^{13}C NMR Reveals No Evidence of $n-\pi^*$ Interactions in Proteins

Bradley Worley, Georgia Richard, Gerard S. Harbison, Robert Powers*

Department of Chemistry, University of Nebraska-Lincoln, Lincoln, Nebraska, United States of America

Abstract

An $n-\pi^*$ interaction between neighboring carbonyl groups has been postulated to stabilize protein structures. Such an interaction would affect the ^{13}C chemical shielding of the carbonyl groups, whose paramagnetic component is dominated by $n-\pi^*$ and $\pi-\pi^*$ excitations. Model compound calculations indicate that both the interaction energetics and the chemical shielding of the carbonyl group are instead dominated by a classical dipole-dipole interaction. A set of high-resolution protein structures with associated carbonyl ^{13}C chemical shift assignments verifies this correlation and provides no evidence for an inter-carbonyl $n-\pi^*$ interaction.

Introduction

Proteins exhibit a diversity of structures, with 1,282 folds or topologies present in the CATH database [1]. Each unique structure is defined by its amino acid composition, where sequence identities greater than 40% imply homologous structures [2]. Predicting the three-dimensional conformation of a protein from its primary sequence is a fundamental challenge of structural biology, and achieving this goal requires a thorough understanding of the underlying interactions and forces that stabilize protein structures [3].

Hydrophobic interactions and hydrogen bonds are two of the most common forces attributed to the overall stability of protein structures [4,5]. The burial of hydrophobic residues is generally considered a major driving force in protein folding [6] and has been predicted to contribute roughly 8 kJ/mol per buried residue. Conversely, the contribution of hydrogen bonds to protein structure stability has been controversial [7]. Hydrogen bonds have been described as destabilizing, partially stabilizing or important driving forces. Of course, hydrogen bonds are a defining feature of α-helices, β-sheets and turns. Thus, the generally accepted view is that hydrogen bonds within a protein structure are marginally favored over hydrogen bonds to water. Hydrogen bonds are estimated to contribute roughly 4 kJ/mol to protein stability, but can vary based on the polarity of the microenvironment [8]. Despite these observations, a satisfying general mechanism for protein folding has not been described [9,10]. This implies that our understanding of the factors involved in protein folding and stability is incomplete.

In a recent paper, Bartlett *et. al.* proposed a new and potentially important interaction analogous to the hydrogen bond [11]. Unfortunately, the predicted $n-\pi^*$ interaction was based on density functional theory and a relatively low-level basis set.

However, conventional Kohn-Sham density functional theory does not properly model virtual orbitals [12] such as the π^* orbital proposed by Bartlett *et. al.* to have a role in protein stabilization. Moreover, the relatively low-level basis set used by the authors is inadequate to model such orbitals, and likely gives rise to substantial basis-set superposition errors. Also, experimental data in support of this prediction was not presented. Nevertheless, the predicted $n-\pi^*$ interaction was suggested to aid in the stabilization of protein structures and contribute roughly 0.4 to 5.4 kJ/mol. This stabilization is predicted to occur through the electron delocalization of the lone pair (n) of a carbonyl oxygen atom to the antibonding orbital (π^*) of a neighboring carbonyl carbon atom. An optimal $n-\pi^*$ interaction was predicted to be restricted to a specific range of structural parameters (Figure 1A) corresponding roughly to the Bürgi-Dunitz trajectory [13]. The distance (d) between the donor oxygen and acceptor carbon must be ≤ 3.2 Å, and the angle between the (donor O)···(acceptor C) vector and the acceptor carbonyl vector (θ) must lie between $99°$ and $119°$. Interestingly, the structural parameters required for an optimal $n-\pi^*$ interaction are prevalent in a wide variety of common secondary structures, including α-helices, 3_{10}-helices and twisted β-sheets, suggesting a potential alternative explanation.

Despite the presence of numerous conformations consistent with the $n-\pi^*$ interaction in protein structures, no experimental evidence has been presented that supports the actual existence of this interaction. NMR chemical shifts of sp^2-hybridized groups contain a paramagnetic component caused by mixing of excited states with non-zero orbital angular momentum into the diamagnetic ground state [14]. These excitations are predominantly $n-\pi^*$ and $\pi-\pi^*$ and are therefore highly sensitive to perturbations of these orbitals. The predicted $n-\pi^*$ interactions between neighboring carbonyls would be expected to modify the local electronic environment of the acceptor carbonyl carbon nucleus,
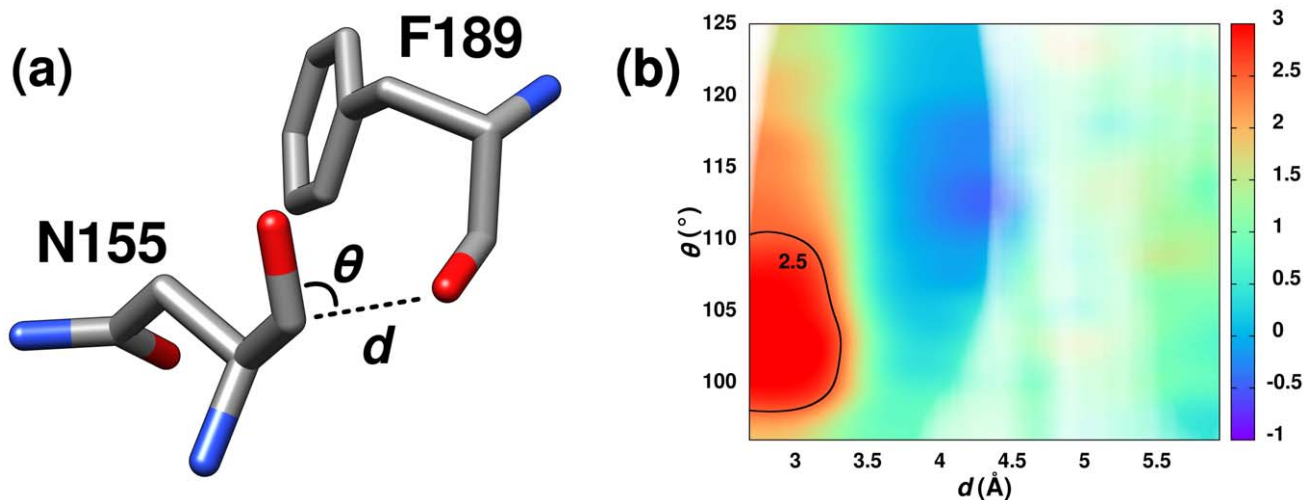
Figure 1. Predicted $n-\pi^*$ Interaction and Associated Carbonyl ^{13}C Chemical Shifts. (**A**) Residues Asn155 and Phe189 from the x-ray structure of *Bacillus amyloliquefaciens* subtillisin BPN' (PDB ID: 1v5i) illustrating the structural features for an optimal $n-\pi^*$ interaction between carbonyl groups. (**B**) 2D contour plot of carbonyl ^{13}C chemical shift differences relative to random coil values as a function of the distance (d) and angle (θ) between carbonyls. A Gaussian smoothing function was applied to the data with Δx and Δy of 0.3 Å and 1.5°, respectively. A transparency mask based on the density of experimental data (see also Figure S1) is overlaid on the contour plot. Regions lacking experimental data are white. Positive values indicate downfield shifts.
doi:10.1371/journal.pone.0042075.g001

and the NMR ^{13}C chemical shift of the acceptor carbonyl carbon would experience a significant chemical shift change in the presence of an $n-\pi^*$ interaction [15]. A large (roughly 20 ppm range) linear relationship has been previously observed between carbonyl ^{13}C chemical shifts and the carbonyl $n-\pi^*$ transition energy [16,17].

An extensive analysis of ^{13}C chemical shifts correlated to high-resolution x-ray structures combined with a detailed analysis of the molecular orbitals of a formamide trimer model complex does not support the postulated $n_\pi-\pi^*$ interaction. In fact, our model indicates that an $n_\pi-\pi^*$ interaction is implausible. Instead, a simple dipole-dipole interaction better explains the observed effects used in support of the $n-\pi^*$ interaction. A prior manuscript by the same authors dismissed the dipole-dipole interaction explanation without elaboration [18].

Results

A total of 2,516,360 residue pairs from a set of 164 high-resolution ($<$1.6 Å) x-ray crystal structures with assigned and uniformly referenced carbonyl ^{13}C chemical shifts were analyzed for potential $n-\pi^*$ interactions. Setting a maximal distance of 6.0 Å between the donor oxygen and acceptor carbon yielded 45,792 potential acceptor carbonyl carbon atoms. The carbonyl ^{13}C chemical shift differences relative to random coil values for each of the 45,792 potential acceptor carbonyls were plotted against the d and θ values for each carbonyl pair (Figure 1B). These chemical shift differences represent the contribution from the local structural environment, and potentially the $n-\pi^*$ interaction. The two-dimensional contour plot indicates a maximal downfield shift of 2.9 ppm centered on the optimal structural parameters predicted for an $n-\pi^*$ interaction.

Of the 45,792 carbonyls, 5,378 had optimal d and θ values for an $n-\pi^*$ interaction and 40,414 were outside this optimal range. The mean carbonyl ^{13}C chemical shift difference for the 40,414 carbonyls labeled as non-interactors is 0.58 $\pm$ 1.98 ppm. In contrast, the mean carbonyl ^{13}C chemical shift difference for the

5,378 interactors is 2.93 $\pm$ 2.41 ppm. A Student's t-test indicates the difference of 2.35 ppm between the two means is statistically significant at the 99.9% confidence level. To address possible errors introduced into the analysis by highly dynamic residues in the x-ray structures, possible carbonyl interactors with B-factors greater than two standard deviations above the mean were omitted from the dataset. In the resultant set of 44,302 potential carbonyl interactors, the 2.33 ppm chemical shift difference was statistically indistinguishable from the original analysis. Similarly, possible interactors predicted at a 95% confidence level to participate in crystal-packing interfaces were also omitted from the dataset. Again, the corresponding set of interactors yielded a chemical shift difference of 2.33 ppm, which is still statistically significant at the 99.9% confidence level.

To address the possiblity that differences between x-ray crystal structures and NMR solution structures may lead to errors in the analysis, a replicate analysis was performed on a set of 137 NMR solution structures corresponding to the same set of ^{13}C and ^{15}N chemical shifts used previously. Structural alignments using MAMMOTH showed a mean rmsd of 1.87 $\pm$ 0.57 Å between the pairs of x-ray and NMR structures. Of the 1,419,547 resulting carbonyl pairs from the NMR structures, 38,534 pairs were found to be potential interactors. Of the carbonyls in that set, 2,510 interactors were found with a mean carbonyl ^{13}C chemical shift difference of 2.84 $\pm$ 1.71 ppm. The remaining 36,024 non-interactors had a mean carbonyl ^{13}C chemical shift difference of 1.02$\pm$2.02 ppm. Again, the 1.82 ppm difference between the two means is statistically significant at the 99.9% confidence level, indicating that differences between x-ray and NMR structures cannot account for the observed downfield ^{13}C chemical shift.

As predicted, a clear correlation is observed between structural regions consistent with an optimal $n-\pi^*$ interaction and a downfield shift of the accepting carbonyl ^{13}C resonance. Interestingly, the potential $n-\pi^*$ interactions were primarily observed between sequential ($|i-j|=1$) carbonyls. Out of the 164 structures and 2,516,360 residue pairs, only four pairs of carbonyls exhibited a through-space ($|i-j|>5$) arrangement consistent with

an optimal $n-\pi^*$ interaction. This result implies any protein stabilization energy obtained from the proposed $n-\pi^*$ interaction is opportunistic, as opposed to a driving force for protein folding. Apparently, the formation of through-space $n-\pi^*$ interactions is simply less favorable than for other interactions, such as hydrogen bonds or salt-bridges. This also implies that the predicted energy of 5.4 kJ/mol for an optimal $n-\pi^*$ interaction is an over-estimate.

In actuality, an $n-\pi^*$ interaction that imparts a stability of 5.4 kJ/mol would likely fix adjacent pairs of carbonyl groups to preferred torsional angles in order to maximize this interaction. Correspondingly, the existence of these highly energetic $n-\pi^*$ interactions would likely be detrimental to properly folding a protein structure. Folding a protein to its native fold would require distorting the majority of carbonyl pairs away from the ideal torsion angles for a proper $n-\pi^*$ interaction. Only 12% (5,378 out of 45,792) of carbonyls from the 164 x-ray structures adopted conformations with optimal d and θ values for an $n-\pi^*$ interaction. As a result, folding every protein structure would incur an initial energetic penalty of nearly 5.4 kJ/mol per carbonyl pair.

A predominant number of the carbonyls consistent with an optimal $n-\pi^*$ interaction and with a downfield shift of roughly 2.5 ppm fall within the typical α-helical region of a Ramachandran plot, where the remaining residues are near the twisted β-sheet region (see Figure S2). Significant chemical shift changes for carbonyl residues within secondary structures are well documented [19]. Previous analyses of structural factors contributing to carbonyl ^{13}C chemical shifts have implicated hydrogen bond formation [20–22] or excluded hydrogen bond formation [23–25], have implicated ϕ, ψ, and χ dihedral angles [24] or have excluded secondary structure parameters [20,23]. Thus, other factors, such as hydrogen bonds or dipole-dipole interactions, may explain the apparent correlation between carbonyl ^{13}C shifts and the optimal d and θ values for an $n-\pi^*$ interaction. This is probable given the association of $n-\pi^*$ interactions with secondary structure elements. The contribution of a dipole-dipole interaction to carbonyl ^{13}C chemical shifts is illustrated in Figure 2. The dipole-dipole potentials were calculated using the high-resolution x-ray structures for each of the 45,792 carbonyl pairs with a maximal distance of 6.0 Å between the donor oxygen and acceptor carbon. While there is significant scatter in the data, there is also a clear trend between a downfield carbonyl ^{13}C chemical shift and an increasing dipole-dipole energy. Importantly, the cluster of acceptor carbonyls in Figure 2 with the largest ^{13}C chemical shift difference (3.15 ± 2.44 ppm) and positive dipole-dipole potentials also conforms to the optimal d and θ values for the predicted $n-\pi^*$ interaction.

The contribution of a hydrogen-bond interaction to the carbonyl ^{13}C chemical shift was similarly evaluated by calculating the shortest oxygen-hydrogen distance (d_{O-H}) for each donor carbonyl. Again, the distances were calculated using the high-resolution x-ray structures for each of the 45,792 carbonyl pairs. A three-dimensional plot comparing the dipole-dipole potentials, oxygen-hydrogen distances, and the associated carbonyl ^{13}C chemical shifts is very revealing. It can be clearly seen from Figure 3 that any contribution from a hydrogen bond to the ^{13}C carbonyl chemical shift is minimal relative to the dipole-dipole contribution. Both the α-helical and β-sheet regions, which obviously contain hydrogen bond interactions, have distinctly different ^{13}C carbonyl chemical shifts. The α-helical region corresponds to a positive dipole-dipole interaction, and correspondingly to a large carbonyl ^{13}C chemical shift difference. Conversely, the β-sheet region has a negative dipole-dipole interaction and a near zero carbonyl ^{13}C chemical shift difference.

These results further indicate a consistency with a dipole-dipole interaction as opposed to the predicted $n-\pi^*$ interaction.

It is important to note that there is a second cluster of carbonyls in Figure 2 with low ^{13}C chemical shifts and negative dipole-dipole potentials that are also consistent with the optimal d and θ values for the predicted $n-\pi^*$ interaction. A visual inspection of the x-ray structures indicates that these carbonyl pairs are actually pointing away from each other and do not form the configuration for an $n-\pi^*$ interaction illustrated in Figure 1A. Clearly, d and θ values alone fail to adequately define the optimal geometry of the dipole-dipole interaction that is apparently responsible for the observed downfield ^{13}C chemical shifts.

To further examine the origin of these effects, quantum chemical calculations were conducted on a model system, a formamide trimer in which molecules 2 and 3 form an approximately planar, head-to-tail hydrogen bonded dimer, and molecule 3 acts as a putative $n-\pi^*$ donor, with the n_π 'donor' oxygen fixed at a distance d which ranges from 2.9 Å and 3.3 Å from the carbonyl carbon of molecule 2, with the $O_3 \cdots C_2$ vector also fixed at angles θ from 70° to 120° from the $C_2 = O_2$ vector. To avoid problems with the use of density functional theory to model virtual orbitals, Möller-Plesset second order perturbation theory (MP2) was instead used, with a substantially larger basis set than in the previous work. The geometry and relevant Hartree-Fock orbitals of the complex used is shown in Figure 4, for $d = 2.9$ Å and $\theta = 100°$. The computed chemical shielding is shown in Figure 5A as a function of d and θ. The shielding decreases monotonically with θ, but, in contrast, the slope of the shielding surface with respect to d changes sign between $\theta = 70°$ and $\theta = 120°$. This shielding surface does not have the geometry expected if the chemical shielding dependencies on θ and d were dominated by an $n_\pi-\pi^*$ interaction, where shielding should be maximum at θ slightly larger than 90° and $d = 2.9$ Å, decreasing rapidly with increasing values of d.

However, the shielding surface does show a remarkable similarity to the dipole-dipole energy between the putative donor and acceptor, as shown in Figure 5B. This energy was computed using a very simple model assuming the electric dipole vector lies along the carbonyl bond for both molecules and has a value of 2.34 D or 7.81×10^{-30} C·m. As can be seen, the dipole energy closely parallels the chemical shielding surface, monotonically increasing with θ and inverting its slope with respect to d as θ increases. This indicates the major influence on the carbonyl ^{13}C chemical shielding is not an $n_\pi-\pi^*$ interaction but rather the electrostatic field from the neighboring carbonyl dipole. The correspondence is not, however, exact: the chemical shielding surface shows a small negative inflection around $\theta = 90°$, which is actually slightly reversed in the dipolar energy plot.

In order to examine whether an $n_\pi-\pi^*$ interaction might be responsible for this inflection, the chemical shielding surface was fit to a function proportional to the dipolar energy, under the assumption the dipole moment vector lies along the $C = O$ bond direction, and the best fit subtracted from the chemical shielding surface (see Figure S3A). The residual shows a minimum at $\theta \sim 95°$, as would be expected for an $n_\pi-\pi^*$ interaction, but the dip does not appear to decrease rapidly as d increases, as an orbital overlap term would. In fact, the residual is slightly larger at $d = 3.3$ Å than at 2.9 Å (1.3 ppm vs. 1.1 ppm).

From the fit of the shielding surface to the estimated dipole interaction energy, with the assumption the magnitude of the electric dipole moment is that of a formamide monomer (3.7 D), a dependence of chemical shielding on field of -190 ppm/a.u. was obtained (1 atomic unit (a.u.) of electric field equals 5.142×10^{11} V/m). Direct calculations of the dependence of the
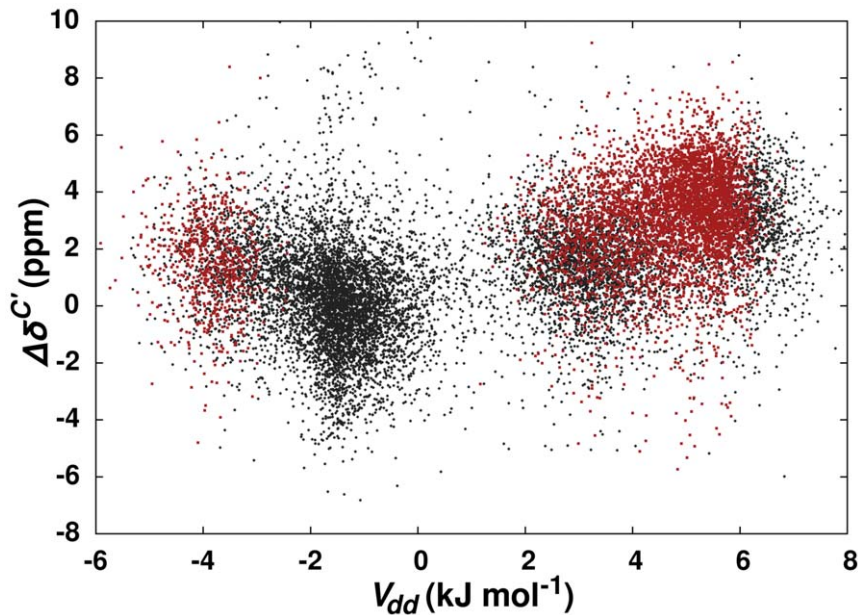
Figure 2. Carbonyl ^{13}C Chemical Shifts and Dipole-Dipole Potential. Carbonyl ^{13}C chemical shift differences relative to random coil are plotted against calculated dipole-dipole potential (V_{dd}). The dipole-dipole potential is calculated from the high-resolution x-ray structure using Equation 1. Pairs of carbonyls with d and θ values within the optimal limits for an $n-\pi^*$ interaction are colored red.
doi:10.1371/journal.pone.0042075.g002

shielding of an isolated formamide on an external applied field along the C = O bond direction gave a value of −150 ppm/a.u. However, it is highly likely that this estimation of the dipole-dipole interaction for two amides is low. Firstly, higher electric multipole terms were neglected in the calculation, and these are likely to be substantial for a moiety as asymmetric as a peptide linkage, at these close proximities. Second, the interaction of the dipoles is likely to be enhanced by the highly polarizable hydrogen bond,

which is necessarily omitted in the monomer model. Agreement of the model with direct estimates of the effect of electric field on shielding is therefore rather good.

The dependence of chemical shielding on hydrogen bonding strength for all combinations of d and θ was examined as a function of the hydrogen bond distance d_{O-N} (see Figure S3B). In accordance with the results of Wishart and others [23–25], and
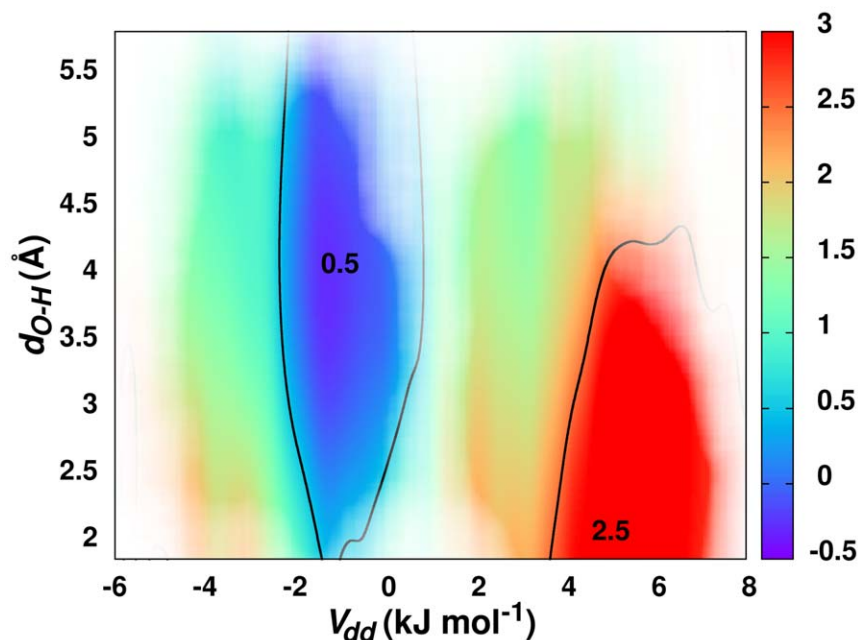

Figure 3. Carbonyl ^{13}C Chemical Shifts and Hydrogen Bonds. Contour plot of ^{13}C carbonyl chemical shift differences as a function of calculated dipole-dipole potential (V_{dd}) and calculated hydrogen bond length (d_{O-H}). See also Figure S2.
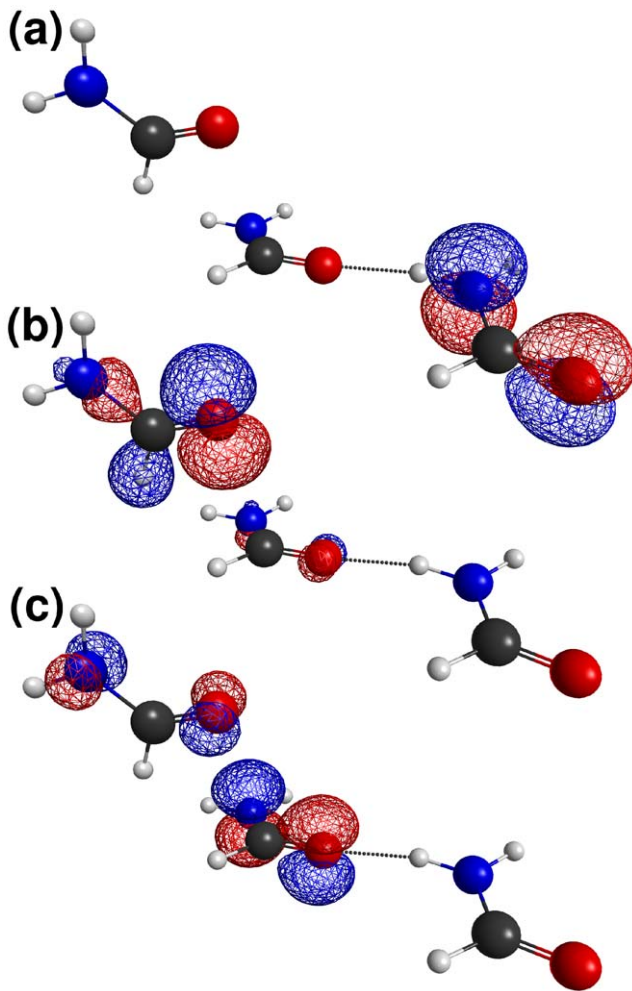doi:10.1371/journal.pone.0042075.g003

Figure 4. Formamide Trimer Model. Molecular orbitals of (**A**) the hydrogen bond donor, (**B**) the putative $n_\pi - \pi^*$ donor and (**C**) the putative $n_\pi - \pi^*$ acceptor, in the trimeric complex used in quantum chemical calculations.
doi:10.1371/journal.pone.0042075.g004

contrary to initial naïve expectations, the effect was very small and independent of the position of the putative $n_\pi - \pi^*$ donor carbonyl.

For the sake of completeness, the effect of an 'end-on' carbonyl-carbonyl interaction was examined, using a dimeric cluster in which the 'donor' carbonyl bond was parallel to the 'donor' oxygen-'acceptor' carbon vector, resulting in a possible $n_\sigma - \pi^*$ interaction. As can be seen in Figure 6, for values of d ranging from 2.9 Å to 4.1 Å, the chemical shielding also follows the negative of the dipolar interaction energy over the range $70° < \theta < 120°$, with little evidence of any effect of orbital overlap on chemical shielding.

One other outcome of the calculations is of possible note. While there was little discernible effect of the proposed $n_\sigma - \pi^*$ or $n_\pi - \pi^*$ interactions on the shielding of the carbonyl carbon or the length of the carbonyl bond, substantial pyramidalization of the amide nitrogen was observed at low values of d and values of θ close to 90°. This would indicate that the primary effect of the 'donor' carbonyl might not be on the carbonyl π bond *per se*, but on its delocalization over the entire amide group. There was also a substantial lengthening of the carbon-nitrogen bond – consistent with a reduced bond order – accompanied by substantial changes in the computed ^{15}N chemical shielding. Thus, while no evidence was found of effects from $n_\sigma - \pi^*$ interactions on the ^{13}C NMR spectroscopy or the energetics of the system, such interactions might be detectable in ^{15}N chemical shifts. Unfortunately, ^{15}N shifts are known to be much more dispersed than carbonyl ^{13}C shifts and are susceptible to a wide range of influences, so disentangling the interaction in real proteins might be a Herculean task.

Discussion

When the molecular orbitals for the trimeric complex are examined in detail, the above results become clear. It is in fact misleading to think of amide groups as being dominated by the carbonyl π bond. The highest occupied molecular orbital (HOMO) of the formamide trimer in fact consists almost entirely of p_z orbitals on the N and O, with wavefunctions of opposite sign. This is depicted in Figure 4A for the Hartree-Fock HOMO of the hydrogen bond donor (energy = −0.377 Ha). The orbital is slightly bonding with respect to the carbonyl, but the carbonyl carbon overall has very little contribution to the molecular orbital.
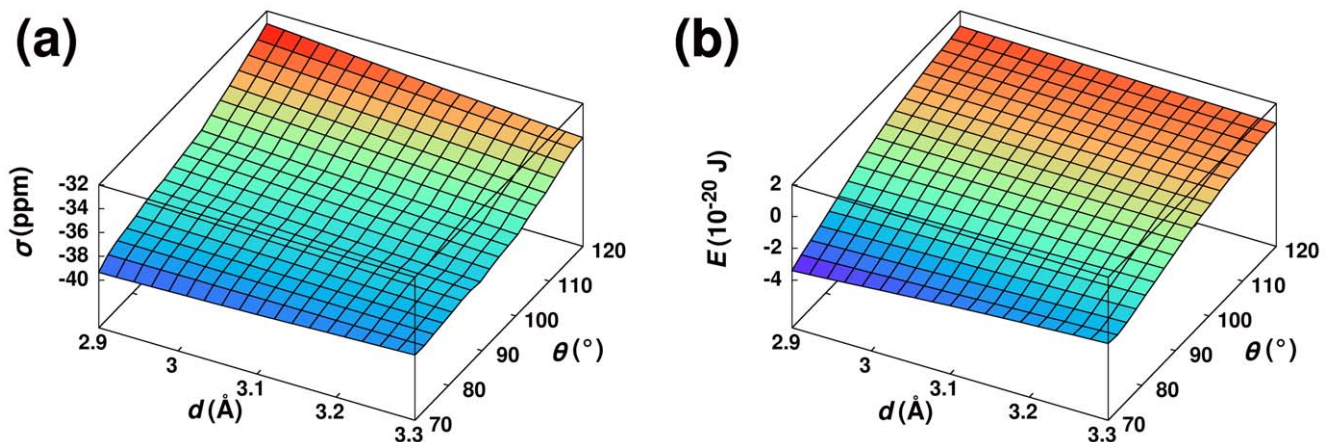

Figure 5. Summary of Quantum Chemical Calculations. Plot of calculated (**A**) carbonyl ^{13}C chemical shielding (σ) and (**B**) dipole-dipole interaction energy (E) as a function of the distance between donor oxygen and acceptor carbon (d) and the angle between carbonyl groups (θ). See also Figure S3.
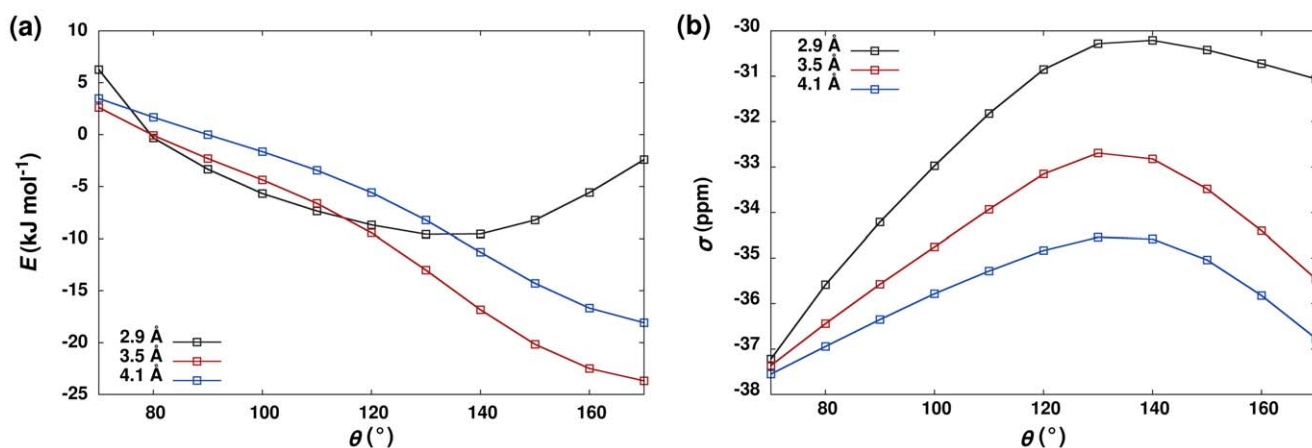doi:10.1371/journal.pone.0042075.g005

Figure 6. Summary of Quantum Chemical Calculations for 'End-On' Dipole Interaction. Plots of (**A**) interaction energy and (**B**) carbonyl ^{13}C chemical shielding (σ) as a function of the angle between the carbonyls (θ) for three different distances (d) between the donor oxygen and acceptor carbon.
doi:10.1371/journal.pone.0042075.g006

The equivalent orbital of the putative acceptor (Figure 4B) has somewhat lower energy (-0.438 Ha) but shows remarkably little mixing with other molecular orbitals, and in particular little mixing with the n_π orbital of the putative $n_\pi - \pi^*$ donor (Figure 4C). That orbital has in fact a very similar energy (-0.465 Ha), and at other geometries – specifically lower values of θ, mixes with the HOMO of the acceptor. The reason for this is quite simple; because the HOMO has only a very small contribution for carbonyl carbon orbitals, bringing the n_π orbital closer to it has very little effect. The mixing that is present at smaller values of θ in fact seems to be partly responsible for the increased pyramidalization of the nitrogen of the acceptor at those orientations. We see no evidence of any orbital mixing that could be attributed to $n_\pi - \pi^*$ interactions. Given the weakness of the mixing with orbitals that are very close in energy to n_π it is implausible that substantial mixing would be observed with an orbital almost a Hartree higher in energy.

In conclusion, quantum chemical calculations, experimental carbonyl ^{13}C chemical shifts and structural data indicate that a simple electrostatic dipole-dipole interaction explains the large downfield carbonyl ^{13}C chemical shift in an α-helix. There is no evidence for a significant contribution from an $n - \pi^*$ interaction to the carbonyl bond. The single indication of $n - \pi^*$ interactions seems to be a substantial lengthening of the carbon-nitrogen bond and pyramidalization of the nitrogen at θ angles favorable for these interactions. In fact, such pyramidalization seems to be a logical consequence of the electronic structure of amides, whose π orbitals are delocalized over the whole system.

Methods

Analysis of Experimental Structures

A detailed statistical analysis was performed to correlate experimentally observed carbonyl ^{13}C chemical shifts with structural parameters between all possible pairs of carbonyls. Specifically, the angle between the carbonyls (θ) and the distance (d) between the oxygen and carbon were compared to experimental carbonyl ^{13}C chemical shifts. The PISCES [26] (http://dunbrack.fcc.edu/pisces) set of 2,885 high-resolution (<1.6 Å) x-ray crystal structures with less than 30% pairwise sequence identity selected from the RCSB Protein Data Bank (PDB) [27] was used for this analysis. Each structure was

associated with assigned NMR ^{13}C and ^{15}N chemical shifts from the Biological Magnetic Resonance Bank (BMRB: www.bmrb.wisc.edu) [28] by FASTA [29] sequence alignments. The best match with an E-value $\leq 1.0 \times 10^{-9}$ and sequence identity $\geq 95\%$ was chosen, where the median E-value was 3.8×10^{-40}. The aligned ^{13}C and ^{15}N chemical shifts were uniformly referenced with the SHIFTCOR software tool [30]. The protein interfaces, surfaces and assemblies software tool (PISA, http://www.ebi.ac.uk/pdbe/prot_int/pistart.html) [31] was used to predict residues involved in crystal packing interfaces. Residues with B-factors two standard deviations from the mean within each structure were identified as dynamic. Also, 3,699 NMR solution structures with PDB depositions cross-linked to the BMRB were used as a validation dataset, with alignments performed in an identical fashion to the analyzed x-ray structures.

A set of Perl scripts was written to extract structural parameters from the x-ray structures. For each structure in the selected set, all pairs of residues were analyzed for the possibility of an $n - \pi^*$ interaction. Values of d and θ were calculated for each residue pair, and torsional angles ϕ and ψ were calculated for the 'acceptor' residue of each pair. Pairs of carbonyls with d and θ values within the optimal limits for an $n - \pi^*$ interaction were labeled as interactors (see Figure S1). Standard random-coil chemical shifts were subtracted from the experimental carbonyl ^{13}C chemical shifts for each residue.

For all pairs of residues, a dipole-dipole potential (V_{dd}) was calculated from the high-resolution x-ray structures using equation 1:

$$V_{dd} = \frac{\vec{\mu}_1 \cdot \vec{\mu}_2 - 3(\vec{\mu}_1 \cdot \hat{r})(\vec{\mu}_2 \cdot \hat{r})}{4\pi\varepsilon_0 |\vec{r}|^3} \qquad (1)$$

where $\vec{\mu}_1$ and $\vec{\mu}_2$ are the two C $=$ O bond vectors, $\vec{r}$ is the vector between the centers of the C $=$ O bonds, and $\hat{r}$ is its unit vector. The nominal value of 2.34 Debye was taken for the carbonyl dipole moment. Similarly, for all residue pairs, the minimum possible hydrogen bond length (d_{O-H}) was calculated from the high-resolution x-ray structures. Hydrogen bond lengths were calculated based on the nearest non-neighboring

backbone amide hydrogen, with a maximal bonding angle of $60°$.

Model Compound Calculations

Quantum chemical calculations were done using the program Gaussian-09 [32]. A nearly planar formamide head-to-tail dimer, composed of a formamide monomer (molecule 1) hydrogen bonded through its $C=O$ group to the N-H group of a second, nearly parallel formamide (molecule 2) was chosen to approximate the hydrogen bonding motif found in both α-helices and antiparallel β-sheets. The dimer was fully optimized at the MP2/6-311++G(2d,p) level; Möller-Plesset second order perturbation theory (MP2) was chosen because it is superior in modeling long-range and dispersive contributions to the electron correlation Hamiltonian. A third formamide (molecule 3) was then added to generate the putative $n−π*$ interaction with molecule 1, imposing the following constraints: (1) $C_3=O_3\cdots C_1$ angle fixed at $90°$, to ensure the n_π orbital of molecule 3 points toward the carbonyl of molecule 1 (2) $O_3\cdots C_1=O_1$ constrained to a set of fixed angles θ, ranging from $70°$ to $120°$ (3) $r_{O3}\cdots C_1$ constrained to a set of fixed distances d, ranging from 2.9 Å to 3.3 Å (5) $r_{O1}\cdots N_2$ constrained to a set of fixed distances, ranging from 2.8 Å to 3.2 Å, to vary the strength of the hydrogen bond. The system of three molecules was then subjected to constrained optimization at the same level as before. The optimized trimolecular system at an angle $\theta=90°$ is shown in Figure 5. Finally, a full set of shielding tensors was computed using standard GIAO methods.

References

1. Knudsen M, Wiuf C (2010) The cath database. Hum Genomics 4: 207–212.
2. Rost B (1999) Twilight zone of protein sequence alignments. Protein Engineering 12: 85–94.
3. Zhang Y (2009) Protein structure prediction: when is it useful? Curr Opin Struct Biol 19: 145–155.
4. Robertson AD, Murphy KP (1997) Protein structure and the energetics of protein stability. Chem Rev 97: 1251–1267.
5. Dill KA (1990) Dominant forces in protein folding. Biochemistry 29: 7133–7155.
6. Kauzmann W (1959) Some factors in the interpretation of protein denaturation. Advances in Protein Chem 14: 1–63.
7. Pace CN (2009) Energetics of protein hydrogen bonds. Nat Struct Mol Biol 16: 681–682.
8. Gao J, Bosco DA, Powers ET, Kelly JW (2009) Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. Nat Struct Mol Biol 16: 684–690.
9. Shakhnovich E (2009) Protein folding roller coaster, one molecule at a time. Proc Natl Acad Sci U S A 106: 11823–11824.
10. Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, et al. (2010) Atomic-level characterization of the structural dynamics of proteins. Science 330: 341–346.
11. Bartlett GJ, Choudhary A, Raines RT, Woolfson DN (2010) n - pi* interactions in proteins. Nat Chem Biol 6: 615–620.
12. Mera H, Stokbro K (2009) Using kohn-sham density functional theory to describe charged excitations in finite systems. Phys Rev B 79: 125109.
13. Burgi HB, Dunitz JD, Shefter E (1973) Geometrical reaction coordinates.2. nucleophilic addition to a carbonyl group. J Am Chem Soc 95: 5065–5067.
14. Ramsey NF (1950) Magnetic shielding of nuclei in molecules. Physical Review 78: 699–703.
15. Abragam A (1961) The Principles of Nuclear Magnetism. Clarendon Press.
16. Savitsky GB, Namikawa K, Zweifel G (1965) A relationship between carbon-13 carbonyl chemical shifts and n-pi transition energies in cyclic and bicyclic ketones. Journal of Physical Chemistry 69: 3105–3109.
17. De JWH (1970) Relation between carbon-13 and oxygen-17 carbonyl chemical shifts for n − n* transition energies. Mol Phys 18: 31–37.
18. Choudhary A, Gandla D, Krow GR, Raines RT (2009) Nature of amide carbonyl-carbonyl interactions in proteins. J Am Chem Soc 131: 7244–7246.
19. Wang Y, Jardetzky O (2002) Probability-based protein secondary structure identification using combined nmr chemical-shift data. Protein Science 11: 852–861.
20. de Dios AC, Pearson JG, Oldfield E (1993) Secondary and tertiary structural effects on protein nmr chemical shifts: an ab initio approach. Science 260: 1491–1496.
21. Asakawa N, Kuroki S, Kurosu H, Ando I, Shoji A, et al. (1992) Hydrogen-bonding effect on c-13 nmr chemical-shifts of l-alanine residue carbonyl carbons of peptides in the solid state. J Am Chem Soc 114: 3261–3265.
22. Wylie BJ, Sperling LJ, Frericks HL, Shah GJ, Franks WT, et al. (2007) Chemical-shift anisotropy measurements of amide and carbonyl resonances in a microcrystalline protein with slow magic-angle spinning nmr spectroscopy. J Am Chem Soc 129: 5318–5319.
23. Cisnetti F, Loth K, Pelupessy P, Bodenhausen G (2004) Determination of chemical shift anisotropy tensors of carbonyl nuclei in proteins through cross-correlated relaxation in nmr. ChemPhysChem 5: 807–814.
24. Neal S, Nip AM, Zhang H, Wishart DS (2003) Rapid and accurate calculation of protein 1h, 13c and 15n chemical shifts. J Biomol NMR 26: 215–240.
25. Markwick PRL, Sattler M (2004) Site-specific variations of carbonyl chemical shift anisotropies in proteins. J Am Chem Soc 126: 11424–11425.
26. Wang G, R L Dunbrack J (2003) Pisces: a protein sequence culling server. Bioinformatics 19: 1589–1591.
27. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The protein data bank. Nucleic Acids Res 28: 235–242.
28. Ulrich EL, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, et al. (2008) Biomagresbank. Nucleic Acids Res 36: D402–D408.
29. Pearson WR (2000) Flexible sequence similarity searching with the fasta3 program package. Methods in Molecular Biology 132: 185–219.
30. Wishart DS, Zhang HY, Neal S (2003) Refdb: A database of uniformly referenced protein chemical shifts. Journal of Biomol NMR 25: 173–195.
31. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. Journal of Molecular Biology 372: 774–797.
32. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, et al. (2009) Gaussian-09, revision a.02.

Supporting Information

Figure S1 *d,θ*-space Analysis of Experimental Structures. Plot of the distance (d) and angle (θ) measured between each of the 45,792 pairs of carbonyls with a potential $n−π*$ interaction. The relative density of points in the occupied d and θ space was used to generate a transparency mask for Figure 1. (TIF)

Figure S2 **Ramachandran-space Analysis of Experimental Structures.** Ramachandran plot of carbonyls with ^{13}C chemical shift differences relative to random coil that are >2.5 ppm. The acceptor carbonyls from each pair of carbonyls with d and θ values within the optimal limits for an $n−π*$ interaction are colored red. (TIF)

Figure S3 **Supplemental Quantum Chemical Calculation Plots.** (**A**) Plot of the residuals for the fit of the chemical shielding surface to a function proportional to the dipole-dipole energy. (**B**) Summary of the quantum chemical calculations of the hydrogen bond contribution to the dipole-dipole interaction; plot of carbonyl ^{13}C chemical shielding (σ) as a function of the hydrogen bond angle (θ) and distance (d_{O-N}). (TIF)

Author Contributions

Conceived and designed the experiments: RP GSH. Performed the experiments: BW GR GSH. Analyzed the data: BW GSH RP. Wrote the paper: BW GSH RP. Designed the software used in analysis: BW.