dihedral angle is shown as a Newmann projection, where the C=O of the $i + 1$ residue is present in front of the plane of the paper, the C_α^{i+2} is drawn as a grey circle behind the plane of the paper, and the NH of $i + 2$ residue is not shown. The Δ dihedral angle is also shown in the Newman projection where the C=O of the $i + 1$ and $i + 2$ residues are in the front and back of the plane of the paper respectively, and the C_α and NH groups of the $i + 2$ residue are not shown.

Table 2. Geometric parameters and stabilisation energies E_{int} (kcal/mol) in the $i+1--i+2$ pair of residues pertaining to the β-turn long-range interactions. Refer Figure 4 for definitions.

β-Variant	$(\phi_{i+2}, \psi_{i+2})°$	$n \rightarrow \pi^*$			$n \rightarrow \sigma^*_{CH}$			$n \rightarrow \sigma^*_{CC}$		$\pi \rightarrow \pi^*$		
		d(Å)	$\theta(°)$	E_{int}	Ω(Å)	$\xi(°)$	E_{int}	$\tau(°)$	E_{int}	λ(Å)	$\Delta(°)$	E_{int}
vT	$(-58.8, -46.6)$	2.9	97.6	0.46	2.6	51.2	-	163.6	0.54	3.3	13.2	0.27
vA	$(-59.7, -42.2)$	3.0	102.2	0.25	2.7	44.6	-	163.6	0.51	3.5	15.4	0.18
vS	$(-66.3, -28.9)$	3.0	112.2	0.20	2.5	43.1	0.12	162.2	0.52	3.7	21.2	0.11
vV	$(-85.1, -12.4)$	3.3	122.8	-	2.4	24.5	0.89	150.1	0.41	4.2	15.8	-

Upon changing the $i+2$ residue from threonine to valine, its (ϕ, ψ) values increasingly deviate from the α-helical values of $(-60°, -45°)$ and accordingly cause the $n \rightarrow \pi^*$ interaction to decrease in its strength [30]. This trend is mirrored by the $\pi \rightarrow \pi^*$ interaction because of the increase in λ ($O^{i+1}-O^{i+2}$ distance). In the case of vV, although Δ is similar to that in vA, there is no $\pi \rightarrow \pi^*$ interaction because of the large λ. This suggests that for $\pi \rightarrow \pi^*$ interaction, λ has to be small enough before considering the Δ coordinate to rationalize the extent of overlap.

For understanding the trends in the two $n \rightarrow \sigma^*$ interactions we analysed the Ω, ξ and τ coordinates. In a comparison of the four β-turn variants, the τ values are about the same, reflecting the near constancy of the $n \rightarrow \sigma^*_{CC}$ interaction strengths. We will see that this is not necessarily the case in the β-turns of other proteins. The $n \rightarrow \sigma^*_{CH}$ on the other hand represents a contrasting situation: as the $n \rightarrow \pi^*$ becomes unfavourable from threonine to valine (with decreasing helix propensity), the $n \rightarrow \sigma^*_{CH}$ becomes favourable due to the decrease in ξ angle with increase in ψ. The decrease in ξ here is also accompanied by the decrease in the Ω ($O^{i+1}-H_\alpha^{i+2}$ distance) suggesting that in this peptide series, $n \rightarrow \sigma^*_{CH}$ plays a compensatory role in locally stabilizing the turn when $n \rightarrow \pi^*$ becomes unfavourable. These results for the four variants thus help in obtaining an atomic level picture of how the microscopic interplay of interactions could help maintain an overall local electronic stability of $\sim$1 kcal/mol in the engineered β-turn of GB1.

2.3. Prevalence of NCIs in β-Turns of PDB Structures

2.3.1. Energetics

In order to test for the prevalence of the numerous overlap interactions highlighted in this work, we chose 22 proteins from the set examined by Khatri et al. [30] (see Table S1 for the list of proteins with one selected β-turn per protein and the relevant internal coordinate values with energies and Table S2 for the amino acid sequence of the β-turn) and subjected them to same kind of analysis (Figure 1B) as the GB1 β-variants to characterize the four overlap interactions. The list of 22 proteins includes all the four types of β-turns (I,I′,II,II′) with randomly chosen structures in each type. Figure 5A shows the total stabilization energy due to these interactions, and it is evident that in contrast to the GB1 variants, NCIs in these proteins contribute varying extents to the electronic energy of the protein. This is primarily because the $i+1$ residue is fixed in GB1 variants whereas here both $i+1$ and $i+2$ residues vary from one protein to another. Another interesting observation is that by no means is the $n \rightarrow \pi^*$ interaction the dominant interaction in cases where it is present; we see instances where the $n \rightarrow \sigma^*_{CH}$ electronic energy stabilization is equal to or even greater than $n \rightarrow \pi^*$. This analysis strongly supports the notion that several interactions comparable to $n \rightarrow \pi^*$ may play a role in influencing protein structure and stability.
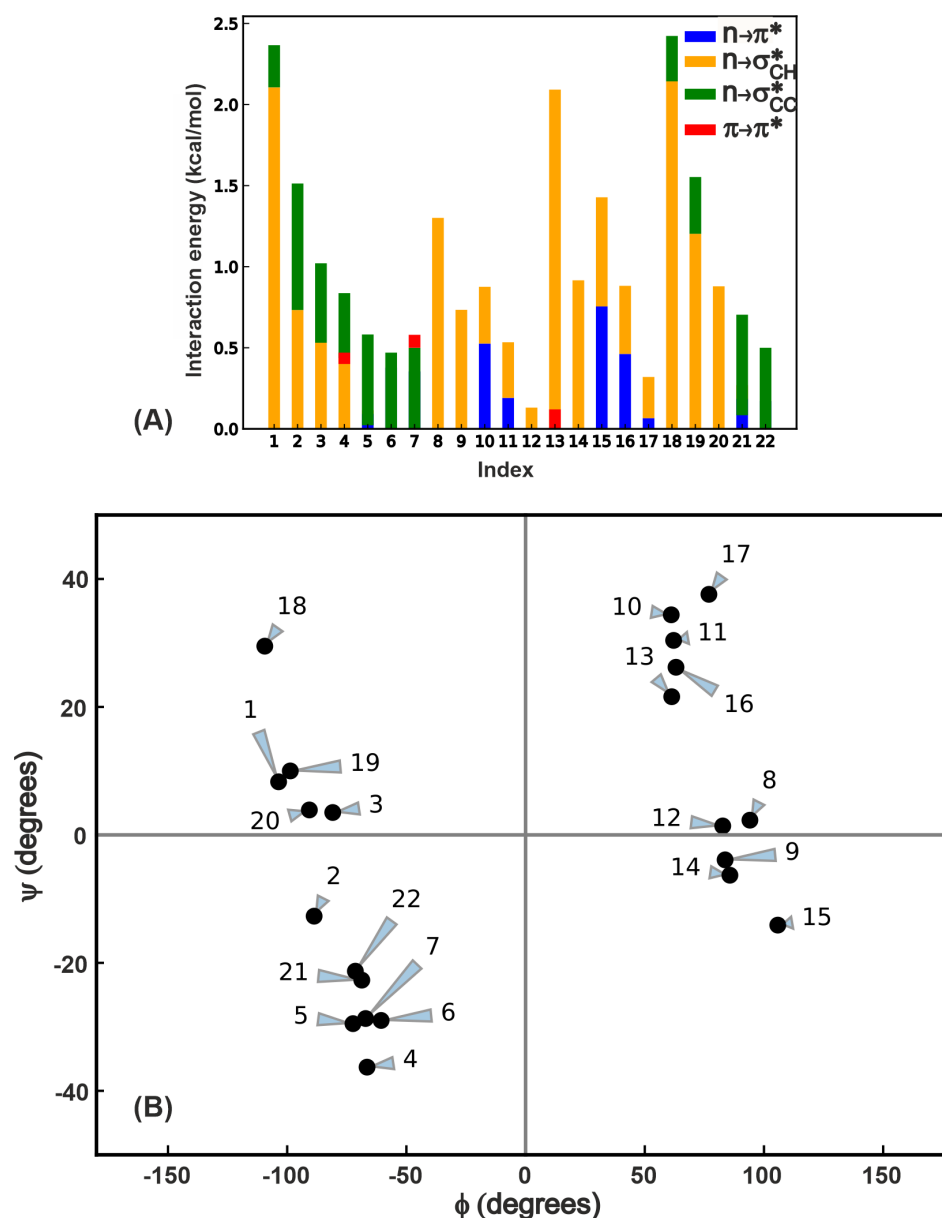
Figure 5. (**A**) Total energy of all long-range interactions in a subset of proteins from the PDB [35] that have β-turns. (**B**) Ramachandran plot of $i + 2$ residues engaged in four types of $i + 1$–$i + 2$ interactions. Each point indicates the (ϕ, ψ) of an $i + 2$ residue of a β-turn in one of the 22 PDB proteins. There may be more than one interaction for a given $i + 2$ residue, and this is indicated by the bar plot in (**A**). See SI (Figure S1) for an overlay of this plot with the traditionally allowed regions of various secondary structures in the Ramachandran map.

2.3.2. Geometric Criteria Based on Variations in β-Turn Sidechain Structure

The internal coordinates defined in Figure 4 are tabulated for the β-turns in the 22 proteins chosen above (Table S1). We observe that some of the trends in the GB1 variants are corroborated here as well, though not in all the cases. An examination of the geometric values reveals the following parameter ranges within which there are non-negligible overlaps leading to stabilisation.

- As λ increases, $\pi \to \pi^*$ decreases in strength. A small value of this coordinate is more important than the dihedral Δ in order for the $\pi \to \pi^*$ interaction to materialise. From this analysis, it is clear that this interaction exists when 0 Å$< \lambda < 3.7$ Å and $0° < \Delta < 28°$.

- The strength of the $n \to \sigma_{CC}^*$ interaction decreases as τ deviates from $180°$, suggesting that this interaction is likely to be present when $130° < \tau < 180°$. Though the lower limit of $130°$ has emerged from the current analysis of only 22 proteins, it is apparent that the extent of overlap will go to zero as we approach $\tau = 90°$ or lower.
- The $n \to \sigma_{CH}^*$ presents two possible scenarios for its existence. The σ^* orbital can overlap with the n orbital either when C=O and $C_\alpha - H_\alpha$ are on the same side of the peptide backbone, or on the opposite sides of the backbone because of the cylindrical nature of the σ^* orbital. Therefore this interaction is possible when (a) 2.2 Å$< \Omega < 2.6$ Å and $|\xi| < 53°$, and (b) 3.8 Å$< \Omega < 4.0$ Å and $160° < \xi < 200°$. Though the latter case was not seen in the GB1 variants, the PDB analysis shows that these interactions can occur in more than one orientation. Note that this is in contrast to expecting this interaction to exist simply based on the distance Ω being less than sum of van der Waals radii of O^{i+1} and H_α^{i+2}, and an exhaustive NBO analysis was essential to establish the criterion for the existence of $n \to \sigma_{CH}^*$ even when Ω is large.

2.3.3. (ϕ, ψ) Analysis

In order to place the four overlap interactions within the context of protein backbone dihedral angles, we performed a Ramachandran map analysis (Figure 5B) where we depict each of the 22 proteins by a point on this map, since there is only one β-turn chosen per protein. A given β-turn can engage in multiple interactions, and this is depicted by the contributions in Figure 5A. For example, the point at $((105°, -14°)$ of $i + 2$ residue in the protein with PDB ID 1C0I in the residue range 1186–1189) has both $n \to \pi^*$ and $n \to \sigma_{CH}^*$ interactions. The $\pi \to \pi^*$ is the least prevalent interaction in terms of the number of β-turns it occurs in (see Table S1 for the data used to generate this plot) and is also the weakest among the interactions. The Ramachandran map plot above also indicates that if the $\pi \to \pi^*$ interaction is present, there should be an $n \to \pi^*$ interaction as well, though the reverse is not necessarily true because of the requirement of lateral overlap of the π and π^* (see Figure 2) orbitals. In contrast to the lack of variation in $n \to \sigma_{CC}^*$ interaction in the β-turn variants in GB1 above, the PDB analysis reveals at least 0.5 kcal/mol variation between different β-turns. This suggests that this interaction might also play a tangible role in the local stability depending on the structural context in which it occurs. Turning to the $n \to \sigma_{CH}^*$ interaction, we see the widest variation in strength ranging from 0–2.1 kcal/mol, and its presence in turns where no other long-range interactions are evident (for instance proteins 8 and 9). These two observations along with the contributions shown in Figure 5A strongly suggest that a comprehensive view of all interactions is essential in understanding the role of NCI in protein conformational changes.

3. Discussion

NCIs form an integral part of the molecular basis determining biomolecular structure, kinetics and stability. In this paper we have identified several QM overlap based interactions in the β-turn motif in the GB1 model protein, and have found that in addition to the $n \to \pi^*$ interaction, $n \to \sigma_{CC}^*$, $n \to \sigma_{CH}^*$ and $\pi \to \pi^*$ play an important role in dictating the electronic stability of the β-turn tetrapeptide. A PDB analysis of these interactions also reveals the presence of all these interactions in β-turns and indicates that $n \to \pi^*$ is not the dominant long-range overlap interaction in several cases.

The aforementioned procedure of identifying donor-acceptor interactions using NBO analysis is general and can be applied to any structural motif in biology. The notion of interaction energy is generally defined in the context of two molecules that are not covalently bonded; however, the NBO approach can aid in the classification of interactions within the same molecule [36] (see Section 4). An important consideration in using this protocol is the size of the motif that is tractable using quantum chemistry methods. A system containing about 50 atoms or less can be studied rigorously and efficiently using the approach in this paper. Once the different types of interactions in various motifs are

identified, one can use more efficient methods of tracking interactions in the biomolecule based on robustly defined geometric criteria.

Our aim is primarily to investigate the complimentary and supplementary roles of attractive overlap interactions in β-turns. In addition to donor-acceptor type attractive interactions, there are filled-filled interactions between orbitals that are repulsive in nature [37]. These filled-filled interactions therefore can be antagonistic to attractive donor-acceptor interactions [38] in terms of defining steric boundaries between various groups. However, these interactions are likely to become important only for short contacts. An integrated view of the electronic energetics contributing to peptide stability therefore will require a consideration of such filled-filled interactions. This is beyond the scope of this study because we are not looking to make an exhaustive list of contributions to the peptide enthalpy and are only demonstrating the need to acknowledge the presence of multiple attractive interactions of similar energy.

It is well-established that the hydrophobic effect contributes to the bulk of the free energy change in the stability of a protein [39,40]. The NCIs, including hydrogen bonds, contribute far less (<5%) towards stability [23,40]. A single NCI is weak (<1 kcal/mol) and not strong enough to compete with the forces of the hydrophobic effect. However, there are studies suggesting that the cooperativity amongst the NCIs [41] is capable of tuning the specific folding pattern and functioning of RNA [42], DNA [43–45], carbohydrates [7] and proteins that mediate liquid-liquid phase separation [8,46,47]. Likewise, we suggest that even in a small motif such as the β-turn, it is the cooperativity of various interactions that dictates the formation of specific conformations. Since the interactions may complement or supplement each other, they therefore must be considered together.

The measurement of melting temperatures for each of the proteins containing the β-turn variants in Ref. [30] showed that the decreasing trend in $n \to \pi^*$ interactions correlated with a decrease in the melting temperature. In order to describe the molecular basis of the decrease in the melting temperatures it will be important to consider the effect of side chain substitution at $i + 2$ on both the unfolded and folded states, and in particular the role of solvent in differential stabilization of the unfolded state, if any [23]. We anticipate that if indeed the reduction in $n \to \pi^*$ strength is the cause of the decrease in the melting temperature, this will be substantiated by the canceling effects of the long-range intra- and inter-molecular interactions in the unfolded and folded states. This will be the subject of an upcoming study.

4. Methods

4.1. Natural Bond Order Analysis

The natural bond order analysis [48] is one of the several [49–54] ways of analysing electronic wavefunctions and densities generated by an *ab initio* quantum chemistry package. The NBO approach is one of the most consistent methods to rationalise chemical behaviour in relation to established knowledge, and has a predictive utility for new or unexpected directions. The NBOs are a result of a series of transformations from the atomic orbitals (AOs) of a chosen electronic basis through the formation of "natural AOs" [48] based on the Carlson-Keller theorem [55]. Each NBO calculation results in a set of 'Lewis-type' (donor) and non-Lewis-type (acceptor) orbitals. Unlike molecular orbitals where there is a clear demarcation between occupied and vacant orbitals, the donor and acceptor NBOs have partial occupancies of electrons. The partial occupancy of the acceptor orbitals represent the 'delocalization' (or deviation) of the donor orbitals from an 'idealised' Lewis-type localised representation. Thus the acceptor orbitals lead to lowering of the covalent bond energy and change in the overall wavefunction. In the perturbative limit where the noncovalent or delocalised contributions comprise <1% of the covalent contribution, the energy of such donor-acceptor type of interaction is estimated as

$$\Delta E_{DA}^{(2)} = -2 \frac{\langle \Phi_D | \hat{F} | \Phi_A \rangle^2}{\epsilon_A - \epsilon_D} \tag{1}$$

where $\hat{F}$ is the Fock operator [56], ϵ_D and ϵ_A are NBO energies, and {Φ_D and Φ_A} are the NBOs. This expression assumes minor energy changes to a selected donor orbital ϕ_D due to the 'interaction' of this orbital with the ϕ_A. However, interactions of ϕ_A with other donor orbitals can take place, and higher order effects need to be taken into account to get the actual change in the energy of the system due to this donor-acceptor pair. This is achieved using a deletion approach, where the interaction terms $\langle\Phi_D|\hat{F}|\Phi_A\rangle$ and $\langle\Phi_A|\hat{F}|\Phi_D\rangle$ are simply deleted from the full Fock matrix. The Fock matrix is then rediagonalised to get new NBOs, NBO energies and the overall ground electronic state energy ($\tilde{E}_0$). The resulting NBO energy for this donor orbital will be higher, and we thus get the true stabilisation energy $\Delta E_{DA} = E_0 - \tilde{E}_0$ of the molecule due to the presence of this donor-acceptor interaction. E_0 and $\tilde{E}_0$ refer to the ground electronic state energy before and after the deletion of the Fock matrix elements respectively.

4.2. Solvent Model

The polarizable continuum solvent (PCM) model involves treating the solvent as a continuum dielectric medium that interacts only electrostatically with the solute [34]. The solute is placed in void cavity that contains within its boundaries the largest possible part of the solute charge distribution. The shape of the cavity broadly outlines the molecular shape in order to avoid any unrealistic deformation due to solvent polarization. The charge distribution of the solute inside the cavity polarizes the dielectric continuum which in turn polarizes the solute charge distribution. This interaction is defined in a self-consistent manner that is numerically solved using an iterative procedure. The implicit solvent models in literature differ in the size and shape of the solute cavity, level of solute description, dielectric medium description, cavity calculation method and representation of charge distribution. In the default model used by the Gaussian program (Integral Equation Formalism-PCM, or IEFPCM) a spherical cavity is used. The solute is described using the model chemistry wB97XD/6-31+G(d,p). The dielectric medium has a fixed dielectric constant (ϵ), and the reaction field is modeled using a continuous surface charge formalism that ensures its smoothness [57]. The mathematical problem involves the solution of the general Poisson equation for the electrostatic potential (V) involving the molecular charge density ρ_{mol}

$$-\nabla^2 V = 4\pi\rho_{mol} \text{ inside the cavity,} \tag{2}$$

$$-\epsilon\nabla^2 V = 0 \text{ outside the cavity,} \tag{3}$$

$$V \to 0(\infty) \text{ where} \tag{4}$$

$$V = V_{mol} + V_R \tag{5}$$

accompanied by conditions [34] to ensure continuity of the potential across the cavity surface, and V_R is the reaction potential described by an apparent surface charge density on the cavity using Green's functions in the IEF-PCM approach. A detailed description of these functions is outside the scope of this paper, and the interested reader is referred to the review by Tomasi et al. [34].

Supplementary Materials: The following are available online, Table S1: Internal coordinate and backbone angle values for all the tetrapeptides used for characterizing the $i + 2$ Ramachandran map and for proposing the geometric criteria, Figure S1: Superposition of Ramachandran map in the main text on top of traditionally allowed regions of the Ramachandran map for any residue, coordinates of four tetrapeptides (hydrogen-optimised geometries of the four β-turn variant tetrapeptides), Figure S2: GB1 protein showing Loop L1 β-turn. The $i + 1$-$i + 2$ 'KT' segment is engineered to 'vV', 'vS', 'vA' and 'vT' where 'v' refers to D-valine and Table S2: Amino acid sequences of the 22 proteins used in this study.

Data Availability Statement: Data is contained within the article or Supplementary Material.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Pauling, L. The nature of the chemical bond. Application of results obtained from the quantum mechanics and from a theory of paramagnetic susceptibility to the structure of molecules. *J. Am. Chem. Soc.* **1931**, *53*, 1367–1400. [CrossRef]
2. Wintjens, R.; Liévin, J.; Rooman, M.; Buisine, E. Contribution of cation-π interactions to the stability of protein-DNA complexes. *J. Mol. Biol.* **2000**, *302*, 393–408. [CrossRef] [PubMed]
3. Gallivan, J.P.; Dougherty, D.A. Cation-π interactions in structural biology. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9459–9464. [CrossRef] [PubMed]
4. Purich, D.L. Enzyme catalysis: A new definition accounting for noncovalent substrate-and product-like states. *Trends Biochem. Sci.* **2001**, *26*, 417–421. [CrossRef]
5. Cooper, G.M. *The Chemistry of Cells: The Central Role of Enzymes as Biological Catalysts*; Sinauer Associates: Sunderland, MA, USA, 2000.
6. Hudson, K.L.; Bartlett, G.J.; Diehl, R.C.; Agirre, J.; Gallagher, T.; Kiessling, L.L.; Woolfson, D.N. Carbohydrate–aromatic interactions in proteins. *J. Am. Chem. Soc.* **2015**, *137*, 15152–15160. [CrossRef] [PubMed]
7. Spiwok, V. CH/π interactions in carbohydrate recognition. *Molecules* **2017**, *22*, 1038. [CrossRef]
8. Vernon, R.M.; Chong, P.A.; Tsang, B.; Kim, T.H.; Bah, A.; Farber, P.; Lin, H.; Forman-Kay, J.D. Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *eLife* **2018**, *7*, e31486–e31534. [CrossRef] [PubMed]
9. 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]
10. Pauling, L.; Corey, R.B.; Branson, H.R. The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 205–211. [CrossRef] [PubMed]
11. Iwaoka, M.; Takemoto, S.; Tomoda, S. Statistical and theoretical investigations on the directionality of nonbonded S. . .O interactions. Implications for molecular design and protein engineering. *J. Am. Chem. Soc.* **2002**, *124*, 10613–10620. [CrossRef]
12. Burgi, H.; Dunitz, J.; Shefter, E. Geometrical reaction coordinates. II. Nucleophilic addition to a carbonyl group. *J. Am. Chem. Soc.* **1973**, *95*, 5065–5067. [CrossRef]
13. Bürgi, H.; Dunitz, J.; Shefter, E. Chemical reaction paths. IV. Aspects of O. . .C= O interactions in crystals. *Acta Crystall. B-Struct.* **1974**, *30*, 1517–1527. [CrossRef]
14. Sahariah, B.; Sarma, B.K. Relative orientation of the carbonyl groups determines the nature of orbital interactions in carbonyl–carbonyl short contacts. *Chem. Sci.* **2019**, *10*, 909–917. [CrossRef]
15. Singh, S.K.; Kumar, S.; Das, A. Competition between n $\rightarrow$ π Ar* and conventional hydrogen bonding (N–H-N) interactions: An ab initio study of the complexes of 7-azaindole and fluorosubstituted pyridines. *Phys. Chem. Chem. Phys.* **2014**, *16*, 8819–8827. [CrossRef]
16. Bartlett, G.J.; Newberry, R.W.; VanVeller, B.; Raines, R.T.; Woolfson, D.N. Interplay of hydrogen bonds and n $\rightarrow$ π* interactions in proteins. *J. Am. Chem. Soc.* **2013**, *135*, 18682–18688. [CrossRef]
17. Allen, F.; Baalham, C.A.; Lommerse, J.; Raithby, P. Carbonyl–carbonyl interactions can be competitive with hydrogen bonds. *Acta Crystall. B-Struct.* **1998**, *54*, 320–329. [CrossRef]
18. Frontera, A.; Quiñonero, D.; Costa, A.; Ballester, P.; Deyà, P.M. MP2 study of cooperative effects between cation–π, anion–π and π–π interactions. *New J. Chem.* **2007**, *31*, 556–560. [CrossRef]
19. Jeffrey, G.A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin/Heidelberg, Germany, 2012.
20. Newberry, R.W.; Raines, R.T. Secondary forces in protein folding. *ACS Chem. Biol.* **2019**, *14*, 1677–1686. [CrossRef] [PubMed]
21. Ludwig, R. Cooperative hydrogen bonding in amides and peptides. *J. Mol. Liq.* **2000**, *84*, 65–75. [CrossRef]
22. Shishkin, O.V.; Gorb, L.; Leszczynski, J. Modeling of the hydration shell of uracil and thymine. *Int. J. Mol. Sci.* **2000**, *1*, 17–27. [CrossRef]
23. Robertson, A.D.; Murphy, K.P. Protein structure and the energetics of protein stability. *Chem. Rev.* **1997**, *97*, 1251–1268. [CrossRef]
24. Weinhold, F. Nature of H-bonding in clusters, liquids, and enzymes: An ab initio, natural bond orbital perspective. *J. Mol. Struct. Theochem.* **1997**, *398*, 181–197. [CrossRef]
25. Hinderaker, M.P.; Raines, R.T. An electronic effect on protein structure. *Prot. Sci.* **2003**, *12*, 1188–1194. [CrossRef] [PubMed]
26. Bartlett, G.J.; Choudhary, A.; Raines, R.T.; Woolfson, D.N. n $\rightarrow$ π* interactions in proteins. *Nat. Chem. Biol.* **2010**, *6*, 615–620. [CrossRef] [PubMed]
27. Sahariah, B.; Sarma, B.K. Spectroscopic evidence of n $\rightarrow$ π* interactions involving carbonyl groups. *Phys. Chem. Chem. Phys.* **2020**, *22*, 26669–26681. [CrossRef] [PubMed]
28. Deb, P.; Jin, G.Y.; Singh, S.K.; Moon, J.; Kwon, H.; Das, A.; Bagchi, S.; Kim, Y.S. Interconverting Hydrogen-Bonding and Weak n $\rightarrow$ π* Interactions in Aqueous Solution: A Direct Spectroscopic Evidence. *J. Phys. Chem. Lett.* **2018**, *9*, 5425–5429. [CrossRef] [PubMed]

29. Singh, S.K.; Mishra, K.K.; Sharma, N.; Das, A. Direct spectroscopic evidence for an n → π* interaction. *Angew. Chem. Int. Ed.* **2016**, *55*, 7801–7805. [CrossRef] [PubMed]
30. Khatri, B.; Majumder, P.; Nagesh, J.; Penmatsa, A.; Chatterjee, J. Increasing protein stability by engineering the n-→ π* interaction at the β-turn. *Chem. Sci.* **2020**, *11*, 9480–9487. [CrossRef]
31. Chai, J.D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620. [CrossRef]
32. 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 D. 01*; Gaussian Inc.: Wallingford, CT, USA, 2009.
33. Glendening, E.D.; Landis, C.R.; Weinhold, F. NBO 6.0: Natural bond orbital analysis program. *J. Comput. Chem.* **2013**, *34*, 1429–1437; Erratum in **2013**, *34*, 2134. [CrossRef]
34. Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999–3094. [CrossRef] [PubMed]
35. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [CrossRef] [PubMed]
36. Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. *Chem. Rev.* **1994**, *94*, 1887–1930. [CrossRef]
37. Zhou, F.; Liu, R.; Li, P.; Zhang, H. On the properties of S⋯O and S⋯π noncovalent interactions: The analysis of geometry, interaction energy and electron density. *New J. Chem.* **2015**, *39*, 1611–1618. [CrossRef]
38. Jakobsche, C.E.; Choudhary, A.; Miller, S.J.; Raines, R.T. n → π* Interaction and n)(π Pauli repulsion are antagonistic for protein stability. *J. Am. Chem. Soc.* **2010**, *132*, 6651–6653. [CrossRef]
39. Dill, K.A. Dominant forces in protein folding. *Biochemistry* **1990**, *29*, 7133–7155. [CrossRef]
40. Honig, B.; Yang, A.S. Free energy balance in protein folding. *Adv. Protein Chem.* **1995**, *46*, 27–58.
41. Mahadevi, A.S.; Sastry, G.N. Cooperativity in noncovalent interactions. *Chem. Rev.* **2016**, *116*, 2775–2825. [CrossRef]
42. Chawla, M.; Chermak, E.; Zhang, Q.; Bujnicki, J.M.; Oliva, R.; Cavallo, L. Occurrence and stability of lone pair–π stacking interactions between ribose and nucleobases in functional RNAs. *Nucleic Acids Res.* **2017**, *45*, 11019–11032. [CrossRef] [PubMed]
43. Poater, J.; Swart, M.; Bickelhaupt, F.M.; Guerra, C.F. B-DNA structure and stability: The role of hydrogen bonding, π–π stacking interactions, twist-angle, and solvation. *Org. Biomol. Chem.* **2014**, *12*, 4691–4700. [CrossRef]
44. Hunter, C.A. Sequence-dependent DNA structure: The role of base stacking interactions. *J. Mol. Biol.* **1993**, *230*, 1025–1054. [CrossRef] [PubMed]
45. Cooper, V.R.; Thonhauser, T.; Puzder, A.; Schröder, E.; Lundqvist, B.I.; Langreth, D.C. Stacking interactions and the twist of DNA. *J. Am. Chem. Soc.* **2008**, *130*, 1304–1308. [CrossRef] [PubMed]
46. Murthy, A.C.; Dignon, G.L.; Kan, Y.; Zerze, G.H.; Parekh, S.H.; Mittal, J.; Fawzi, N.L. Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nat. Struct. Mol. Biol.* **2019**, *26*, 637–648. [CrossRef]
47. Nott, T.J.; Petsalaki, E.; Farber, P.; Jervis, D.; Fussner, E.; Plochowietz, A.; Craggs, T.D.; Bazett-Jones, D.P.; Pawson, T.; Forman-Kay, J.D.; et al. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **2015**, *57*, 936–947. [CrossRef] [PubMed]
48. Weinhold, F. Natural bond orbital methods. *Encycl. Comput. Chem.* **1998**, *3*, 1792–1811.
49. Bader, R.F.W. A quantum theory of molecular structure and its applications. *Chem. Rev.* **1991**, *91*, 893–928. [CrossRef]
50. Von Hopffgarten, M.; Frenking, G. Energy decomposition analysis. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1345. [CrossRef]
51. Mulliken, R.S. Electronic population analysis on LCAO–MO molecular wave functions. I. *J. Chem. Phys.* **1955**, *23*, 1833–1840. [CrossRef]
52. Mayer, I. Charge, bond order and valence in the AB initio SCF theory. *Chem. Phys. Lett.* **1983**, *97*, 270–274. [CrossRef]
53. Khaliullin, R.Z.; Cobar, E.A.; Lochan, R.C.; Bell, A.T.; Head-Gordon, M. Unravelling the origin of intermolecular interactions using absolutely localized molecular orbitals. *J. Phys. Chem. A* **2007**, *111*, 8753–8765. [CrossRef] [PubMed]
54. Weiner, P.K.; Langridge, R.; Blaney, J.M.; Schaefer, R.; Kollman, P.A. Electrostatic potential molecular surfaces. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 3754–3758. [CrossRef] [PubMed]
55. Carlson, B.; Keller, J.M. Eigenvalues of density matrices. *Phys. Rev.* **1961**, *121*, 659–661. [CrossRef]
56. Szabo, A.; Ostlund, N.S. *Modern Quantum Chemistry: An Introduction to Advanced Electronic Structure Theory*; Dover Publications: Mineola, NY, USA, 1989.
57. Scalmani, G.; Frisch, M.J. Continuous surface charge polarizable continuum models of solvation. I. General formalism. *J. Chem. Phys.* **2010**, *132*, 114110-1–114110-15. [CrossRef] [PubMed]
58. Ponce, M.; van Zon, R.; Northrup, S.; Gruner, D.; Chen, J.; Ertinaz, F.; Fedoseev, A.; Groer, L.; Mao, F.; Mundim, B.C.; et al. Deploying a Top-100 Supercomputer for Large Parallel Workloads: The Niagara Supercomputer. In *Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (Learning) (PEARC'19)*; Association for Computing Machinery: New York, NY, USA, 2019.
59. Loken, C.; Gruner, D.; Groer, L.; Peltier, R.; Bunn, N.; Craig, M.; Henriques, T.; Dempsey, J.; Yu, C.H.; Chen, J.; et al. SciNet: Lessons Learned from Building a Power-efficient Top-20 System and Data Centre. *J. Phys. Conf. Ser.* **2010**, *256*, 012026. [CrossRef]