Reciprocal carbonyl–carbonyl interactions in small molecules and proteins

Abdur Rahim1, Pinaki Saha1, Kunal Kumar Jha1, Nagamani Sukumar1 & Bani Kanta Sarma1

Carbonyl-carbonyl $n\rightarrow\pi^*$ interactions where a lone pair ($n$) of the oxygen atom of a carbonyl group is delocalized over the $\pi^*$ orbital of a nearby carbonyl group have attracted a lot of attention in recent years due to their ability to affect the 3D structure of small molecules, polyesters, peptides, and proteins. In this paper, we report the discovery of a “reciprocal” carbonyl-carbonyl interaction with substantial back and forth $n\rightarrow\pi^*$ and $\pi\rightarrow\pi^*$ electron delocalization between neighboring carbonyl groups. We have carried out experimental studies, analyses of crystallographic databases and theoretical calculations to show the presence of this interaction in both small molecules and proteins. In proteins, these interactions are primarily found in polyproline II (PPII) helices. As PPII are the most abundant secondary structures in unfolded proteins, we propose that these local interactions may have implications in protein folding.
nature effectively uses combinations of weak noncovalent interactions in the functional forms of various biologically important molecules such as nucleic acids and proteins. Intermolecular noncovalent interactions of varying magnitude are also responsible for the existence of different states of matter. Carbonyl-carbonyl ($C\equiv\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\·

Results

Reciprocal carbonyl-carbonyl interactions in N,N'-diacylhydrazines. To test our hypothesis of reciprocal $n\rightarrow\pi^*$ interactions, we have synthesized $N,N'$-diacylhydrazines 1-8 having various substituents on either side of the carbonyl groups (Fig. 2a). In $N,N'$-diacylhydrazines 1-8, the two amide carbonyl groups (CO-I and CO-II; Fig. 2b) are separated by three covalent bonds and 1,5-type $n\rightarrow\pi^*$ interactions are feasible from both sides. We propose that due to the repulsion between the nitrogen lone pairs, the $N,N'$-diacylhydrazines should be nonplanar with the carbonyl groups oriented favorably for reciprocal $n\rightarrow\pi^*$ interactions. Incorporation of electron donating and withdrawing substituents near the carbonyl groups in 1-8 should help us to tune these interactions.

As anticipated, the $N,N'$-diacylhydrazines (1-8) crystallized in nonplanar form with the carbonyl groups oriented almost orthogonal to each other ($C\equiv\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\cdot\·

Fig. 1 Schematic illustration of a one-sided and b reciprocal $n\rightarrow\pi^*$ interactions. Curved dotted arrows indicate $n\rightarrow\pi^*$ interactions.
calculations are also relatively higher for these compounds (Table 1). The NBO orbital overlaps between the oxygen lone pairs (n_\text{O}) and π^*_{\text{C=O}} orbitals in compound 6 are shown in Figs. 2c, d. Note that in compounds 4 and 6–8 where X = CH_2Cl or CH_2Br, d_2 is shorter than d_1 and stronger n→π* interactions from CO-II to CO-I are observed by NBO analysis. This is due to the electron donation from the Cl or Br atom to the σ^*_{\text{C=O}} and π^*_{\text{C=O}} orbitals of CO-II, which increases the donor ability of the CO-II oxygen atom (Supplementary Table 4). Such electron donations from α-halogens to carbonyl groups and their effect on n→π* interactions were previously reported in the literature.36, 37.

Interestingly, the values of \angle O⋯C = O angles \theta_1 and \theta_2 are much smaller (~82°) in compounds 4–8 where reciprocal n→π* interactions are observed than in 1–3 that lack n→π* interactions (Table 1). In fact, the values of \theta_1 and \theta_2 are much smaller than what is expected for one-sided n→π* interactions (\angle O⋯C = O = 109°) reported previously.38. This may be due to the geometrical arrangement required for reciprocal n→π* interactions, which forces \theta_1 and \theta_2 away from the Bürgi-Dunitz trajectory.

Another important signature of n→π* interactions is the pyramidalizaton of the acceptor carbonyl carbon atom measured by parameters \Delta (and \Theta)9, 14, 17, 25, 29. Positive values of \Delta and \Theta indicate pyramidalization of the acceptor carbonyl carbon towards the donor oxygen atom whereas negative values of \Delta and \Theta indicate pyramidalization of the acceptor carbon away from the donor oxygen atom. In compounds 1–8, however, we have not observed a correlation of pyramidalization (\Theta) with O⋯C distance and the strength of n→π* interactions. One reason for this could be the stronger donation from the α-halogen atoms to the nearby carbonyl, which would force the acceptor carbonyl carbons towards the halogen atoms away from the donor oxygen atoms. Also, the crystal packing forces may have some influence in the observed geometries and the pyramidalization of the two nitrogen atoms between the carbonyl groups may influence the pyramidalization of the acceptor carbonyl carbons. Moreover, the individual n→π* interactions in compounds 1–8 may not be strong enough to exert a significant effect on pyramidalization of the carbonyl carbons.

Overall, these data suggest, in compounds 1–8, the geometrical constraints imposed by the repulsion between

---

**Fig. 2** Model compounds synthesized to study reciprocal n→π* interactions. a Chemical structures of N,N'-diacylhydrazines 1–8. b Definition of different structural parameters in N,N'-diacylhydrazines 1–8: d_1 = O_1⋯C_2^2; d_2 = O_2⋯C_1^1; \theta_1 = \angle O_1⋯C_2^2 = \Theta_1; \theta_2 = \angle O_2⋯C_1^1 = \Theta_2. c NBO orbital overlap between oxygen lone pair (n_\text{O}) of CO-I and π^*_{\text{C=O}} of CO-II of compound 6. d NBO orbital overlap between oxygen lone pair (n_\text{O}) of CO-II and π^*_{\text{C=O}} of CO-I of compound 1. e Plot showing correlation between O⋯C distances (d_1 and d_2) in compounds 1–8 [Linear fitting; Pearson correlation coefficient = 0.9906]. f Plot showing correlation between reciprocal n→π* interaction energies ([E_1(n→π*) and E_2(n→π*)]) in compounds 1–8 [Linear fitting; Pearson correlation coefficient = 0.938]. Curved dotted arrows indicate n→π* interactions.
the nitrogen lone pairs orient the two carbonyl groups favorably for reciprocal \( n \rightarrow \pi^* \) interactions. We could tune these interactions by introducing electron donating or withdrawing substituents near the carbonyl groups. Interestingly, we observed that an increase in \( n \rightarrow \pi^* \) interaction from one side also leads to an increase in the \( n \rightarrow \pi^* \) interaction from the other side in compounds 1–8. This correlation suggests that \( n \rightarrow \pi^* \) interactions in these compounds could be synergistic (Figs. 2e, f). For example, shorter \( d_1 \) and higher \( E_1^{n \rightarrow \pi^*} \) values are observed in 4 compared to 3 although 3 and 4 have the same the substituent (4-CH$_2$-Ph) attached to CO-I. Similarly, higher donation from CO-I to CO-II is observed in 6 compared to 5 although 5 and 6 have the same substituent (4-OCH$_2$-Ph) attached to CO-I.

To find out if geometry optimization has any effect on the computed \( n \rightarrow \pi^* \) interactions in comparison to the unrevised X-ray geometries, we also carried out geometry optimizations in compounds 1–8 by freezing the dihedral angles of the side chains involved in reciprocal interactions to their X-ray values and freely optimizing the remaining degrees of freedom (bond lengths, angles, and dihedrals) (Supplementary Fig. 2). We observed that reciprocal \( n \rightarrow \pi^* \) interactions were retained after geometry optimizations but they became slightly weaker than what were observed from the NBO calculations on the crystal geometries (Supplementary Table 5). The coordinates of the optimized geometries of 1–8 are provided in Supplementary Data 1. We also observed that, during gas phase geometry optimization, in absence of any packing and intermolecular forces that are present in the X-ray geometries, the Cl or Br atoms attached to the methylene carbons in 4, 6–8 moved to an anti-periplanar geometry (trans) with respect to the oxygen atom of the nearby carbonyl group (CO-II). This is probably due to higher hyperconjugative delocalization between the halogen lone pairs and carbonyl \( \pi^* \) orbital in the anti-periplanar geometry that would provide more stability to the isolated gas phase molecule. Note that such elongation of carbonyl-carbonyl (O–C) short contacts (weakening of \( n \rightarrow \pi^* \) interactions) in gas phase optimized geometry relative to the X-ray geometries are well known.$^9, 13, 14$

### Reciprocal carbonyl-carbonyl interactions in small organic molecules

To probe whether intramolecular reciprocal C=O···C=O interactions are also present in other small molecules we carried out a CSD search. In our search, we looked for organic molecules having at least two carbonyl groups with intramolecular O=C=C (\( d_1 \)) and O=C=C (\( d_2 \)) distances \( \leq 3.2 \) Å (Supplementary Fig. 3). The search was carried out for cases where the two carbonyl groups are separated by at least three covalent bonds (1.5-type interaction). No restriction was imposed on the \( \angle \text{O} \cdots \text{O} \cdots \text{C} = \text{O} \) (\( \theta_1 \) and \( \theta_2 \)) during the search. The CSD search provided 1432 molecules which fulfilled our search criteria (Supplementary Table 6).

The plots showing the distribution of O–C distances (\( d_1 \) and \( d_2 \)) and O–O···C = O angles (\( \theta_1 \) and \( \theta_2 \)) of all the molecules obtained from the CSD search are shown in Figs. 3a, b, respectively. As can be seen from Fig. 3a, in most cases \( d_1 \) and \( d_2 \) fall in 2.90–3.20 Å range indicating that reciprocal interactions are in general weak. The values of \( \theta_1 \) and \( \theta_2 \) are mainly concentrated in the 70–100° range with majority of the molecules having \( \theta_1 \) and \( \theta_2 \) in the range 80–90°. Interestingly, we also observed similar values for O···C distances (\( d_1 \) and \( d_2 \)) and O–O···C = O angles (\( \theta_1 \) and \( \theta_2 \)) in compounds 4–8 that showed reciprocal \( n \rightarrow \pi^* \) interactions. Therefore, it is quite clear that the O–O···C = O (\( \theta \)) angle deviates significantly from the Bürgi-Dunitz trajectory in reciprocal C=O···C=O short contacts. The \( d_1 \) vs. \( \theta_1 \) and \( d_2 \) vs. \( \theta_2 \) plots (Figs. 3c, d, respectively) show that when the angle of approach of donor oxygen atoms to the acceptor carbonyl C=O bonds deviates from Bürgi-Dunitz trajectory, the O···C distances (\( d_1 \) and \( d_2 \)) increase, suggesting weakening of interactions. NBO analyses of crystal geometries of 30 randomly chosen molecules (Supplementary Fig. 4) having \( d_1 \) and \( d_2 \) \( \leq 3.20 \) Å and covering the range of observed O–O···C = O angles (\( \theta \)) values (70–100°) showed the presence of reciprocal interactions in them (Table 2). The NBO orbital overlaps between the oxygen lone pairs (\( n(O) \) and \( \pi^* \)) in one such molecule (Fig. 3e) (CCDC ref. code: JUHQEK) are shown in Figs. 3f, g.

In most of the molecules obtained from the CSD search, reciprocal C=O···C=O interactions were stabilized by both \( n \rightarrow \pi^* \) and \( \pi \rightarrow \pi^* \) interactions between the carbonyl groups (Table 2 and Supplementary Table 7). We observed substantial C=O···C=O \( \pi \rightarrow \pi^* \) interactions in molecules having \( \theta_1 \) and \( \theta_2 \) values \( \geq 90° \) (Supplementary Table 7). In some cases, \( \pi \rightarrow \pi^* \) interactions are even stronger than \( n \rightarrow \pi^* \) interactions. When \( \theta_1 \) and \( \theta_2 \) values were \( <90° \), \( \pi \rightarrow \pi^* \) interactions were observed for molecules having relatively shorter O···C distances (both \( d_1 \) and \( d_2 \) \( <2.90 \) Å) and stronger \( n \rightarrow \pi^* \) interactions. We propose that although the contribution of individual orbital interaction is small, the overall contribution of two \( n \rightarrow \pi^* \) and two \( \pi \rightarrow \pi^* \) interactions to the stabilization of molecules having reciprocal C=O···C=O interactions could be significant. Based on the NBO calculations at B3LYP/6-311+ G(2d,p) level, we observed that reciprocal C=O···C=O interactions contribute 0.11–3.37 kcal mol$^{-1}$ (with an average value of 0.98 kcal mol$^{-1}$) to the stabilization of small molecules (see the last column in Supplementary Table 7).

We observed positive values of \( \Delta \) and \( \Theta \) for the carbonyl carbons in most of the molecules from the CSD listed in Table 2,
which indicate their pyramidalization towards the donor oxygen atoms. The plots of $\Theta$ with $O\cdots C$ distances and the strength of the reciprocal interactions in compounds obtained from the CSD search are shown in Supplementary Fig. 5. Although the correlation between pyramidality of second carbonyl (CO-II) carbon ($\Theta_2$) and $d_1$ looks better than the correlation between pyramidality of first carbonyl (CO-I) carbon ($\Theta_1$) and $d_2$, the CO-I and CO-II are chosen completely randomly in these molecules. As the pyramidalization also depends on other factors like $\phi$ and the elasticity of the carbonyl group, a strong correlation between pyramidalization and the O$\cdots C$ distance and strength of $n\rightarrow\pi^*$ interactions may not be observed in these molecules having different types of carbonyl groups as well as different $\theta$ values.

To get some insights into the structures of the small molecules having reciprocal C=O$\cdots$C=O interactions, we manually analyzed small molecules from the CSD having 1,5-type reciprocal interactions with both $d_1$ and $d_2 \leq 3.00$ Å. A total of 249 molecules fulfill the above criteria [1, 5-interaction; both $d_1$ and $d_2 \leq 3.00$ Å]. As can be anticipated, the nature of the two atoms/groups between the interacting carbonyl groups plays a key role in keeping the two carbonyl groups non coplanar and provides them the conformation required for reciprocal interactions (Supplementary Table 8). Interestingly, majority of these molecules (117, ~47%) have one heteroatom and one chiral carbon between the two interacting carbonyl pairs, a feature that resembles peptides and proteins.

Fig. 3 X-ray crystallographic data and NBO overlap diagrams for CSD molecules. a Plot showing the distribution of O$\cdots C$ distances ($d_1$ and $d_2$) in molecules obtained from the CSD search. b Plot showing the distribution of $\angle O\cdots C=O$ angles ($\theta_1$ and $\theta_2$) in molecules obtained from the CSD search. c Plot of distance $d_1$ vs. angle $\theta_1$ in molecules obtained from the CSD search. d Plot of distance $d_2$ vs. angle $\theta_2$ in molecules obtained from the CSD search. e Chemical structure of a molecule (CCDC reference code: JUHQEK) obtained from the CSD search. The amide carbonyl group is taken as CO-I and the ester carbonyl group is taken as CO-II here. f NBO orbital overlap between oxygen lone pair ($n_O$) of CO-I and $\pi^*$C=O orbital of CO-II of JUHQEK. g NBO orbital overlap between oxygen lone pair ($n_O$) of CO-II and $\pi^*$C=O orbital of CO-I of JUHQEK. $d_1 = O_2\cdots C_5$; $d_2 = O_6\cdots C_1$; $\theta_1 = \angle O_2\cdots C_5 = O_6$; $\theta_2 = \angle O_6\cdots C_1 = O_2$ (Supplementary Fig. 3). Curved dotted arrows indicate $n\rightarrow\pi^*$ interactions.
Reciprocal carbonyl-carbonyl interactions in proteins. The presence of reciprocal C=O···C=O interactions in the X-ray crystal geometries of small organic molecules inspired us to look for their presence in protein crystal structures. To probe the presence of reciprocal C=O···C=O interactions in proteins, we analyzed a total of 2269 protein crystal structures with resolution ≤ 1.6 Å from the PDB with redundancy (pairwise sequence identity) less than 10%, out of which 2184 showed the presence of reciprocal interactions in them. The PDB protein structures ranked by the number of reciprocal C=O···C=O interactions present in them are included in Supplementary Data 2. For the PDB search, the distance between the carbonyl oxygen of i_th amino acid residue and the carbonyl carbon of (i + 1)th amino acid residue is defined as d_1. The distance between the carbonyl oxygen of (i + 1)th residue and carbonyl carbon of i_th residue is defined as d_2. The corresponding C=O···C=O angles are defined as \( \Theta_1 \) and \( \Theta_2 \), respectively (Supplementary Fig. 6). During the search, both d_1 and d_2 were kept ≤ 3.20 Å but no restriction was imposed on \( \Theta_1 \) and \( \Theta_2 \). The plot of d_1 and d_2 values obtained from the search show that most of them fall in 2.90–3.20 Å range (Fig. 4a). The angles \( \Theta_1 \) and \( \Theta_2 \) (−85 ± 15°) deviates significantly from the Bürgi-Dunitz trajectory (Fig. 4b). These observations are consistent with the trend that was observed for small molecules discussed above. Analyses of d_1 and d_2 for all amino acid residues in all proteins (2184) studied here show that shorter distances d_1 and d_2 ≤ 3.2 Å fall within the tail of the full distribution (Supplementary Fig. 7).

In a previous study, Bartlett et al reported one-sided \( n \to \pi^* \) interactions with \( d \leq 3.20 \) Å and 99° ≤ \( \theta \) ≤ 119°. As we have applied the same distance (\( d \leq 3.20 \) Å) and resolution (\( < 1.6 \) Å) criteria, the reciprocal interactions observed here for angles 99° ≤ \( \theta_1 \), \( \theta_2 \) ≤ 119° would be observed as one-sided \( n \to \pi^* \) interactions by using the criteria of Bartlett et al. As can be seen from Fig. 4b, the distribution of \( \Theta_1 \) and \( \Theta_2 \) in the range of 99°–119° (regions II, III, and IV) is very a small percentage (6.5%) of the total number of reciprocal C=O···C=O interactions that are being reported here. This indicates that reciprocal C=O···C=O interactions are novel and distinct from one-sided \( n \to \pi^* \) interactions reported previously.

NBO analysis of 30 amino acid pairs (Supplementary Fig. 8) with short O···C distances (both d_1 and d_2 ≤ 3.20 Å) that covers the complete range of observed \( \angle \text{O} \to \text{C} \to \text{O} \) angle (70°–110°) clearly showed the presence of reciprocal \( n \to \pi^* \) interactions (Table 3, Figs. 4c, d). Similar to CSD molecules, in proteins also we observed substantial C=O···C=O \( n \to \pi^* \) interactions between the amino acid pairs having \( \Theta_1 \) and \( \Theta_2 \) values >90° (Supplementary Table 9). \( \pi \to \pi^* \) NBO orbital overlap between the two carbonyl groups in an amino acid pair is shown in Figs. 4c, f [Leu-Pro (141–142); [PDB: 2x5o]. For molecules having relatively stronger \( n \to \pi^* \) interactions (both d_1 and d_2 < 2.90 Å), \( \pi \to \pi^* \) interactions were observed for \( \Theta_1 \) and \( \Theta_2 \) values <90° also (Table 3). This indicates that the overall contribution of reciprocal interactions (two \( n \to \pi^* \) and two \( \pi \to \pi^* \) interactions) could be substantial to protein stabilization. Based on the NBO calculations at B3LYP/6-311 + G(2d,p) level, we observed that reciprocal C=O···C=O interactions contribute 0.27–4.41 kcal mol\(^{-1}\) (with an average value of 1.34 kcal mol\(^{-1}\)) to the stabilization of proteins per amino acid pair (see the last column in Supplementary Table 9).
The plot of torsion angles ($\phi$, $\psi$) (Supplementary Fig. 6) of the residue between the two interacting carbonyl groups along with other residues in the proteins show that the reciprocal interactions are mainly concentrated in the polyproline II (PPII), $\beta$-turn and right-twisted $\beta$-strand regions (Fig. 5a). Unlike the one-sided $n\rightarrow\pi^*$ interactions reported previously22, 23 that are abundant in proteins, the abundance of these newly discovered reciprocal $\text{C}=\text{O}$$\cdots$$\text{C}=\text{O}$ interactions is low (~7.2%). Secondary structure analyses using Stride38 show that reciprocal $\text{C}=\text{O}$$\cdots$$\text{C}=\text{O}$ interactions have considerable abundance in random coils (~20%) and turn regions (10%) of proteins but negligible presence in $\alpha$-helices (0.35%) (Table 4). This is in contrast to the one-sided $n\rightarrow\pi^*$ interactions that are most abundant in $\alpha$-helices22, 23. As PPII helix is not included as an independent secondary structure in most secondary structure prediction programs, many PPII helices remain unassigned even though they are present in the experimentally solved structures. We observed that the coil regions having reciprocal $\text{C}=\text{O}$$\cdots$$\text{C}=\text{O}$ interactions are dominated by PPII structures ($\phi, \psi$):$(-75^\circ, 145^\circ)$). We have confirmed this by plotting the $\phi, \psi$ angles of residues in the random coil regions having reciprocal interactions (Fig. 5b). This is not surprising given that PPII conformations are known to dominate coil regions of folded proteins39.

We also manually analyzed 789 reciprocal $\text{C}=\text{O}$$\cdots$$\text{C}=\text{O}$ interactions in 10 proteins having the highest numbers of reciprocal $\text{C}=\text{O}$$\cdots$$\text{C}=\text{O}$ interactions (Supplementary Table 10). In agreement with Stride prediction, manual inspection revealed that reciprocal $\text{C}=\text{O}$$\cdots$$\text{C}=\text{O}$ interactions are mostly present in...
In 41 cases, the reciprocal pairs found in Table 1 show that most of the reciprocal C=O···C=O interactions were present between the interactions of the largest number of reciprocal C=O···C=O interactions in various proteins followed by glutamic acid and serine (Fig. 5c). This trend is different from what was previously observed for one-sided n→π* interactions in α-helices and β-sheets.22 (Pro > Gly > Ala). Analysis of distribution of reciprocal C=O···C=O interactions among the amino acid pairs in various proteins reveals that Pro–Pro is the most abundant pair (Fig. 5d). The 10 most prominent amino acid pairs that participate in reciprocal C=O···C=O interactions, all contain a proline residue (Fig. 5d). These results may be expected given the abundance of reciprocal interactions in PPII regions.

**Table 3 X-ray crystallographic structural and NBO data for amino acid pairs from the PDB**

| Amino acid pair | PDB code | Residues | d₁ (Å) | d₂ (Å) | θ₁ (°) | θ₂ (°) | n→π* (kcal mol⁻¹) | Total n→π* (kcal mol⁻¹) |
|-----------------|----------|----------|--------|--------|--------|--------|-------------------|------------------------|
| Ile-Pro         | 2opc     | 135-136  | 2.675  | 2.768  | 75.6   | 71.8   | 1.75              | 0.74                   |
| Lys-Pro         | 1k3i     | 50-51    | 2.975  | 2.938  | 78.9   | 80.5   | 0.44              | 0.38                   |
| Cys-Pro         | 1gc      | 251-252  | 2.815  | 2.834  | 80.6   | 79.8   | 1.24              | 0.60                   |
| Leu-Pro         | 1g5a     | 379-380  | 2.956  | 2.978  | 80.7   | 79.6   | 0.61              | 0.27                   |
| Ile-Pro         | 1o7i     | 107-108  | 2.986  | 3.082  | 81.6   | 77.8   | 0.44              | 0.19                   |
interactions could play a role in protein folding. Also, turn regions that are stabilized by reciprocal interactions are known to act as nucleation sites for protein folding. Therefore, an open question is how important such reciprocal interactions might be for protein folding.

Nature of reciprocal carbonyl-carbonyl interactions. The nature of C=O···C=O interactions has been debated in the literature. While some consider them n→π* orbital interactions, others believe them to be dipolar in nature. We have so far discussed reciprocal C=O···C=O interactions as n→π* and π→π* orbital interactions because of the following reasons. Firstly, the plots of the n→π* and sum of n→π* and π→π* orbital interaction energies against the O···C distances (d) show a strong correlation (Figs. 7a, b). In Fig. 7a, we have plotted the distances (d1 and d2 values) against the stabilization energies due to n→π* interactions [NBO second order perturbation energies E1(n→π*) and E2(n→π*)] reported in Tables 1–3. The plot suggests that the stabilization energies E1(n→π*) for n→π* interactions decreases with an increase in the O···C (d) in synthetic molecules 1–8, molecules taken from CSD and interacting amino acid pairs obtained from PDB (Tables 1–3). The overall orbital interaction energies (sum of n→π* [E1(n→π*)] and π→π* [E2(π→π*)] interaction energies reported in Tables 1–3) plotted in Fig. 7b also show a similar correlation with O···C (d) distances. These correlations indicate that orbital interaction is the major mechanism for the stabilization of these reciprocal C=O···C=O short contacts. Secondly, we carried out NBO deletion analysis on all the molecules reported in Tables 1–3 (Supplementary Table 13) and observed that deletion of n→π* interactions increases charge on donor oxygen lone pair (nO) and depletes it on acceptor carbonyl π*C=O orbital, which correlate well with the strength of O···C distances (Supplementary Fig. 10a, b). Similarly, deletion of π→π* interactions increases charge on π*C=O orbital of donor carbonyl and depletes it on π*C=O orbital of the acceptor carbonyl (Supplementary Table 14), which also can be correlated to the strength of C=O···C=O short contacts (Supplementary Fig. 11a, b). The overall accumulation of

![Ramachandran plots and analyses of reciprocal interactions in proteins.](image-url)
charges on the acceptor carbonyl $\pi^*_{\mathrm{C=O}}$ orbitals due to donation from the oxygen lone pairs and $\pi_{\mathrm{C=O}}$ orbital of donor carbonyl is shown in Figs. 7c–d, which correlate well with the strength of C–O···C=O short contacts. This also suggests that electron delocalization is a major contributor in reciprocal C=O···C=O interactions. Finally, C=O···C=O torsion angles of the carbonyl groups involved in reciprocal interactions indicate a net zero dipole-dipole interaction eliminating the possibility of these interactions being dipolar in nature. To emphasize this point, in Figs. 7e–f, we have plotted the values of C=O···C=O torsion angles of the 1432 molecules obtained from the CSD search. The torsion angle ($T$) between two dipoles could be used to understand the dipolar nature of interaction between them. As we know, antiparallel ($T \sim 180^\circ$) dipoles attract and parallel dipoles ($T \sim 90^\circ$) have net zero dipolar interaction. In case of reciprocal interaction, the C=O···C=O torsion angles show an orientational preference [C=O···C=O torsion angle falls in 60° to 90° (or −60° to −90°) range] as a consequence of the simultaneous restrictions on $d_1$ and $d_2$ ($\leq 3.2$ Å). However, the values of the C=O···C=O torsion angles ($\sim 90^\circ$) suggest that there would be almost net zero interaction between the dipoles, eliminating the possibility of strong dipolar interactions. Therefore, we conclude that orbital delocalization is the major driving force for the stabilization of reciprocal C=O···C=O interactions. An elaborate energy decomposition analysis may be required for the accurate deconvolution of various factors contributing to the stabilization of reciprocal C=O···C=O short contacts.

We conclude that reciprocal carbonyl-carbonyl interactions exist both in small organic molecules and proteins. However, due to geometrical constraints associated with such interactions, the approach of the donor oxygen atoms to the acceptor carbon atoms deviates significantly from the Bürgi-Dunitz trajectory, and therefore, electron delocalization between the oxygen lone pair ($n_O$) and $\pi^*_{\mathrm{C=O}}$ orbital is weak. This weak donation from the first carbonyl group to the second is compensated by a back donation from the second carbonyl group to the first. In many cases, reciprocal $\pi\rightarrow\pi^*$ interactions were also observed along with reciprocal $n\rightarrow\pi^*$ interactions and their overall contributions to the stabilization of molecules having reciprocal C=O···C=O short contacts could be significant. In proteins, C=O···C=O $n\rightarrow\pi^*$ interactions are present in all types of secondary structures. While one-sided $n\rightarrow\pi^*$ interactions are prevalent in $\alpha$-helices22, 23, reciprocal interactions are abundant in PPII helices and turn regions. Prevalence of reciprocal C=O···C=O interactions in PPII helices and turn regions of proteins suggests a possible role for these interactions in protein folding. Further, the presence of reciprocal C=O···C=O interactions in twisted $\alpha$-helices and distorted PPII helices and turn regions suggests that these interactions could stabilize secondary structures that deviate from their regular geometries. The reciprocal C=O···C=O interactions at the interface of proteins suggest that these interactions could stabilize secondary structures.

Fig. 6 Reciprocal carbonyl-carbonyl interactions in various secondary structures. a PPII-helix; b $\beta$-turn; c Right-twisted $\beta$-strand; d $\alpha$-helix; e interface of $\alpha$-helix and $\beta$-sheet. The Figures are generated by using PyMOL.
helices and β-turns. It would also be interesting to investigate if some non-peptidic fragments obtained from the CSD search having strong reciprocal C=O⋯C=O interactions could be used to stabilize PPII conformation or design peptide-turns. Finally, an energy decomposition analysis would provide better understanding of the forces that contributes to the stabilization of reciprocal C=O⋯C=O interactions.

Methods

Crystallization method. Single crystals of compounds 1-8 were grown by slow evaporation. Various solvent combinations were used to crystallize the compounds either at room temperature or low temperature (4°C). Details of the crystallization conditions are given in Supplementary Table 1.

X-ray crystal structure determination method. Single crystal structures of compound 1-8 were determined by measuring X-ray intensity data. Brucker D8 Venture APEX 342 single crystal home source X-ray diffractometer equipped with CMOS PHOTON 100 detector and Monochromated microfocus sources Mo Kr radiation (λ = 0.71073 Å) were used for data collection in phi (φ) and omega (ω) scan strategy at room temperature (298 K). The data was processed using SAINT11 and absorption correction was done using SADABS43 implemented in APEX 3. For structure solution XSHELL program based on SHELX45 was used. The non-hydrogen atoms were refined anisotropically and located in successive difference Fourier syntheses. The hydrogen atoms were added with riding models and their positions were refined using restraints.

Fig. 7 Delocalization energies, charge redistribution and torsion angles. a Plot of n→π* interaction energies between the interacting carbonyl pairs against crystallographic O⋯C distances (d1 and d2) in molecules shown in Tables 1-3. When the x-axis is d1, E1(n→π*) is plotted in the y-axis and when the x-axis is d2, E2(n→π*) is plotted in the y-axis. The d1, d2, E1(n→π*) and E2(n→π*) values are taken from Tables 1-3. The n→π* interaction energies were computed at B3LYP/6-311+G(2d,p) level of theory. b Plot of overall orbital interaction energy (sum of n→π* and π→π* interaction energies) between the interacting carbonyl pairs against crystallographic O⋯C distances (d1 and d2) in molecules shown in Tables 1-3. When the x-axis is d1, E1(n→π*) + E2(n→π*) is plotted in the y-axis and when the x-axis is d2, E2(n→π*) + E1(n→π*) is plotted in the y-axis. d1, d2, E1(n→π*) and E2(n→π*) values are taken from Tables 1-3. E1(n→π*) and E2(n→π*) values are from Supplementary Table 3, Supplementary Table 7 and Supplementary Table 9. The orbital interaction energies were computed at B3LYP/6-311+G(2d,p) level of theory. c Plot of accumulation of charges on the n' C=O orbital of CO-II due to donation from lone pairs of oxygen and πC=O orbital of CO-I against d1. d Plot of accumulation of charges on the n' C=O orbital of CO-I due to donation from lone pairs of oxygen and πC=O orbital of CO-II against d2. The solid curves in a-d are drawn for convenience. e Histogram plot showing the frequency of the C=O⋯C=O dihedral angles (see Supplementary Fig. 3 for atom numbers) for 1432 molecules obtained from the CSD search. f Histogram plot showing the frequency of the C=O⋯C=O dihedral angles (see Supplementary Fig. 3 for atom numbers) for 1432 molecules obtained from the CSD search.
ARTICLE

NATURE COMMUNICATIONS | DOI: 10.1038/s41467-017-00081-x

and Cl1B was refined using the PART command. Similar ADP restraint SIMU46 and rigid bond restraint DELU46 was applied to stabilize the anisotropic restraint DELU46 instruction was used to restrain the distance to equal. The anisotropic displacement parameter for disordered chlorine atom was fixed using EADP46 constraint.

The Gaussian09 suite of quantum chemistry programs48. The Hartree-Fock (HF)49 and structural data were retrieved from Cambridge Structural Database33 (CSD version 5.21 Nov. 2015) using Conquest47 (version 1.18) program. The fragment charge from the CCDC via www.ccdc.cam.ac.uk/. deposition number CCDC 1486577-1486584. These data can be obtained free of charge from the CCDC via www.ccdc.cam.ac.uk/.

The CSD analysis of this study are available within the paper and its Supplementary Information. The authors declare that the data supporting the conclusions of this study are available within the paper and its Supplementary Information. The authors declare that the data supporting the conclusions of this study are available within the paper and its Supplementary Information. The authors declare that the data supporting the conclusions of this study are available within the paper and its Supplementary Information.

References

1. Strekowski, L. & Wilson, B. Noncovalent interactions with DNA: An overview. Mutat. Res. Fund. Mol. Mech. Mut. 623, 3–13 (2007).
2. Baldwin, R. L. Protein Folding Handbook Ch. 6 (Wiley-VCH Verlag GmbH, 2003).
3. Karabasoff, A. Non-covalent Interactions in Proteins (Imperial College Press, 2006).
4. Hill, J. W., Petrucci, R. H., McCreary, T. W. & Perry, S. S. General Chemistry Ch. 11 (Pearson Prentice Hall, 2005).
5. Burgi, H. B., Dunitz, J. D. & Shafter, E. Chemical reaction paths. IV. Aspects of C=C interactions in crystals. Acta Cryst B30, 1517–1527 (1974).
6. Singh, S. K. & Das, A. The n→* interaction: a rapidly emerging non-covalent interaction. Phys. Chem. Chem. Phys. 17, 9596–9612 (2015).
7. Paulini, R., Müller, K. & Diederich, F. Orthogonal multipolar interactions in structure chemistry and biology. Angew. Chem. Int. Ed. 44, 1788–1805 (2005).
8. Fischer, F. R., Wood, P. A., Allen, F. H. & Diederich, F. Orthogonal dipolar interactions between amide carbonyl groups. Proc. Natl Acad. Sci. USA 105, 17290–17294 (2008).
9. Choudhary, A., Gandla, D., Krow, G. R. & Raines, R. T. Nature of amide carbonyl-carbonyl interactions in proteins. J. Am. Chem. Soc. 131, 7244–7246 (2009).
10. Worley, B. R., Garbison, G. S. & Powers, R. 13C NMR reveals no evidence of n→* interactions in proteins. PlaSo ONE 7, e42075 (2012).
11. Kamer, K. J., Choudhary, A. & Raines, R. T. Intimate interactions with carbonyl groups: dipole–dipole or n→*? J. Org. Chem. 78, 2099–2103 (2013).
12. Choudhary, A., Kamer, K. J. & Raines, R. T. An n→* Interaction in aspirin: implications for structure and reactivity. J. Org. Chem. 76, 7933–7937 (2011).
13. Newberry, R. W. & Raines, R. T. A Key n→* interaction in N-acyl homoserine lactones. ACS Chem. Biol. 9, 880–883 (2014).
14. Newberry, R. W., VanVeller, B., Guzei, I. A. & Raines, R. T. n→* Interactions of amides and thioamides: implications for protein stability. J. Am. Chem. Soc. 135, 7843–7846 (2013).
15. DeRider, M. L. et al. Collagen stability: insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on prolyl ring conformations. J. Am. Chem. Soc. 124, 2497–2505 (2002).
16. Newberry, R. W. & Raines, R. T. n→* interactions in poly(lactic acid) suggest a role in protein folding. Chem. Commun. 49, 7690–7701 (2013).
17. Choudhary, A. & Raines, R. T. Signature of n→* Interaction in α-helices. Protein Sci. 20, 1077–1081 (2011).
18. Gorske, B. C., Bastian, B. L., Geske, G. D. & Blackwell, H. E. Local and tunable n→* Interactions regulate amide isomerism in the peptide backbone. J. Am. Chem. Soc. 129, 8928–8930 (2007).
19. Gorske, B. C., Stringer, J. R., Bastian, B. L., Fowler, S. A. & Blackwell, H. E. New strategies for the design of folded peptides revealed by a survey of noncovalent interactions in model systems. J. Am. Chem. Soc. 131, 16555–16567 (2009).
20. Stringer, J. R., Crapster, J. A., Guzei, I. A. & Blackwell, H. E. Extraordinary Radial probability map of a helical protein. J. Am. Chem. Soc. 133, 15559–15567 (2011).
21. Laursen, J. S., Engel-Andreasen, J., Fristrup, P., Harris, P. & Olsen, C. A. Cis-Trans Amide Bond Rotamers in β-Peptoids and Peptoids: evaluation of Stereoelectronic Effects in Backbone and Side Chains. J. Am. Chem. Soc. 135, 2833–2844 (2013).
22. Bartlett, G. J., Choudhary, A., Raines, R. T. & Woolfson, D. N. n→* interactions in proteins. Nat. Chem. Biol. 6, 615–620 (2010).
23. Fuzean, C. The role of Buerji-Dunitz interactions in the structural stability of proteins. Proteins 78, 2831–2838 (2010).
24. Bartlett, G. J., Newberry, R. W., VanVeller, B., Raines, R. T. & Woolfson, D. N. Investigation of Hydrogen Bonds and n→* Interactions in Proteins. J. Am. Chem. Soc. 135, 18682–18688 (2013).
25. Newberry, R. W., Bartlett, G. J., VanVeller, B., Woolfson, D. N. & Raines, R. T. Signatures of n→* interactions in proteins. Protein Sci. 23, 284–288 (2014).
26. Pal, T. K. & Sankararamakrishnan, R. Quantum chemical investigations on intrareside carbonyl–carbonyl contacts in aspartate of high-resolution protein structures. J. Phys. Chem. B 114, 1038–1049 (2010).
27. Pal, T. K. & Sankararamakrishnan, R. Self-contacts in Axs and Glx residues of high-resolution protein structures: role of local environment and tertiary interactions. J. Mol. Graph. Model. 27, 20–33 (2008).
28. Bondi, A. van der Waals volumes and radii. J. Phys. Chem. 68, 441–451 (1964).
29. Esposito, L., Vitagliano, L., Zagari, A. & Mazzarella, L. Pyramidalization of backbone carbonyl carbon atoms in proteins. Protein Sci. 9, 2038–2042 (2000).
30. Singh, S. K., Mishra, K. K., Sharma, N. & Das, A. Direct spectroscopic evidence for an n→* interaction. Angew. Chem. Int. Ed. 55, 1–6 (2016).
31. Allen, F. H., Baalham, C. A., Lommerse, J. P. M. & Raithby, P. R. Carbonyl–carbonyl interactions can be competitive with hydrogen bonds. Acta Cryst B54, 320–329 (1998).
32. Macallum, P. H., Poet, R. & Müllner-White, E. J. Coulombic attractions between partially charged main-chain atoms stabilise the right-handed twist found in most β-Strands. J. Mol. Biol. 248, 374–384 (1995).
33. Groom, C. R. & Allen, F. H. The Cambridge structural database in retrospect and prospect. Angew. Chem. Int. Ed. 53, 662–671 (2014).
34. Berman, H. M. et al. The protein data bank. Nucleic Acids Res. 28, 235–242 (2000).
35. Glendening, E. D., Reed, A. E., Carpenter, J. E. & Weinhold, F. NBO Version 3.1. Choudhary, A., Fry, C. G. & Raines, R. T. Modulation of an n→* interaction with α-fluoro groups. Arkivoc viii, 251–262 (2010).
36. Cravatt, F., Tschopp, M., Inema, C. F., Freitas, M. P., Birmer, R. & Abraham, R. J. Conformational analysis of 2-haloacyclohexanones: an NMR, theoretical and solution study. J. Chem. Soc. Perkin Trans 2, 1494–1498 (2002).
38. Frishman, D. & Argos, P. Knowledge-based protein secondary structure assignment. Proteins 23, 566–579 (1995).
39. Adzhubei, A. A., Sternberg, M. J. E. & Makarov, A. A. Polyproline-II helix in proteins: structure and function. J. Mol. Biol. 425, 2100–2132 (2013).
40. Shi, Z., Chen, K., Liu, Z. & Kallenbach, N. R. Conformation of the backbone in unfolded proteins. Chem. Rev. 106, 1877–1897 (2006).
41. Levinthal, C. How to fold graciously. in Mossbauer Spectroscopy in Biological Systems: Proceeding of a meeting held at Allerton House (eds De Brunner, J.T.P. & Munck, E.) 22–24, (Monticello, Illinois, University of Illinois Press, 1969).
42. Bruker APEX 3, Version 2, User Manual S86 –EXS053 (Bruker AXS Inc., 2015).
43. Siemens, SMART system (Siemens Analytical X-ray Instruments Inc., 1995).
44. Sheldrick, G. M. SADABS (Bruker AXS Inc., 2007).
45. Sheldrick, G. M. A short history of SHELX. Acta Cryst. Sect. A 64, 112–122 (2008).
46. Muller, P., Herbst-Irmer, R., Spek, A. L., Schneider, T. R. & Sawaya, M. R. Crystal Structure Refinement - A Crystallographer’s Guide to SHELXL (Oxford University Press, 2006).
47. Bruno, I. J. et al. New software for searching the Cambridge Structural Database and visualizing crystal structures. Acta Crystalogr. Sect. B 58, 389–397 (2002).
48. Frisch, M. J. et al. Gaussian 09, Revision D.01 (Gaussian, Inc., 2013).
49. Slater, J. C. A simplification of the Hartree-Fock method. Phys. Rev 81, 385–390 (1951).
50. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys. Rev. B. 37, 785–789 (1988).
51. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. J. Chem. Phys. 98, 5648–5652 (1993).
52. The PyMOL Molecular Graphics System, Version 1.8, Schrödinger, LLC.

Acknowledgements
We thank E. Arunan, M. Goswami, B.B. Dhar, R.P. Vivekananth and R. Ayana for helpful discussions. This project was funded by Shiv Nadar University and Early Career Research Grant (ECR/2015/000337) from Science and Engineering Research Board (SERB), Department of Science and Technology (DST), Government of India.

Author contributions
B.K.S. conceived the project. A.R. carried out the synthesis and characterization of the compounds. A.R. crystallized the compounds, K.K.J. collected the X-ray data and K.K.J. and A.R. solved the structures. B.K.S. and A.R. designed the CSD analyses. A.R. performed the CSD analyses. B.K.S. and P.S. designed the PDB analyses. P.S. performed the PDB analyses. B.K.S. designed the computational studies and A.R. performed them. B.K.S. wrote the manuscript. B.K.S., A.R., N.S., P.S., and K.K.J. discussed the results and edited the manuscript.

Additional information
Supplementary Information accompanies this paper at doi:10.1038/s41467-017-00081-x.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017