Carbon Bonds as Ubiquitous Hydrophobic Interactions in Proteins

V. Rao Mundlapati[1,2], Dipak Kumar Sahoo[1,2], Suman Bhaumik[1,2], Subhrakant Jena[1,2], Amol Chandrakar[1,2], and Himansu S. Biswal[1,2]*

[1]*School of Chemical Sciences, National Institute of Science Education and Research (NISER), PO- Bhimpur-Padanpur, Via-Jatni, District- Khurda, PIN - 752050, Bhubaneswar, India*

[2]*Homi Bhabha National Institute, Training School Complex, Anushakti Nagar, Mumbai 400094, India*

* Corresponding Author's E-mail: himansu@niser.ac.in, Phone No: - +91-674-2494 185/186

Abstract: A genre of noncovalent interactions such as C5-hydrogen bond, halogen bond, reciprocal carbonyl-carbonyl interactions involving carbonyl groups of proteins discovered in recent years are proved to be useful in de novo protein structure and function prediction. However, in proteins, the occurrences and importance of carbon bonds (C-bonds), the highly directional non-covalent interactions between carbonyl-oxygen acceptors and the sp^3-hybridized-carbon σ-hole donors through n→σ* electron delocalization are yet to be conceived. For the first time, we discovered ubiquitous existences of C-bonds in proteins with the help of careful protein structure analysis, quantum calculations, nuclear magnetic resonance and infrared spectroscopy and determined C-bond energies precisely. We demonstrated the implications of C-bonds in explaining photochemistry of oxygen-storage protein myoglobin and protein-DNA interaction. We anticipate inclusion of C-bonds in computational force fields would unravel many more implications of C-bonds in structure, function and dynamics of proteins and protein-ligand/drug complexes.

Significance: The importance of conventional non-covalent interactions such as hydrogen bond (H-bond) and halogen bond (X-bond) in structure and function of biological molecules are well established while carbon bond (C-bond), yet another non-covalent interaction is to be recognized. C-bond is a σ-hole interaction, discovered in the year 2003 in small organic molecules. However, the occurrence, strength and significance of C-bonds in proteins are still unknown. Here by using solution NMR spectroscopy, detail PDB analysis, and quantum calculations, we investigated the C-bonds in proteins and revealed that a significant number of C-bonds are present in proteins, contribute enthalipcally to the overall hydrophobic interaction and play a significant role in photo dissociation mechanism of myoglobin and binding of nucleobase to the protein.

The cumulative contributions from diverse type of several non-covalent interactions govern protein folding and establish their structures[1-8]. Exploring novel non-covalent interactions in biomolecules is fascinating; determining their strength leading to the discovery of many new non-covalent interactions in proteins has remained even compelling and challenging in the last decade[6, 9-14]. It is pertinent to mention few noteworthy breakthroughs as depicted in Fig. S1, highlighting the relevance and significance of this work in bringing out new non-covalent interactions that yet to be conceived by the scientific community in explaining the structures and functions of biomolecules. For the first time in the year 2009, P. Shing Ho et al.[9] showed halogen bonds (X-bonds) can be perpendicular to hydrogen bonds (H-bonds) in protein-ligand complexes. In these interactions both the halogen (X) and hydrogen (H) are non-covalently bonded to the oxygen (O) of the amide carbonyl group (C=O) of peptides. These are considered as orthogonal molecular interactions and termed as hX-bonds. Robert W. Newberry and Ronal T. Raines[11] showed that the close proximity of amide-NH and carbonyl (C=O) oxygen of the same residue can lead to the formation of amide-NH•••O H-bonds, designated as C5 H-bonds. The C5 H-bonds are commonly observed in β-sheets and polyproline II (PPII) secondary structures. In another report Sharma et al.[13] demonstrated reciprocal C=O•••C=O interactions i.e. back and forth n→π* and π→π* interactions in PPII helices. It is also demonstrated that less electronegative element such as selenium of selenomethionine (MSE) is capable of forming strong amide-NH•••Se H-bonds in proteins.[12]

The cited few examples demonstrate that in proteins the C=O group can be engaged in varieties of non-covalent interactions, viz. X-H•••O=C (H-bond), X•••O=C (X-bond), and C=O•••C=O interactions through n→π* electron delocalization[15]. In proteins the possibility of Z-C•••O=C (Z=C, N, O & S) interaction through n→σ* electron delocalization has never been

reckoned and explored. The O atom of C=O can interact with the sp^3-hybridized carbon (C-sp^3) through n→σ* electron delocalization, termed as Z-R_3C•••O=C (R = H & C) interaction or "Carbon bond" (C-bond).

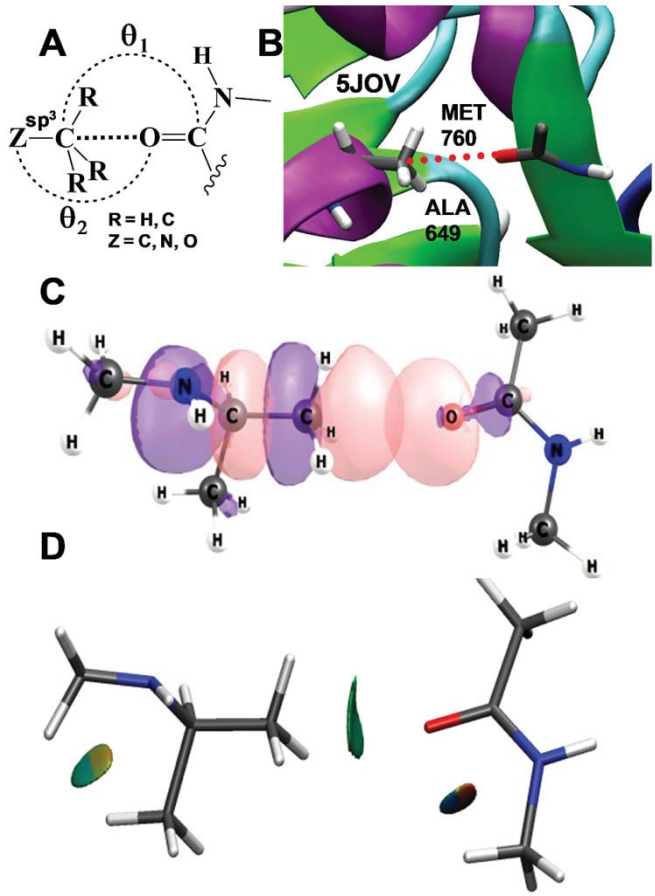

Fig. 1. Carbon bond as a σ-hole interaction in proteins ***(A)*** *Schematic representation of Z-R_3C^{sp3}•••O=C carbon bond.* ***(B)*** *The formation of Z-R_3C^{sp3}•••O=C carbon bond (C-bond) in XyGUL proteins (PDB: 5JOV), representing a highly directional hydrophobic interactions between an electron-rich carbonyl-oxygen acceptor and an electron-deficient sp^3-hybridized carbon σ-hole donor through n→σ* electron delocalization.* ***(C)*** *The $n_O \rightarrow \sigma^*_{C\text{-}C}$ electron delocalization due to the overlap of p-type oxygen lone pair (n) and antibonding C-C (σ*) orbital.* ***(D)*** *Reduced density gradient isosurfaces in real 3D space for Ala^{649}-C•••O=C-Met^{760} C-*

bond, the green isosurface indicates attractive noncovalent interaction between the carbonyl and methyl group.

Based on the quantum chemistry calculations E. Arunan et al.[16] for the first time reported the existences of the C-bonds in small organic molecules. The estimated C-bond energy at the CCSD(T) level can be as much as ~-15 kJ/mol. Later on through the X-ray charge density analysis T. N. Guru Row et al.[17] provided the experimental evidences of C-bonds in natural products and hormones. The C-bond concept was further extended to protein-ligand systems[18] and smaller intermolecular complexes[19-23]. To our surprise we did not come across a single report exploring the C-bonds in proteins that involve C=O groups and sp^3-hybridized carbons (Fig. 1). Owing to the ubiquitous presence of C=O and aliphatic side chains in proteins it is expected that the C^{sp3}•••O=C C-bonds be present in large numbers that would lead to better understanding of the hydrophobic interactions in proteins.

The objective of this work is to scrutinize the protein structures thoroughly to concur C^{sp3}•••O=C C-bonds in proteins, to estimate their enthalpic contributions to the stability of protein structures and to address the role of C-bonds in protein structure and function. Herein, we for the first time provide evidences for C-bonds in proteins with the help of various experimental and computational methods. First, we carried out structural analysis of protein data bank (PDB) to attest that significant numbers of C-bonds are present in proteins. By using several spectroscopic methods on model molecular complexes, we determined C-bond energy very precisely. We then combined quantum chemical calculations and spectroscopic data to formulate an empirical equation that can be adopted to estimate the C-bond energy just by knowing the

C•••O distance. Finally we delved into the adoption of the concept of C-bonds in explaining the biomolecular structures and functions.

Results and Discussion:

Carbon Bonds in Proteins through PDB Structure Analysis: The protein structure coordinates were retrieved from the RCSB website[24] which satisfied the following criteria: structure by X-ray crystallography was at less than 2.0 Å resolution and with less than 30% sequence identity among the proteins. The distance and angle criteria used to search and identify the Z-R_3C•••O=C are (a) $d_{C\cdots O} \leq 3.6$ Å, (b) $160^o \leq \theta_1$ (∠C•••O=C) $\leq 180^o$, (c) (b) $160^o \leq \theta_2$ (∠O•••C-Z) $\leq 180^o$. We found about 7600 interactions satisfying these structural criteria and forming C-bonds. A histogram of C•••O=C C-bond distance distribution was produced using 0.1 Å bar width (Fig. 2A). The shorter distances between O and C in proteins could be attributed to the interpenetration of donor-acceptor orbitals. The donor-acceptor orbital interactions that lead to C-bonds were investigated using natural bond orbital (NBO) analysis. This will be discussed in the next section. Figure 2B shows the histogram of the propensity of different amino acids participating in C-bonds. We found almost all amino acids can contribute to the formations of C-bonds in proteins. Among them ALA, LEU, ASP, and GLU are the potential C-bond acceptors summing up to 40% of total interactions whereas LYS is engaged most frequently as C-bond donors. Figure 2C presents a representative example of C•••O=C interactions in a RAD51 protein (PDB ID: 3NTU). The C=O of ALA^{269} interacts with the C_α of LYS^{261}. The distance between O and C is 2.872 Å and the angles (θ_1 and θ_2) are very close to 170º, suggesting the formation of a strong C-bond of binding energy of -21.7 kJ/mol (*vide infra*). The donor- acceptor

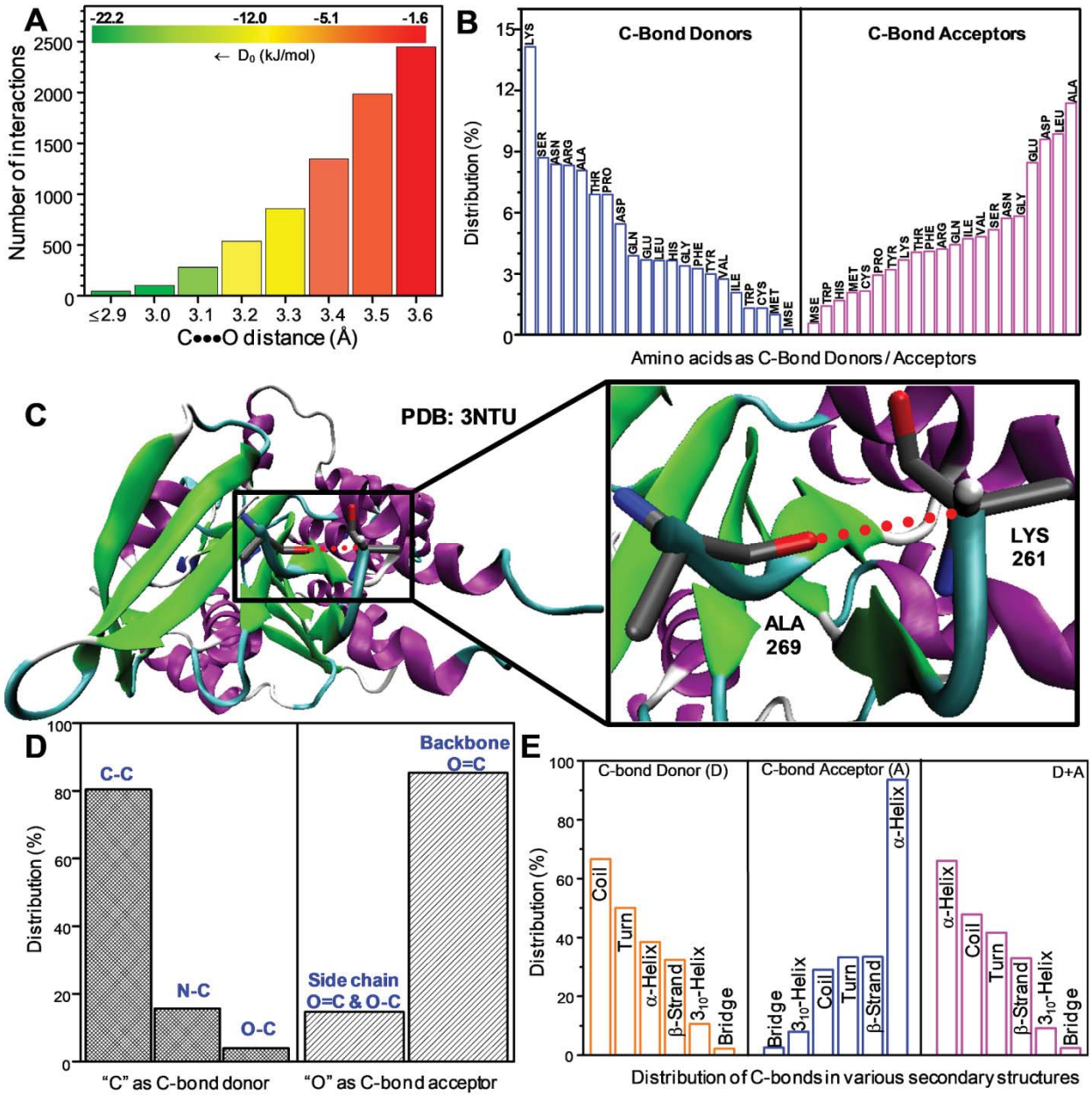

Fig. 2. Carbon bonds in proteins through PDB analysis. ***(A)*** *Frequencies of C•••O carbon bond donor–acceptor distances ($d_{C\cdots O}$) in proteins. Values of the C•••O distances are from protein crystal structures. The histogram represents the distribution of C-bond length within the +.0.4 Å of the sum of the van der Waal's radii of O and C (r_O = 1.5 Å, r_C = 1.7 Å).* ***(B)*** *Plot showing the percentage distribution of amino acids involved in C-bonds.* ***(C)*** *A representative example of Z-*

C•••O=C C-bond observed in RADA recombinase D302K mutant in complex with AMP-PNP (PDB: 3NTU). The C-bond acceptor residue is alanine and the Ala^{269} and the donor residue is Lys^{261}. ***(D)*** *Percentage distribution plot for different type of C-bond acceptors and donors in proteins.* ***(E)*** *Histogram plot showing percentage location of C-bond acceptor and donor residues in various secondary structures of the proteins.*

pair analysis suggests that LYS-ALA, LYS-ASP, and ASN-ALA pairs are the prominent ones (Fig. S2). Figure 2D displays different types of C-bond acceptors and donors involved in C-bond formation in proteins. In all most all the cases, the C-bond acceptors are the backbone or side chain carbonyls whereas in 80.4 % cases C-C bonds are involved as C-bond donor followed by C-N and C-O bonds. Several examples of Z-R_3C•••O=C (R = H & C; Z = C, N, & O) interactions are shown in Supplementary Fig. S3. To explore the possible location of C-bonds in proteins, we made a secondary structure analysis in detail. A histogram of secondary structure assignment and distribution is shown in Fig. 2E. We found C-bond donors are evenly distributed in coils, turns, α-helix and β-strand, whereas ~50% of C-bond acceptors are located in α-helix. The analysis also indicates that the C-bonds are mostly observed in α-helix, coils, turns, summing up to ~70% of C-bonds found in proteins. To ascertain that the C-bonds in proteins were the consequences of n→σ* electron delocalization, we carried out natural bond orbital (NBO) analysis of 22 amino acid pairs covering complete distance (2.75 Å ≤ $d_{C\cdots O}$ ≤ 3.6Å) and angle (160º ≤ θ_i≤ 180º). The donor-acceptor interaction energy (E_{DA}) or the $n_O \rightarrow \sigma^*_{C-Z}$ electron delocalization energy were calculated using second order perturbation theory and are provided in Table S1. There is substantial donor-acceptor orbital overlap i.e. the overlap between the oxygen lone pair orbital (n) with the antibonding sigma orbital of C-Z (σ*). The $n_O \rightarrow \sigma^*_{C-Z}$ interaction energies are very similar to those of two n→π* interaction energies observed through reciprocal

C=O•••C=O interaction.[13] Hence, we would like to draw attention to the fact that one $n_O \rightarrow \sigma^*_{C-Z}$ interaction is equivalent to two $n \rightarrow \pi^*$ interactions, confirming that C-bonds are strong and viable in proteins. Hence their effects on protein structures and functions should not be ignored.

Strength of Carbon Bonds in Proteins: The determination of accurate strength of C-bonds in protein is an insurmountable task owing to the interferences of other non-covalent interactions. The alternative strategy is to choose some model molecular systems that mimic C-bonds in proteins and the C-bond should be the only strongest interaction in those systems. Here we used simple inter-molecular complexes to carry out the experiments in a non-polar solvent so as to get rid of other interactions and to avoid the solvent effect. We chose N,N-dimethylacetamide (NNDMA) as the C-bond acceptor mimicking the amide-C=O group in proteins and several C-bond donors bearing different functional groups. As shown in Fig. 3A the C-bond donors include acetonitrile (ACN), nitromethane (NM), alanine (ALA), acetic acid (AA), acetylchloride (AcetylCl), and trifluoroacetic acid (TFACE). We performed several quantum chemical analyses to theoretically establish the C-bond formation in the model donor-acceptor pairs. Figure 3B displays the interacting natural bond orbitals. The NBO[25] analysis reveals that there is substantial donor-acceptor orbital overlap i.e. the in phase orbital overlap between oxygen lone pair (n_O) of NNDMA and the C-C antibonding sigma orbital (σ^*_{C-C}) of ACN leading to ~3 kJ/mol of $n \rightarrow \sigma^*$ interaction energy. The C-C•••O=C interaction was further supported by Bader's atoms in molecules (AIM) electron density analysis. The molecules presented in Fig. 3C show a bond critical point (BCP) along the C•••O bond vector, inferring accumulation of electron density in between non-covalently bonded O and C atoms forming the C-bond. We also performed Non-Covalent Interaction (NCI) index analysis and molecular electrostatic potential (MESP) calculations to establish that the C•••O=C interactions are attractive and σ-hole interactions.

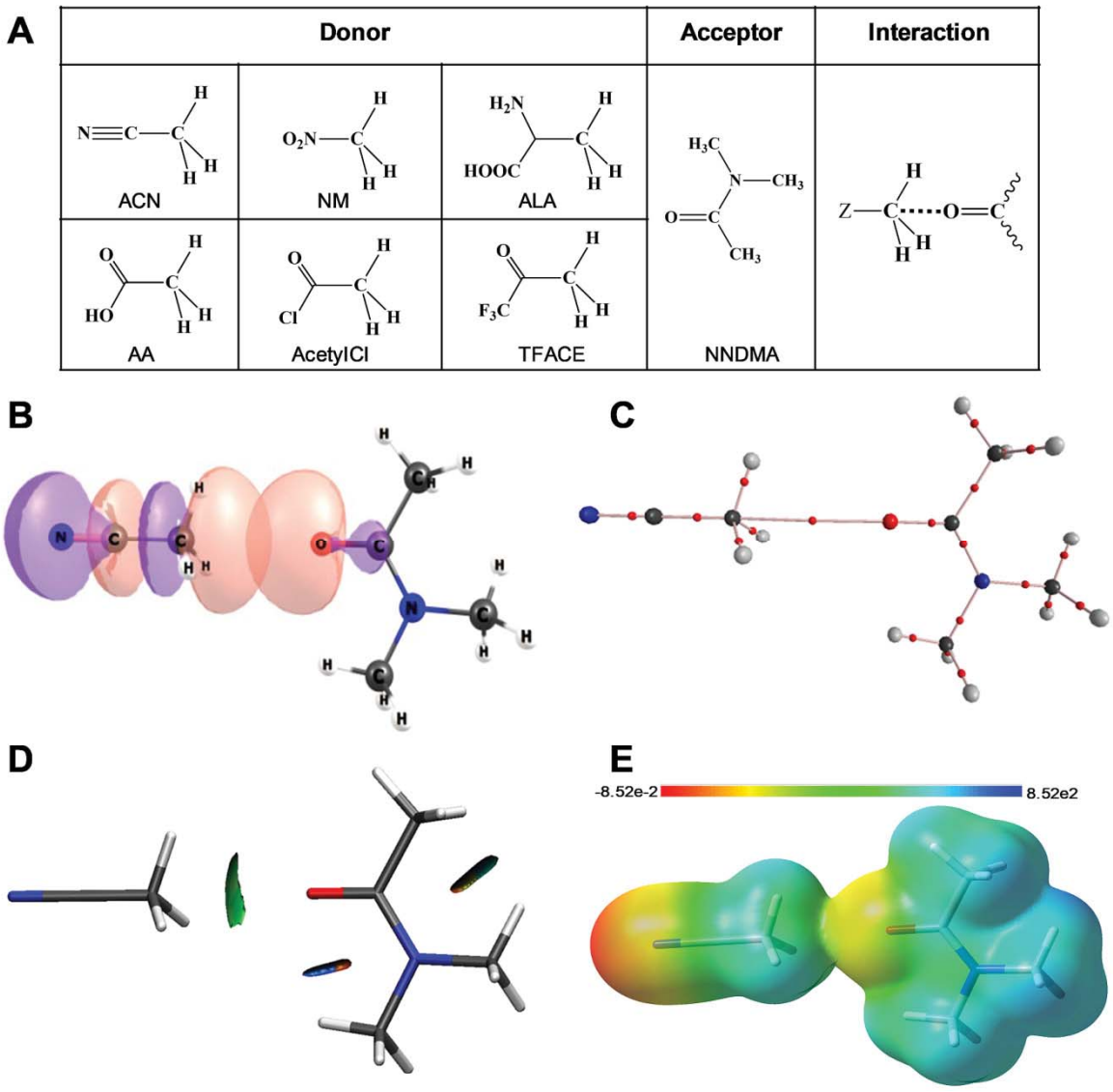

Fig. 3. Model compounds used to mimic C-bonds in proteins. ***(A)*** *Model C-bond donors and acceptors used to mimic C-bonds in proteins* ***left:*** *four different C-bond donors such as acetonitrile (ACN), nitromethane (NM), alanine (ALA), acetic acid (AA), acetylchloride (AcetylCl) and 1,1,1-trifluoroacetone (TFACE),* ***middle:*** *N,N-dimethylacetamide (NNDMA) as carbon bond acceptor,* ***right:*** *schematic representation of C-bond formation between acceptor and donors.* ***(B)*** *The C-C•••O=C C-bond in NNDMA-ACN C-bond characterized by overlap of p-type carbonyl-oxygen lone pair and C-C σ* orbital i.e.* $n_O \rightarrow \sigma^*_{C\text{-}C}$ *electron delocalization.* ***(C)***

Electron density topology analysis and molecular graph for NNDMA-ACN, the presence of bond critical point (shown as small red ball) along the C•••O bond vector confirms the formation of C-bond. ***(D)*** *Reduced density gradient isosurfaces in 3D real space, green isosurface represents the attraction between acceptor and donor.* ***(E)*** *Molecular electrostatic potential showing a σ-hole on C atom of C-donor. The molecular surface is defined by the isovalue of the electron density (0.004 au). Negative values are shown in red and positive values in blue.*

Figure 3D depicts reduced density gradient isosurfaces in real 3D space for C•••O=C C-bond, the blue isosurface indicates attractive noncovalent interaction between NNDMA and ACN. Figure 3E displays MESP contours of NNDMA-ACN complex attesting the σ-hole interaction. We carried out all the aforementioned analyses for the rest of the donor-acceptor pairs and confirmed the existences of C-bonds of variable strengths in these complexes (Supplementary Fig. S4-S8).

From the six complexes we chose NNDMA-ACN complex for the experimental purpose to determine the C-bond energy. The reasoning of the choice are many fold: (a) there are two functional groups i.e. C=O in the acceptor and C≡N in the donor that are very sensitive probe for NMR and IR spectroscopy. From monomer (NNDMA or ACN) to the complex (NNDMA-ACN) formation we can expect changes in ^{13}C chemical shift and in the C=O and C≡N vibrational stretching frequencies. (b) There are no other noncovalent interactions of measurable magnitude present in between these two molecules. (c) Both the molecules are soluble in an inert and non-polar solvent (CCl_4) that helps in avoiding solute-solvent interaction.

NMR titration is a very conventional and useful experimental method that can be exploited to determine the host-guest interaction energy[26]. In these titrations, the host (H) concentration is kept constant and the guest (G) concentration is gradually varied. The NMR

peaks of the host sensitive to the host-guest interactions are monitored. We treated NNDMA as host and ACN as guest in one case and vice-versa in other case. Figure 4B shows the downfield shifts of $^{13}C_{CO}$ of NNDMA with the gradual increase of ACN concentration. Similar shifts of $^{13}C_{CN}$ and $^{13}C_{CH3}$ peaks were observed for ACN while varying NNDMA concentration (Fig. 4D and F). The association constant (K_{eq}) for the host-guest association equilibrium represented as H+G⇋HG at the experimental temperature can be deduced from the chemical shift versus concentration of the guest plot as shown in Figures 4A, C, and E The expression of K_{eq} as a function of guest concentration is provided in the supplementary information. From K_{eq} we calculated the free energy of association (ΔG) and subsequently from experimental free energy (ΔG) and computational entropy (ΔS) we could be able to estimate the C-bond energy $(D_0)_{expt.}$ by using the expression,

$$(D_0)_{expt.} = \Delta G_{expt.} + T\,(\Delta S - R),\ R = \text{universal gas constant.} \qquad (1)$$

The experimental D_0 value for the C-bonded NNDMA-ACN complex averaged for the three experiments is ~-18.6 kJ/mol. We calculated binding energy of NNDMA-ACN complex with a very accurate method like coupled-cluster with single and double and perturbative triple (CCSD(T)) to compare with the experimental value. The $(D_0)_{expt.}$ value obtained at the CCSD(T) *ca*. -18.4 kJ/mol is very close to the experimentally determined D_0 (-18.6 kJ/mol). This suggests that our experimental methodology is reliable to determine the C-bond energy very precisely. The C-bond formation in the NNDMA-ACN is accompanied by the red shifts of C=O and C≡N vibrational stretching frequencies. We observed a linear correlation between the red sifts of the IR frequencies and downfield field shifts of ^{13}C NMR peaks (Fig. S9).

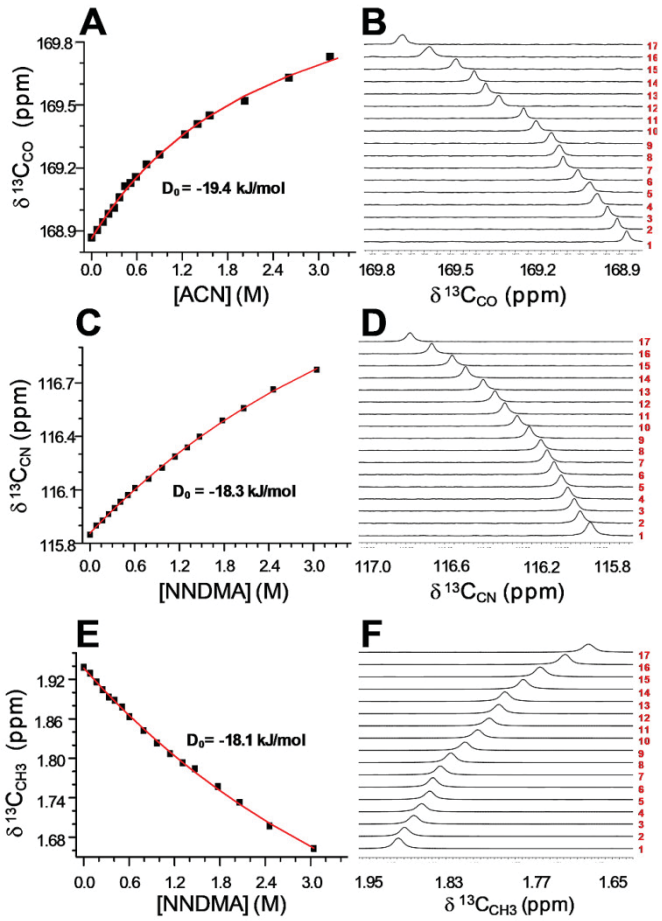

Fig. 4. Determination of C-bond energy by NMR spectroscopy. ***(A)*** *Chemical shift of the* $^{13}C_{CO}$ *peak of NNDMA as a function of concentration of ACN.* ***(B)*** *Stacked* ^{13}C *NMR spectra of carbonyl peak of NNDMA at different concentrations of ACN which shows down field shifts.* ***(C) & (E)*** *Chemical shifts of the* $^{13}C_{CN}$ *peak and* $^{13}C_{CH3}$ *of ACN, respectively as function of concentration of NNDMA.* ***(D) & (F)*** *stacked* ^{13}C *NMR spectra of cyano (CN) and methyl (*CH_3*) peak of ACN at different concentrations of NNDMA. The red line is the fitted curve to determine the equilibrium constant and C-bond energy (see the main text and supplementary information for the fitting equation). The numbers shown right side in the stacked diagrams (B), (D) and (F) represent serial number (S.No.) of titration experiments at different concentrations. The*

concentration ranges are provided in table S2. The ^{13}C NMR spectra were recorded in CCl_4 solvent with DMSO-d_6 as external standard.

After benchmarking the CCSD(T) C-bond energy with the experimental data, we embarked to compute CCSD(T) energetics of the rest of the five C-bond complexes mentioned in Fig. 3. We also calculated n→σ* interaction energy (E_{DA}) for these complexes. We found an exponential correlation between D_0 and E_{DA} as shown in Fig. 5A D_0 and E_{DA} are related as

$$D_0 = 22.2 \times e^{-E_{DA}/1.6} - 22.2 \qquad (2)$$

Since the calculation of E_{DA} is not as computationally expensive as CCSD(T), the above expression will be useful in predicting CCSD(T) binding energy with the prior knowledge of n→σ* interaction energy. We, then exploited this expression in predicting CCSD(T) C-bond energetics in proteins. We computed E_{DA} for 22 amino acid pairs covering complete distance (2.75 Å ≤ $d_{C\bullet\bullet\bullet O}$ ≤ 3.6Å) and angle range (160º ≤ θ_i≤ 180º). The E_{DA} values were plotted against the distance between interacting O and C atoms and the graph is provided in Fig. 5B. An exponential relation between E_{DA} and $d_{C\bullet\bullet\bullet O}$ was deduced as shown in equation 3.

$$E_{DA} = 3.5 \times 10^7 \times e^{-d_{C\bullet\bullet\bullet O}/0.184} \text{ for } 2.75 \text{ Å} \leq d_{C\bullet\bullet\bullet O} \leq 3.6\text{Å} \;\&\; (160^o \leq \theta_i \leq 180^o) \quad (3)$$

By combining equation (2) and (3), we derived another expression as shown below

$$D_0 \approx e^{(\ln(22.2) - 22.2\times10^6 \times (e^{-d_{C\bullet\bullet\bullet O}/0.184}))} - 22.2 \qquad (4)$$

The equation 4 can be used to estimate Z-C•••O=C C-bond energy very precisely just by knowing the C•••O distances. The estimated D_0 of NNDM-ACN C-bond complex by using the

above formula is -17 kJ/mol that is comparable with the experimentally determined binding energy. We are now in a convincing position to use this empirical equation to estimate the C-bond energy in proteins from the C•••O distances. The estimated C-bond energies in proteins are in the range of -2 to -22 kJ/mol. As shown in Fig. 5(c), we devised a C-bond energy protractor that can be used to measure C-bond strengths in proteins.

Possible Role of C-bonds in Structure and Function of Proteins: Since the C-bonds are discovered few years back, we did not find a single report on the role of C-bonds in protein functions. Here we provide two examples where the concept of C-bonds can be evoked to explain many biophysical phenomena. In a recent report[27] entitled "Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation", the authors attribute the ultrafast structural changes in the carbonmonoxy myoglobin complex upon photolysis of the Fe-CO bond to the coupling of the ultrafast heme doming mode to the large scale modes of the proteins through the low frequency vibrations. The low frequency modes of the protein are mostly assigned to the H-bonded residues, e.g. Lys^{98} O–Lys^{42} N_{ξ} distance shows oscillations with a period of 500±150 fs through an N-H•••O=C H-bond. However these H-bonded residues are remotely located from the heme center and the oscillations are transmitted from residues that are directly bonded to the heme center through multiple H-bonds. A careful structure analysis of the CD corner of CO myoglobin revealed that Lys^{42}O is in direct contact with the heme through a C•••O=C C-bond (Fig. 5E). Using our aforementioned empirical relation we estimated the C-bond energy from the C•••O distance and it is found to be ~15.7 kJ/mol.

We performed relaxed potential energy scan as a function $d_{C\cdots O}$ (Fig. 5D) and fitted to Morse potential to deduce intermolecular stretching frequency ($\sigma_{C\cdots O}$) for NNDMA-ACN that has

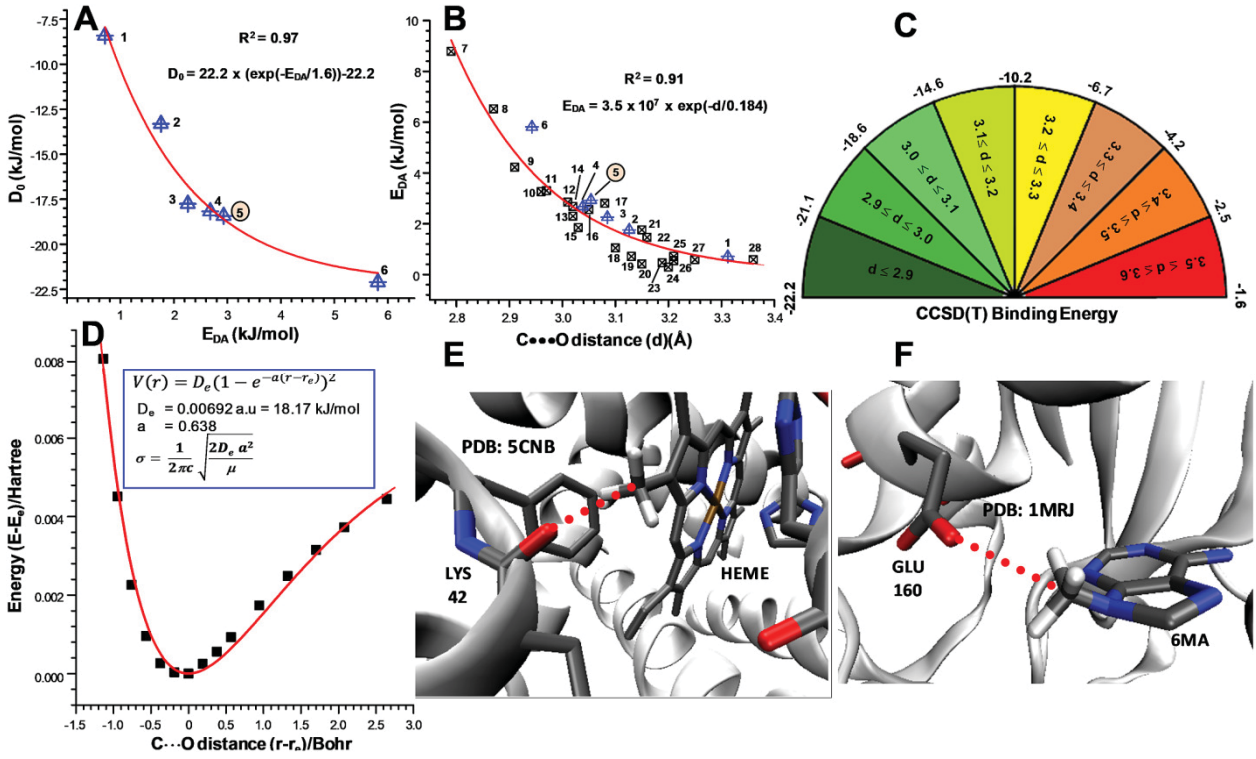

Fig. 5. Strength and Significance of Carbon Bonds in proteins. ***(A)*** *The exponential correlation plot between carbon bond energies (D_0) obtained at CCSD-T level and donor-acceptor interaction energies (E_{DA}) of model C-bond complexes.* ***(B)*** *The exponential correlation plot between the donor-acceptor energies (E_{DA}) of model C-bond complexes and proteins and C•••O bond distances (d), refer the supplementary information for detail of numbering and C-bond donor-acceptor pairs. The NNDMA-ACN complex is highlighted in circle for which NMR experiments were performed to determine the C-bond energy.* ***(C)*** *C-bond energy proposed by combing the two exponential relations displayed in Fig 5(A) and (B). This ruler is useful to estimate the strength of C-bonds in proteins from the C•••O distances.(D) Potential energy scan of energy differences ($E-E_e$) over the change in C···O C-bond distances($r-r_e$). E is the energy for C···O C-bond distance(r) and E_e is the minimum energy for C···O equilibrium C-bond distance (r_e) for* ***C****-bond interaction.* *The solid red line represents Morse potential fitted curve.*

Representative examples of implication of C-bonds in protein structure and function Z-C•••O=C C-bonds observed in ***(E)*** *C-C•••O=C C-bond in myoglobin (PDB: 5CNB) formed by* Lys^{42}*-C=O and Heme-*CH_3*, responsible for the transmission of the heme doming oscillation directly through* Lys^{42} *during CO photolysis in myoglobin to other parts of the protein.* ***(F)*** *N-C•••O=C C-bonds in ribosome-inactivating protein (PDB: 1MRJ) formed by* Glu^{160}*-C=O and* CH_3 *of 6-methyladenine (6MA), contributing significantly to bind 6MA in the active centre of the protein.*

similar $d_{C\cdots O}$ distance as in myoglobin. The $\sigma_{C\cdots O}$ is ~75 cm^{-1} (equivalent to 2.25 THz). The corresponding time period for this C-bond is ~450 fs which is highly in synergy with the ultrafast heme doming mode (417-430 fs)[27]. Hence the heme doming oscillation can easily be transmitted directly through Lys^{42} to other parts of the protein via a C-bond. However it is not explored in the aforementioned article as the C-bond is yet to be recognized as another important non-covalent interaction that can influence the structures and functions of proteins. One more example as shown in Fig. 5F can be cited where 6-methyleadenine (6MA) can be held in the active center of the ribosome-inactivating proteins with a C-bond. In their article Yu Wang and coworkers[28] reported that the ligand (6MA) and protein interaction is due to "hydrophobic forces", aromatic stacking and H-bonds. We observed a C-bond between Glu^{160}-C=O and CH_3-6MA. The estimated C-bond energy (using equation 4) in this case is -21.7 kJ/mol. This C-bond energy is strong enough to be considered as a prominent non-covalent interaction to bind 6MA in the active center of the protein. Hence, the concepts of C-bonds can unequivocally be considered to explain a debatable term frequently used in biomolecules called "hydrophobic forces or hydrophobic interactions". These two are just representative examples. We expect the effects of C-bonds would be explored in many more systems in near future.

Conclusion:

Our data after careful examination of various protein structures suggest that C-bonds are ubiquitous in proteins. NNDMA and ACN in a non interacting solvent (CCl_4) bestow an appropriate opportunity to exploit them as the model C-bond acceptor and donor, respectively to mimic the C-bonds in proteins and to determine their strength experimentally. The concentration dependent NMR experiments enabled us to determine the C-bond energy in NNDMA-ACN complex very precisely ($(D_{0,expt})_{avg}$.= -18.6 kJ/mol) that can be used to bench mark the CCSD(T) binding energy (-18.4 kJ/mol). With the help of experimental and computational data, we proposed a functional to estimate the C-bond energy in proteins from the C•••O distances. We further established that the C-bonds can contribute to the functions and structures of proteins, e. g. the transmisson of the heme doming oscillation in monocarboxy myoglobin can be possible directly through Lys^{42} to other parts of the protein via a C-bond. The C-bond was also found to be strong enough to be considered as a prominent non-covalent interaction to bind 6MA in the active centre of ribosome inactivating proteins.

We hope that this work will encourage further research on the implications of C-bonding in biomolecules. The C-bonds will compliment other non-conventional noncovalent interactions in proteins such as halogen bond, C5 H-bond, reciprocal carbonyl-carbonyl interactions, sulfur and selenium centred H-bonds and they deserve to be included in the new generation force field for bio-molecular structure and function simulation.

Materials and Methods

PDB analysis was carried out with the help of different inbuilt programs written in python language. Protein secondary structure analysis was performed with STRIDE[29, 30]. 1D-^{13}C NMR spectra were acquired from the 400 MHz (9.41 Tesla) Avance-III Bruker liquid state NMR spectrometer for titration experiments. PerkinElmer Frontier™ FT-IR spectrometer was used to record the IR spectra of model C-bond complex and its monomers. All the geometry optimization and frequency calculations were performed using the dispersion corrected DFT functional (B97D) with aug-cc-pVDZ basis set. The C-bond energies were estimated at CCSD(T)/aug-cc-pVDZ level. The donor-acceptor interaction energies were calculated using natural bond orbital (NBO) analysis at MP2/aug-cc-pVDZ level. The non-covalent interaction (NCI)[31] and atoms in molecule (AIM)[32] analyses were done at MP2/aug-cc-pVDZ level. Turbomole 6.5[33], NCI-PLOT[31], Gaussian09[34], and NBO-6.0[35] were used to perform all the computations. Further details of experimental and computational methods used in this study are provided in the supplementary information.

Acknowledgements

The authors thank Prof. A. C. Dash for the stimulating discussion and careful reading of the manuscript. HSB acknowledges financial support from Department of Atomic Energy (DAE), India. *This manuscript is dedicated to the memory of our revered teachers, Prof. Rabindra Kumar Nanda and Prof. Dinabandhu Bhatta who recently passed away.*

References:

1. Crabtree RH (1998) A New Type of Hydrogen Bond. *Science* 282: 2000-2001.

2. Desiraju GR, Steiner T *The Weak Hydrogen Bond: In Structural Chemistry and Biology*. (Oxford University Press, 1999).

3. Auffinger P, Hays FA, Westhof E, Ho PS (2004).Halogen bonds in biological molecules. *Proc. Natl. Acad. Sci. USA* 101: 16789-16794.

4. Kolano C, Helbing J, Kozinski M, Sander W, Hamm P (2006).Watching hydrogen-bond dynamics in a β-turn by transient two-dimensional infrared spectroscopy. *Nature* 444: 469.

5. Joh NH. *et al.* (2008) Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *Nature* 453: 1266.

6. Bachmann A, Wildemann D, Praetorius F, Fischer G, Kiefhaber T (2011) Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction. *Proc. Natl. Acad. Sci. USA* 108: 3952-3957.

7. Goymer P (2012).100 years of the hydrogen bond. *Nat. Chem.* 4: 863.

8. Nisius L, Grzesiek S (2012) Key stabilizing elements of protein structure identified through pressure and temperature perturbation of its hydrogen bond network. *Nat. Chem.* 4: 711.

9. Voth AR, Khuu P, Oishi K, Ho PS (2009) Halogen bonds as orthogonal molecular interactions to hydrogen bonds. *Nat. Chem.* 1: 74.

10. Plevin MJ, Bryce DL, Boisbouvier J (2010) Direct detection of CH/π interactions in proteins. *Nat. Chem.* 2: 466.

11. Newberry RW, Raines RT (2016) A prevalent intraresidue hydrogen bond stabilizes proteins. *Nat. Chem. Biol.* 12: 1084.

12. Mundlapati VR *et al.* (2017) Spectroscopic Evidences for Strong Hydrogen Bonds with Selenomethionine in Proteins. *J. Phys. Chem. Lett.* 8: 794-800..

13. Rahim A, Saha P, Jha KK, Sukumar N, Sarma BK (2017).Reciprocal carbonyl–carbonyl interactions in small molecules and proteins. *Nat. Commun.* 8: 78.

14. Wildemann D *et al.* (2007).A Nearly Isosteric Photosensitive Amide-Backbone Substitution Allows Enzyme Activity Switching in Ribonuclease S. *J. Am. Chem. Soc.* 129: 4910-4918.

15. Newberry RW, Raines RT (2017).The n→π* Interaction. *Acc. Chem. Res.* 50: 1838-1846.

16. Mani D, Arunan E (2013).The X-C···Y (X = O/F, Y = O/S/F/Cl/Br/N/P) 'carbon bond' and hydrophobic interactions. *Phys. Chem. Chem. Phys.* 15:14377-14383.

17. Thomas SP, Pavan MS, Guru Row TN (2014). Experimental evidence for 'carbon bonding' in the solid state from charge density analysis. *ChemComm* 50: 49-51.

18. García-Llinás X, Bauzá A, Seth SK, Frontera A (2017) Importance of R–CF_3···O Tetrel Bonding Interactions in Biological Systems. *J. Phys. Chem. A* 121: 5371-5376.

19. Bauzá A, Frontera A (2016) $RCH_3 \cdots O$ Interactions in Biological Systems: Are They Trifurcated H-Bonds or Noncovalent Carbon Bonds? *Crystals* 6: 26.

20. Bauzá A, Mooibroek TJ, Frontera A (2016) Tetrel Bonding Interactions. *Chem. Rec.* 16: 473-487.

21. Marín-Luna M, Alkorta I, Elguero J (2016) Cooperativity in Tetrel Bonds. *J. Phys. Chem. A* 120: 648-656.

22. Liu M, Li Q, Li W, Cheng J (2017) Carbene tetrel-bonded complexes. *Struct. Chem.* 28: 823-831.

23. Scheiner S Steric (2018) Crowding in Tetrel Bonds. *J. Phys. Chem. A* 122: 2550-2562.

24. Berman HM *et al.* (2000) The Protein Data Bank. *Nucleic Acids Res.* 28: 235-242.

25. Reed AE, Curtiss, LA, Weinhold F (1988) Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* 88: 899-926.

26 Hristova YR, Smulders MMJ, Clegg JK, Breiner B, Nitschke JR (2011).Selective anion binding by a "Chameleon" capsule with a dynamically reconfigurable exterior. Chem. Sci. 2: 638-641.

27. Barends TRM *et al.* (2015).Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science* 350: 445-450.

28. Huang Q, Liu S, Tang Y, Jin S, Wang Y (1995) Studies on crystal structures, active-centre geometry and depurinating mechanism of two ribosome-inactivating proteins. *Biochem. J.* 309: 285-298.

29. Frishman D, Argos P (1995).Knowledge-based protein secondary structure assignment. *Proteins* 23: 566-579.

30. Coe, JV *et al.* (2015) Extracting Infrared Spectra of Protein Secondary Structures Using a Library of Protein Spectra and the Ramachandran Plot. *J. Phys. Chem. B* 119: 13079-13092.

31. Contreras-García J *et al.* (2011).NCIPLOT: A Program for Plotting Noncovalent Interaction Regions. *J. Chem. Theory Comput.* 7: 625-632.

32. Bader RFW *Atoms in Molecules.A quantum theory*. (Oxford University Press, 1990).

33. Furche F *et al.* (2014) Turbomole. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 4: 91-100.

34. Frisch, M. J. *et al.* Gaussian 09, Revision B.01. (2009).

35. Glendening ED, Landis CR, Weinhold F (2013).NBO 6.0: Natural bond orbital analysis program. *J. Comput. Chem.* 34: 1429-1437.

Figure Legends

Fig. 1. Carbon bond as a σ-hole interaction in proteins ***(A)*** *Schematic representation of Z-R_3C^{sp3}•••O=C carbon bond.* ***(B)*** *The formation of Z-R_3C^{sp3}•••O=C carbon bond (C-bond) in XyGUL proteins (PDB: 5JOV), representing a highly directional hydrophobic interactions*

between an electron-rich carbonyl-oxygen acceptor and an electron-deficient sp^3-hybridized carbon σ-hole donor through n→σ electron delocalization.* ***(C)*** *The $n_O \rightarrow \sigma^*_{C\text{-}C}$ electron delocalization due to the overlap of p-type oxygen lone pair (n) and antibonding C-C (σ*) orbital.* ***(D)*** *Reduced density gradient isosurfaces in real 3D space for Ala^{649}-C•••O=C-Met^{760} C-bond, the green isosurface indicates attractive noncovalent interaction between the carbonyl and methyl group.*

Fig. 2. Carbon bonds in proteins through PDB analysis. ***(A)*** *Frequencies of C•••O carbon bond donor–acceptor distances ($d_{C\cdots O}$) in proteins. Values of the C•••O distances are from protein crystal structures. The histogram represents the distribution of C-bond length within the +.0.4 Å of the sum of the van der Waal's radii of O and C (r_O = 1.5 Å, r_C = 1.7 Å).* ***(B)*** *Plot showing the percentage distribution of amino acids involved in C-bonds.* ***(C)*** *A representative example of Z-C•••O=C C-bond observed in RADA recombinase D302K mutant in complex with AMP-PNP (PDB: 3NTU). The C-bond acceptor residue is alanine and the Ala^{269} and the donor residue is Lys^{261}.* ***(D)*** *Percentage distribution plot for different type of C-bond acceptors and donors in proteins.* ***(E)*** *Histogram plot showing percentage location of C-bond acceptor and donor residues in various secondary structures of the proteins.*

Fig. 3. Model compounds used to mimic C-bonds in proteins. ***(A)*** *Model C-bond donors and acceptors used to mimic C-bonds in proteins* ***left:*** *four different C-bond donors such as acetonitrile (ACN), nitromethane (NM), alanine (ALA), acetic acid (AA), acetylchloride (AcetylCl) and 1,1,1-trifluoroacetone (TFACE),* ***middle:*** *N,N-dimethylacetamide (NNDMA) as carbon bond acceptor,* ***right:*** *schematic representation of C-bond formation between acceptor and donors.* ***(B)*** *The C-C•••O=C C-bond in NNDMA-ACN C-bond characterized by overlap of*

p-type carbonyl-oxygen lone pair and C-C σ orbital i.e. $n_O \rightarrow \sigma^*_{C\text{-}C}$ electron delocalization.* ***(C)*** *Electron density topology analysis and molecular graph for NNDMA-ACN, the presence of bond critical point (shown as small red ball) along the C•••O bond vector confirms the formation of C-bond.* ***(D)*** *Reduced density gradient isosurfaces in 3D real space, green isosurface represents the attraction between acceptor and donor.* ***(E)*** *Molecular electrostatic potential showing a σ-hole on C atom of C-donor. The molecular surface is defined by the isovalue of the electron density (0.004 au). Negative values are shown in red and positive values in blue.*

Fig. 4. Determination of C-bond energy by NMR spectroscopy. ***(A)*** *Chemical shift of the $^{13}C_{CO}$ peak of NNDMA as a function of concentration of ACN.* ***(B)*** *Stacked ^{13}C NMR spectra of carbonyl peak of NNDMA at different concentrations of ACN which shows down field shifts.* ***(C) & (E)*** *Chemical shifts of the $^{13}C_{CN}$ peak and $^{13}C_{CH3}$ of ACN, respectively as function of concentration of NNDMA.* ***(D) & (F)*** *stacked ^{13}C NMR spectra of cyano (CN) and methyl (CH_3) peak of ACN at different concentrations of NNDMA. The red line is the fitted curve to determine the equilibrium constant and C-bond energy (see the main text and supplementary information for the fitting equation). The numbers shown right side in the stacked diagrams (B), (D) and (F) represent serial number (S.No.) of titration experiments at different concentrations. The concentration ranges are provided in table S2. The ^{13}C NMR spectra were recorded in CCl_4 solvent with DMSO-d_6 as external standard.*

Fig. 5. Strength and Significance of Carbon Bonds in proteins. ***(A)*** *The exponential correlation plot between carbon bond energies (D_0) obtained at CCSD-T level and donor-acceptor interaction energies (E_{DA}) of model C-bond complexes.* ***(B)*** *The exponential correlation plot between the donor-acceptor energies (E_{DA}) of model C-bond complexes and proteins and C•••O bond distances (d), refer the supplementary information for detail of numbering and C-bond donor-acceptor pairs. The NNDMA-ACN complex is highlighted in circle for which NMR experiments were performed to determine the C-bond energy.* ***(C)*** *C-bond energy proposed by combing the two exponential relations displayed in Fig 5(A) and (B). This ruler is useful to estimate the strength of C-bonds in proteins from the C•••O distances.(D) Potential energy scan of energy differences ($E\text{-}E_e$) over the change in C···O C-bond distances($r\text{-}r_e$). E is the energy for C···O C-bond distance(r) and E_e is the minimum energy for C···O equilibrium C-bond distance (r_e) for* ***C****-bond interaction****.*** *The solid red line represents Morse potential fitted curve.*

Representative examples of implication of C-bonds in protein structure and function Z-C•••O=C C-bonds observed in ***(E)*** *C-C•••O=C C-bond in myoglobin (PDB: 5CNB) formed by* Lys^{42}*-C=O and Heme-*CH_3*, responsible for the transmission of the heme doming oscillation directly through* Lys^{42} *during CO photolysis in myoglobin to other parts of the protein.* ***(F)*** *N-C•••O=C C-bonds in ribosome-inactivating protein (PDB: 1MRJ) formed by* Glu^{160}*-C=O and* CH_3 *of 6-methyladenine (6MA), contributing significantly to bind 6MA in the active centre of the protein.*

Supplemental information (SI) for

Carbon Bonds as Ubiquitous Hydrophobic Interactions in Proteins

V. Rao Mundlapati[1,2], Dipak Kumar Sahoo[1,2], Suman Bhaumik[1,2], Subhrakant Jena[1,2], Amol Chandrakar[1,2], and Himansu S. Biswal[1,2]*

[1]*School of Chemical Sciences, National Institute of Science Education and Research (NISER), PO- Bhimpur-Padanpur, Via-Jatni, District- Khurda, PIN - 752050, Bhubaneswar, India*

[2]*Homi Bhabha National Institute, Training School Complex, Anushakti Nagar, Mumbai 400094, India*

* Corresponding Author's E-mail: himansu@niser.ac.in, Phone No: - +91-674-2494 185/186

Table of Contents

Methods and Materials:

Protein data Bank (PDB) Analysis:

The PDB analyses were carried out by retrieving protein structure coordinates from the RCSB website[1] which satisfy the following criteria: structure resolved by X-ray crystallography at less than 2.0 Å resolution and less than 30% sequence identity among the proteins. With these criteria, we could retrieve 11501 protein structures (Download date: 30th January 2018). In each of the protein structures, only single chain A was considered and other chains were removed by using a script for simplification and error minimization in carbon bond (C-bond) identification. An in-house program written in python was used to identify the C-bonds represented by Z-C•••O=C where the carbon atom bonded to Z = N, C, O atoms acts as C-bond donor and the carbonyl group acts as C-bond. The following criteria were considered to identify the C-bond interaction:

1) The minimum and maximum distances for C=O are 1.18 and 1.35 Å respectively.
2) The minimum and the maximum distances between the O of carbonyl group and the C of C-Z (Z=N,O, and C) bond were kept at 2.5 and 3.6 Å respectively.
3) The minimum and the maximum distances of C-Z (Y=N, O and C) bond were kept at 1.41 Å and 1.56 Å respectively.

The minimum distance of C-Z bond was important to circumvent the reciprocal C=O•••C=O C-bonds. Total of 7599 Z-C•••O=C C-bonds were obtained by considering aforementioned criteria. The secondary structure information about the C-bond donor and the acceptor residues were obtained using STRIDE[2] software and subsequently assigned by another in-house developed python code.

Materials and Experimental Methods:

N, N-Dimethylacetamide (NNDMA), acetonitrile (ACN) were purchased from Sigma-Aldrich and used directly without further purification for Nuclear Magnetic Resonance (NMR) and Infrared (IR) experiments. HPLC grade (purity 99.9%) carbon tetrachloride (CCl_4) solvent was purchased from Spectrochem Pvt. Ltd. This was dried by distillation followed by the addition of molecular sieves and was used to prepare the NNDMA and ACN sample solutions for NMR and IR experiments. PerkinElmer FTIR spectrometer of Frontier™ MIR/NIR/FIR model was used to record the vibrational spectra of monomers and complexes. IR Spectra were recorded for both monomers and complexes by using the horizontal attenuated total internal reflectance (HATIR) technique with ZnSe crystal. 2 cm^{-1} resolution with data intervals of 0.5 cm^{-1} and 16 scans were set for the IR experiment to record samples and background correction. ^{13}C NMR experiments were carried out using 400MHz (9.41 Tesla) Avance-III Bruker liquid state NMR spectrometer. For NMR experiements, 50 µL of deuterated dimethyl sulfoxide (DMSO-d_6) was taken in capillary tube for locking and external reference. The association constant for the C-bond complex formation was determined using the ^{13}C NMR concentration titration experiments. ^{13}C NMR titration experiments and IR experiments were carried out in two ways: (i) keeping NNDMA concentration (0.21M) in CCl_4 constant and varying the ACN concentration at 17 intervals from 0 M to 3.15 M, and (ii) keeping ACN concentration (1.7M) in CCl_4 constant and varying the NNMDA concentration at 17 intervals from 0 M to 3 M. For measuring the equilibrium constant of C-Bond complex, the Host-Guest complexation equilibrium was used as mentioned in Jonathan R. Nitschke et al.[3]

The Host-Guest equilibrium is represented as :

$$H + G \rightleftharpoons HG \qquad (1)$$

where H is the Host molecule and G is the Guest molecule

The equilibrium constant (K) for the host-guest complex is given by

$$K = \frac{[HG]}{[H][G]} \qquad (2)$$

$$\Rightarrow [HG] = K[H][G] \qquad (3)$$

The total concentration of guest $[G]_0$ and host $[H]_0$ in the solution is given

$$[G]_0 = [G] + [HG] = [G] + K[H][G] \qquad (4)$$

$$[H]_0 = [H] + [HG] = [H] + K[H][G] = [H] \text{ x } (1+K[G]) \qquad (5)$$

The fraction of host molecule (f_H) forming the H-G complex is given by

$$f_H = \frac{[HG]}{[H]_0} = \frac{K[G]}{1+K[G]} = y \qquad (6),$$

The equilibrium constant or complexation constant or binding constant (K) can be obtained from equation (6) by knowing the guest concentration at equilibrium ([G]). [G] can be expressed in terms of $[G]_0$, $[H]_0$ and K.

$$[G] = [G]_0 - [HG] = [G]_0 - K[H][G] \qquad (7)$$

From equations (5) & (6) it follows that the free host concentration at equilibrium is equal to

$$[H] = (1-y)[H]_0 \qquad (8)$$

Combining equation (7) and (8), we have

$[G] = [G]_0 - K(1\text{-}y)[H]_0[G]$ (9)

Rearranging above equation (9) we get

$(1+K(1\text{-}y)[H]_0) \times [G] = [G]_0$ (10)

From equation (10) it follows that the free guest concentration at equilibrium is equal to

$$[G] = \frac{[G]_0}{1+K(1-y)[H]_0} \quad (11)$$

Equation (6) can rearranged as

$y + yK[G] = K[G]$ (12)

Equation (11) substituted in equation (12) gives

$$y + yK\left(\frac{[G]_0}{1+K(1-y)[H]_0}\right) = K\left(\frac{[G]_0}{1+K(1-y)[H]_0}\right) \quad (13)$$

Rearranging the equation (13), we get

$y(1 + K[H]_0 - Ky[H]_0) + yK[G]_0 - K[G]_0 = 0$ (14)

Rearranging equation (14), we get

$$-K[H]_0y^2 + (1 + K([H]_0 + [G]_0))y - K[G]_0 = 0 \qquad (15)$$

Finally, we can rearrange the equation (15) and get

$$K[H]_0y^2 - (1 + K([H]_0 + [G]_0))y + K[G]_0 = 0 \qquad (16)$$

Equation (16) is a quadratic equation in y and was solved to give two roots as below

$$y = \frac{[1+K([H]_0+[G]_0)] \pm \sqrt{[1+K([H]_0+[G]_0)]^2 - 4K^2[H]_0[G]_0}}{2K[H]_0} \qquad (17)$$

From the two roots of equation (17), it was found that the physically meaningful root is given by:

$$y = \frac{[1+K([H]_0+[G]_0)] - \sqrt{[1+K([H]_0+[G]_0)]^2 - 4K^2[H]_0[G]_0}}{2K[H]_0} \qquad (18)$$

Equation (18) was rearranged to the equation (19) as mentioned below and used in origin software to fit the curve obtained from the chemical shifts versus Hosts / Guests concentration.

$$Y = \frac{Y_0 + \Delta Y(K([H]_0+[G]_0)+1) - \sqrt{[K([H]_0+[G]_0)+1]^2 - 4K^2[H]_0[G]_0}}{2K[H]_0} \qquad (19)$$

The unknown parameters K, Y_0 and ΔY were obtained from the fitted curve.

Where, Y is a measured chemical shift, Y_0 is a chemical shift of only host in solution, ΔY is the maximal change in chemical shift i.e. the difference in chemical shift of a fully occupied host and an empty host, K is the association constant, $[H]_0$ is a total host concentration in solution, $[G]_0$ is the total guest concentration in solution.

Quantum Chemical Calculations:

All the geometry optimizations and numerical frequency calculations of monomers and complexes of model compounds were performed using standard Kohn–Sham density theory with dispersion correction at B97-D3/aug-cc-pVDZ level of theory.[4] The binding energies of carbon bond (C-bond) complexes of model compounds were calculated with gold standard Coupled Cluster Singles Doubles and Triples (CCSD-T) at aug-cc-pVDZ basis set. Out of the protein structures obtained from the dataset (7599 C-bond interactions), 22 protein structures were chosen for quantum chemistry calculations to study Z-C•••O=C C-bonds. Then the C-bond interacting residues in these 22 proteins were truncated followed by the addition of hydrogen atoms and optimizations of the added hydrogen positions while freezing other nuclei at their crystal structure positions. The optimization was carried out at B97D/aug-cc-pVDZ level of theory. The donor-acceptor interaction energy (E_{DA}) between the donor and acceptor residues in proteins and also for model C-bond complexes were obtained by Natural Bond Orbital (NBO) analysis at MP2/aug-cc-pVDZ level of theory. All the geometry optimizations and energy calculations were carried out using Turbomole 6.5 package[5] and the NBO calculations using NBO 6.0 program[6]. Non-Covalent Interaction (NCI) analysis and Atoms in Molecules (AIM) analysis were executed for the confirmation of existences of C-bonds using the wave functions of the optimized structures at MP2/aug-cc-pVDZ level. Molecular Electrostatic Potential (MESP) analysis was performed using Gaussian09 software[7] and subsequently MESP surfaces were plotted using Gauss View software. Binding energies for C-bond formation were computed at CCSD (T)/aug-cc-pVDZ level of theory as follows:

<u>Binding energy of C-Bond complexes</u>

$E'^{CCSD-T}_{Gas} = E^{CCSD-T}_{Gas} + ZPE^{B97D}_{Gas}$

E^{CCSD-T}_{Gas} = Electronic energy at CCSD-T/aug-cc-pVDZ level in gas phase

ZPE^{B97D}_{Gas} = Zero-point energy at B97D/aug-cc-pVDZ level

Binding energy of C-Bond complexes

$D_0 = E'_{Complex} - (E'_{monomer1} + E'_{monomer2})$ (20)

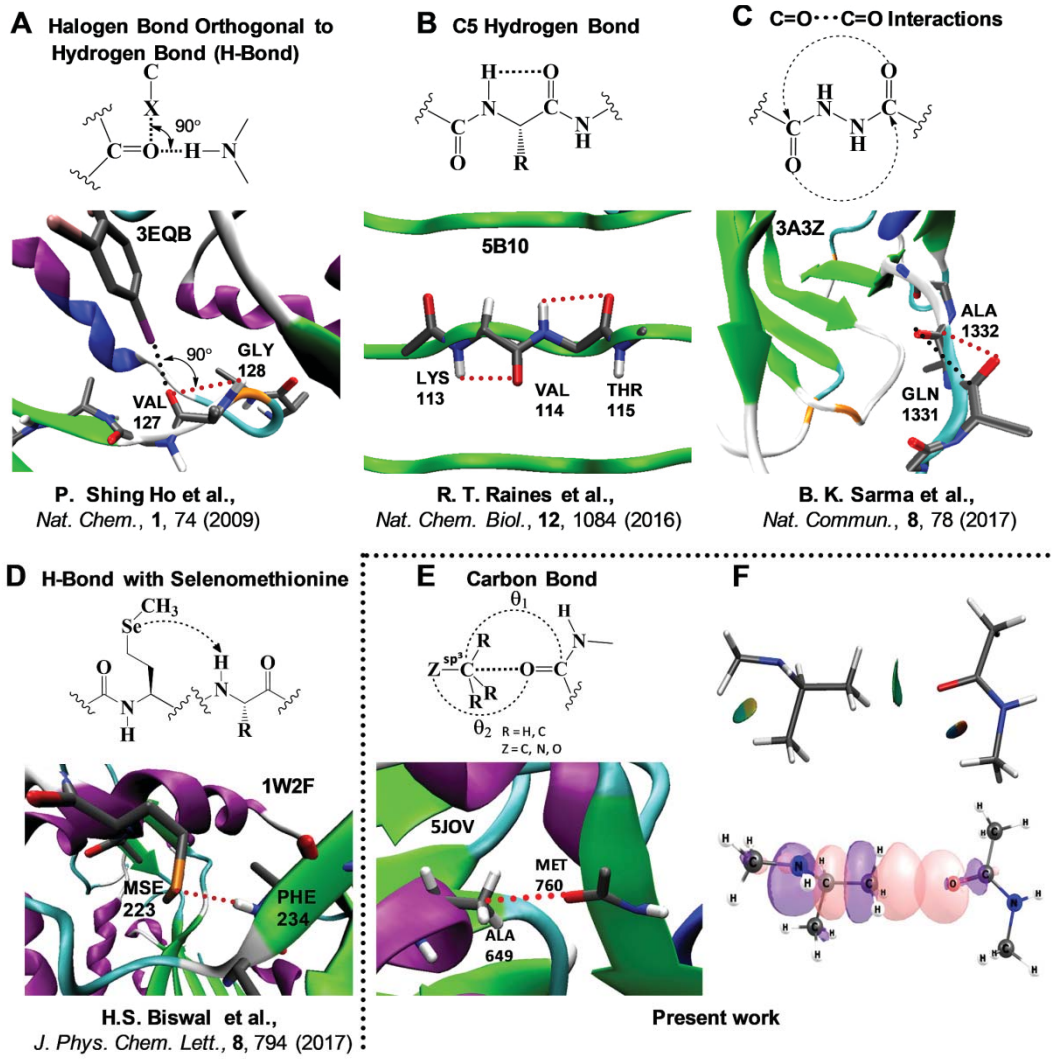

Fig. S1. Recent discovery of novel non-covalent interactions in proteins. ***(A)*** *Formation of orthogonal* ***I•••O=C-VAL^{127}*** *halogen bond (X-bond) and* ***GLY^{128}-N-H•••O=C-VAL^{127}*** *hydrogen bond (H-bond) in human MAP kinase1 (PDB:3EQB) forms halogen bond (X-bond). In this case, the X-bond and H-bond are perpendicular to each other, representing an example of orthogonal molecular interactions (hX-bond) in proteins.* ***(B)*** *Formation of inraresidue intramolecular C5-H-bonds between amide-NH and O of carbonyl group of same residue in a mutant of OspA protein (PDB: 5B10).* ***(C)*** *Reciprocal carbonyl-carbonyl (C=O•••C=O) interactions in vitamin-D nuclear receptor (PDB: 3A3Z), due to back and forth electron delocalization between the*

carbonyl groups, that are frequently observed in polyproline II helices. ***(D)*** *The amide-N-H•••Se H-bond in Selenomethionine containing protein (PDB: 1W2F).* ***(E)*** *The formation of Z-R_3C^{sp3}•••O=C carbon bond (C-bond) in XyGUL proteins (PDB: 5JOV), representing a highly directional hydrophobic interactions between an electron-rich carbonyl-oxygen acceptor and an electron-deficient sp^3-hybridized carbon σ-hole donor through n→σ* electron delocalization.* ***(F) Top:*** *Reduced density gradient isosurfaces in real 3D space for Ala^{649}-C•••O=C-Met^{760} C-bond, the green isosurface indicates attractive noncovalent interaction between the carbonyl and methyl group.* ***Bottom:*** *The $n_O \rightarrow \sigma^*_{C\text{-}C}$ electron delocalization due to the overlap of p-type oxygen lone pair (n) and antibonding C-C (σ*) orbital.*

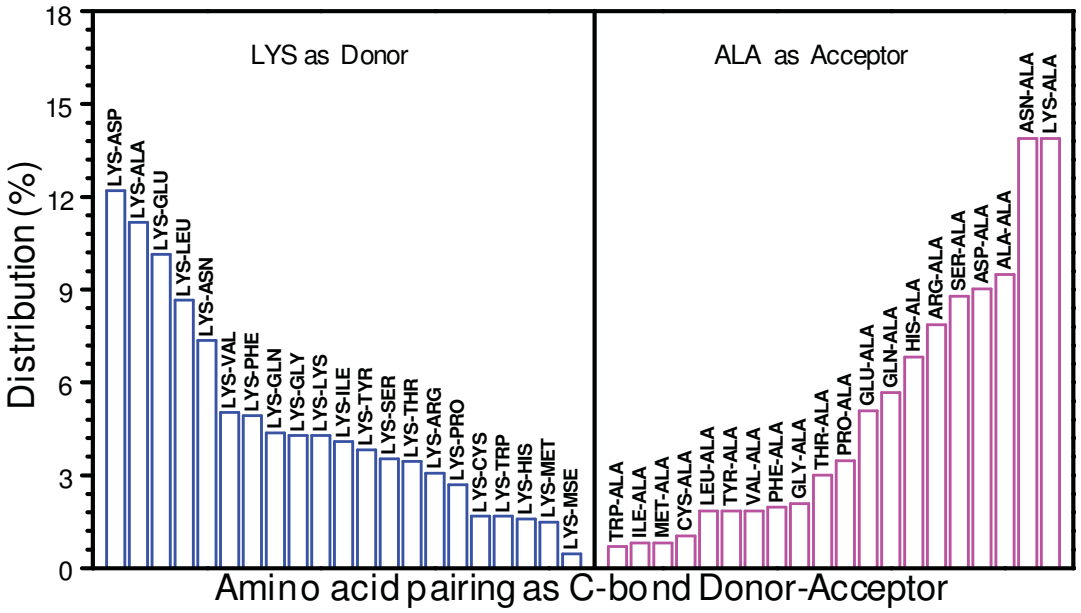

Fig. S2|. Distribution of amino acid pairs as C-bond acceptor-donor in proteins through PDB analysis. Left Panel: Histograms representing the percentage distribution of amino acids involved in C-bonds where LYS as C-bond donor pairs with other amino acids. **Right Panel:**

Histograms representing the percentage distribution of amino acids involved in C-bonds where ALA as C-bond acceptor pairs with other amino acids.

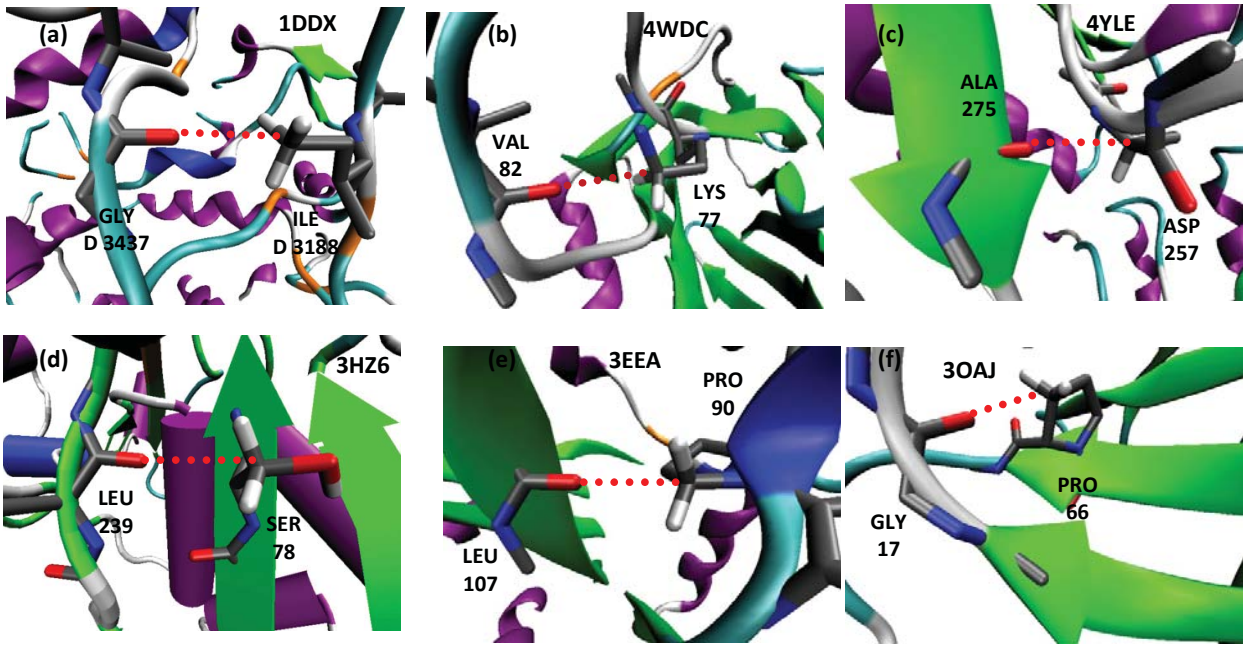

Fig. S3|. Representative examples of Z-R_3C ···O=C C-bonds in proteins. **(a)** C-C···O=C C-bond formed by C=O of GLY^{3437} with CH_3 primary carbon of ILE^{3188} (PDB: 1DDX). **(b)** C-C···O=C C-bond formed by C=O of VAL^{82} with secondary carbon of LYS^{77} (PDB: 4WDC). **(c)** C-C···O=C C-bond formed by C=O of ALA^{275} with tertiary carbon atom of ASP^{257} (PDB: 4YLE). **(d)** O-C···O=C type C-bond formed by C=O of LEU^{239} with C-O of SER^{78} in (PDB: 3HZ6) **(e)** N-C···O=C type C-bond formed by C=O of LEU^{107} with C-N of PRO^{90} (PDB:3EEA) **(f)** C-C···O=C type C-bond formed by C=O of GLY^{17} with C-C of PRO^{66} (PDB: 3OAJ).

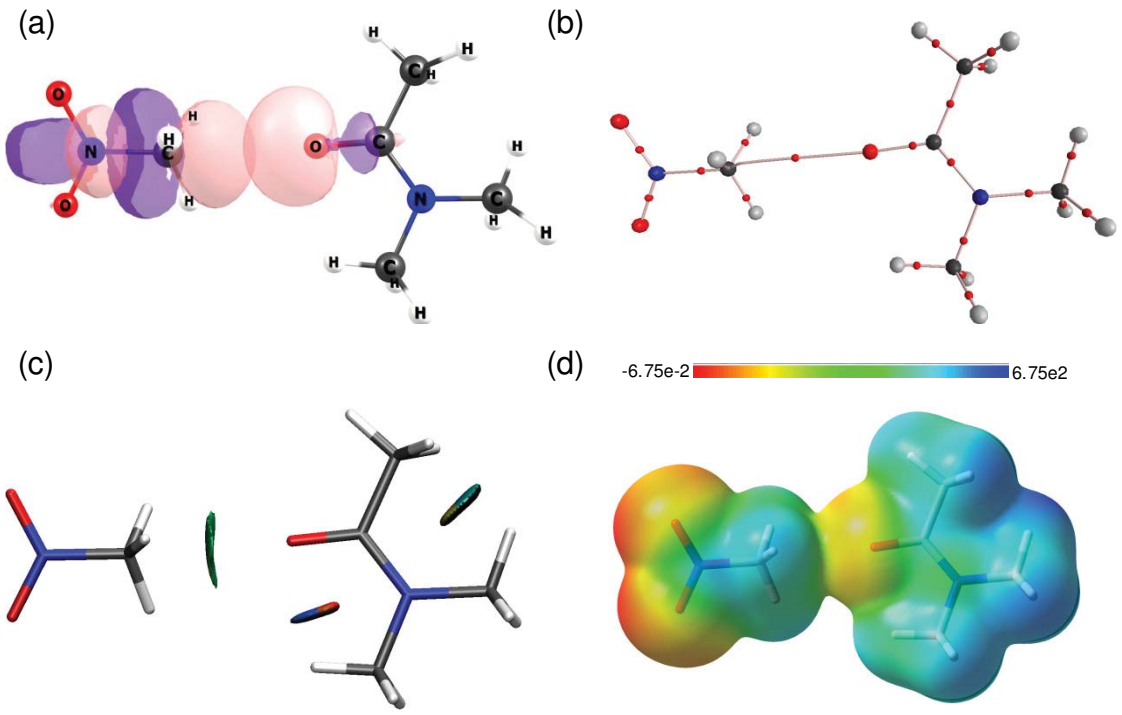

Fig. S4|. Topological analysis for evidence of C-bond in NNDMA-NM. The different topological analyses show the existence of C-N···O=C C-bond in NNDMA-NM complex, **(a)** overlap of p-type carbonyl-oxygen lone pair and C-N σ* orbital i.e. $n_O \rightarrow \sigma^*_{C\text{-}N}$ electron delocalization **(b)** Molecular graph and electron density topology analysis showing bond critical point(red dot) along C···O C-bond **(c)** Reduced density gradient isosurfaces in 3D real space, green isosurface represents the existence of attraction between donor and acceptor. **(d)** Molecular electrostatic potential showing a σ-hole on C atom of C-bond donor and mapped on the 0.004 au surface of electron density. Negative values are shown in red and positive values in blue.

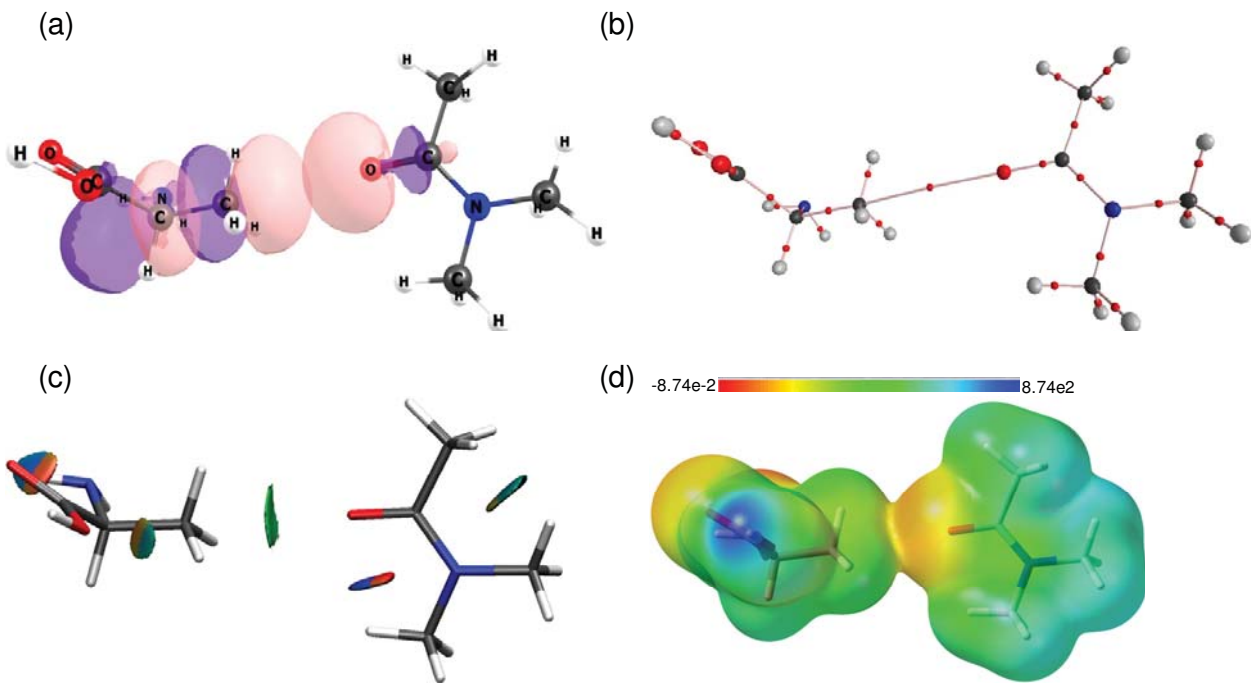

Fig. S5|. Topological analysis for evidence of C-bond in NNDMA-ALA. The different topological analyses show the existence of C-C···O=C C-bond in **NNDMA-ALA** complex, **(a)** overlap of p-type carbonyl-oxygen lone pair and C-C σ* orbital i.e. $n_O \rightarrow \sigma^*_{C\text{-}C}$ electron delocalization **(b)** Molecular graph and electron density topology analysis showing bond critical point(red dot) along C···O C-bond **(c)** Reduced density gradient isosurfaces in 3D real space, green isosurface represents the existence of attraction between donor and acceptor. **(d)** Molecular electrostatic potential showing a σ-hole on C atom of C-bond donor and mapped on the 0.004 au surface of electron density. Negative values are shown in red and positive values in blue.

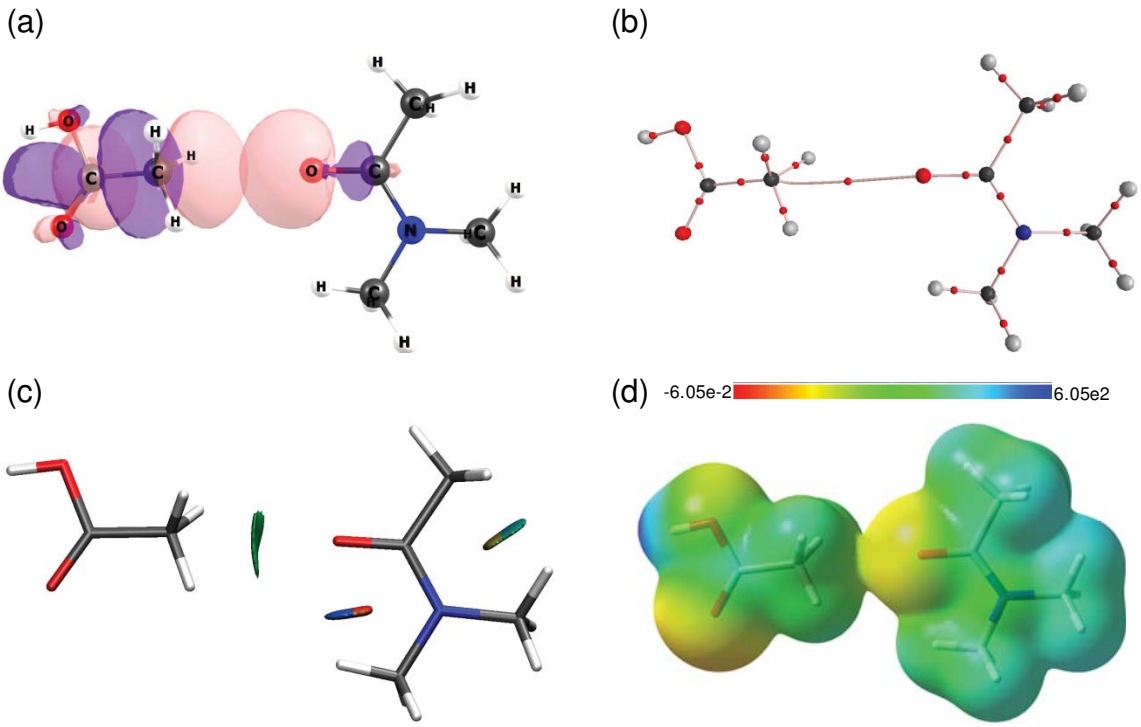

Fig. S6l. Topological analysis for evidence of C-bond in NNDMA-AA. The different topological analyses show the existence of C-C···O=C C-bond in **NNDMA-AA** complex, **(a)** overlap of p-type carbonyl-oxygen lone pair and C-C σ* orbital i.e. $n_O \rightarrow \sigma^*_{C\text{-}C}$ electron delocalization **(b)** Molecular graph and electron density topology analysis showing bond critical point(red dot) along C···O C-bond **(c)** Reduced density gradient isosurfaces in 3D real space, green isosurface represents the existence of attraction between donor and acceptor. **(d)** Molecular electrostatic potential showing a σ-hole on C atom of C-bond donor and mapped on the 0.004 au surface of electron density. Negative values are shown in red and positive values in blue.

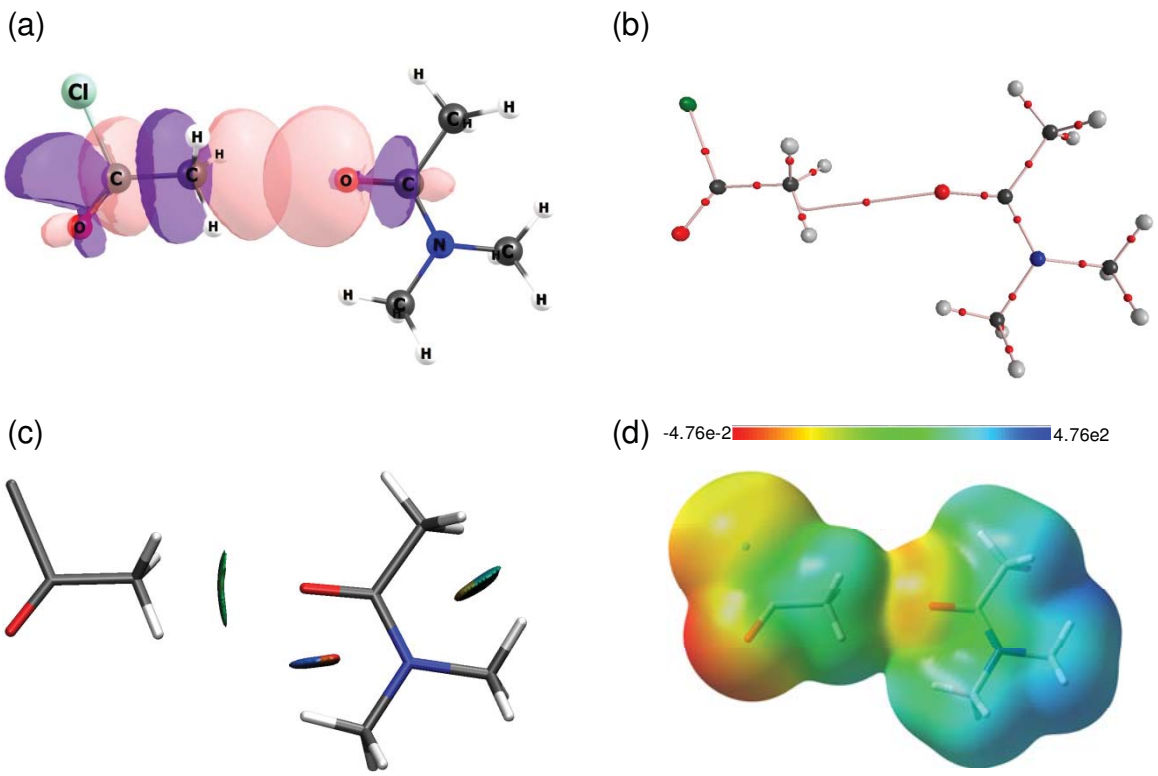

Fig. S7l. Topological analysis for evidence of C-bond in NNDMA-AcetylCl. The different topological analyses show the existence of C-C···O=C C-bond in **NNDMA-AcetylCl** complex, **(a)** overlap of p-type carbonyl-oxygen lone pair and C-C σ* orbital i.e. $n_O \rightarrow \sigma^*_{C\text{-}C}$ electron delocalization **(b)** Molecular graph and electron density topology analysis showing bond critical point(red dot) along C···O C-bond **(c)** Reduced density gradient isosurfaces in 3D real space, green isosurface represents the existence of attraction between donor and acceptor. **(d)** Molecular electrostatic potential showing a σ-hole on C atom of C-bond donor and mapped on the 0.004 au surface of electron density. Negative values are shown in red and positive values in blue.

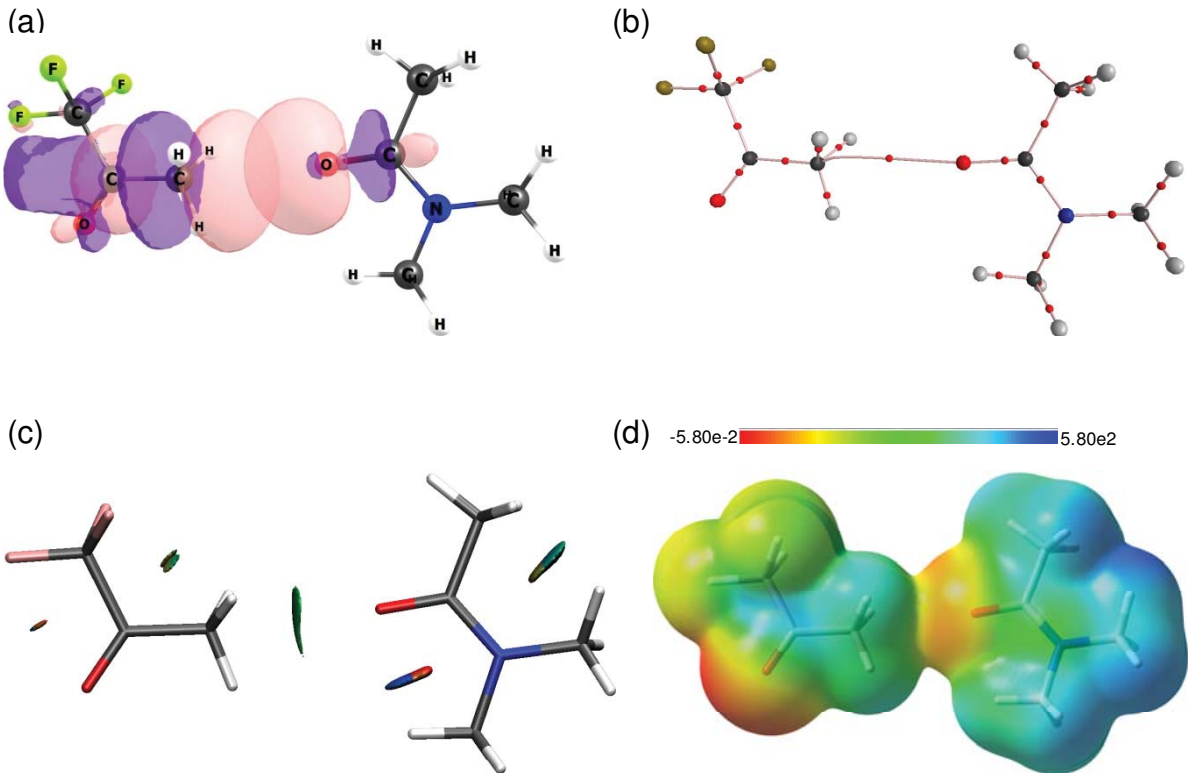

Fig. S8|. Topological analysis for evidence of C-bond in NNDMA-TFACE. The different topological analyses show the existence of C-C···O=C C-bond in **NNDMA-TFACE** complex, **(a)** overlap of p-type carbonyl-oxygen lone pair and C-C σ* orbital i.e. $n_O \rightarrow \sigma^*_{C\text{-}C}$ electron delocalization **(b)** Molecular graph and electron density topology analysis showing bond critical point(red dot) along C···O C-bond **(c)** Reduced density gradient isosurfaces in 3D real space, green isosurface represents the existence of attraction between donor and acceptor. (d) Molecular electrostatic potential showing a σ-hole on C atom of C-bond donor and mapped on the 0.004 au surface of electron density. Negative values are shown in red and positive values in blue.

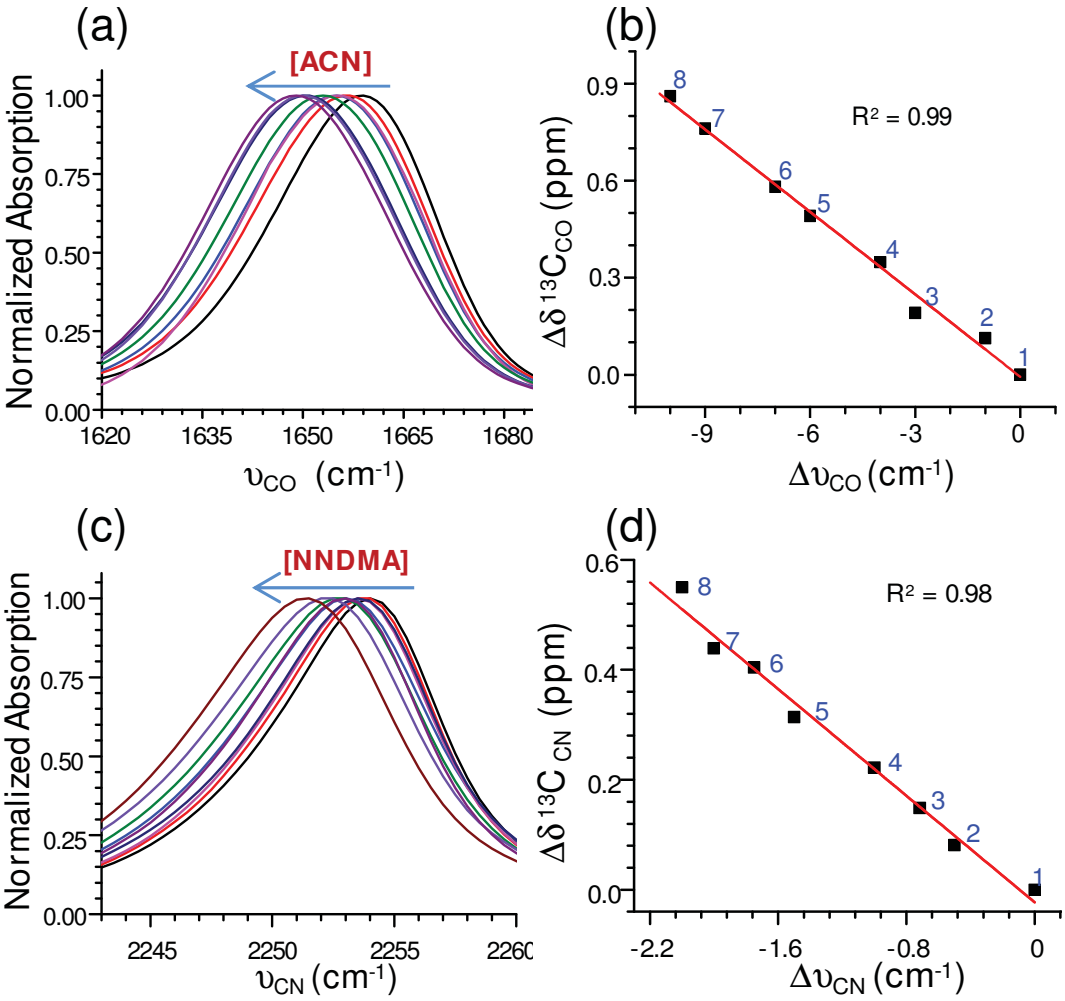

Fig. S9l. IR spectra and correlation plots between IR and NMR spectral data in C-bond complexes. IR Spectra and correlation between redshifts in IR and downfield chemical shifts in NMR, **(a)** Represents the C=O IR peak shifts towards lower frequencies with increased [ACN], **(b)** Correlation plot between the changes in chemical shifts(Δδ) and vibrational stretching frequencies(Δυ) of C=O, **(c)** Represents the C≡N IR peak shifts towards lower frequencies with increased [NNDMA], **(d)** Correlation plot between the changes in chemical shifts(Δδ) and vibrational stretching frequencies(Δυ) of C≡N.

Table S1l. The values of carbon bond distance (d) in Å, the donor-acceptor interaction energies (E_{DA}) in kJ/mol, calculated binding energies (Calc.BE) and estimated binding energies (Est.BE) in kJ/mol of carbon bond complexes of model compounds and proteins at CCSD-T/aug-cc-pVDZ level.

S.No	PDB/Model compounds	Distance (Å)	E_{DA} (kJ/mol)	Calc.BE(CCSD-T) (kJ/mol)	Est.BE(CCSD-T) (kJ/mol)
1	NNDMA-ALA	3.31	0.711	-8.4	-6.6
2	NNDMA-AA	3.13	1.757	-13.3	-13.6
3	NNDMA-TFACE	3.08	2.259	-17.8	-15.8
4	NNDMA-AcetylCl	3.04	2.678	-18.2	-17.4
5	NNDMA-ACN	3.05	2.929	-18.4	-17.0
6	NNDMA-NM	2.94	5.816	-22.1	-20.6
7	3EEA	2.79	8.786		-22.1
8	3DNX	2.87	6.527		-21.7
9	3A57	2.91	4.226		-21.2
10	3DXY	2.96	3.264		-20.1
11	4PGM	2.97	3.305		-19.9
12	2O7C	3.08	2.845		-15.8
13	3OAJ	3.02	2.301		-18.2
14	5JOV	3.02	2.678		-18.2
15	3LGI	3.03	1.841		-17.8
16	3HZ6	3.05	2.552		-17.0
17	4Q6U	3.08	2.803		-15.8
18	4WDC	3.1	1.046		-14.9
19	3ZK4	3.13	0.711		-13.6
20	1RY9	3.15	0.418		-12.7
21	4IVN	3.15	1.757		-12.7
22	1DDX	3.16	1.464		-12.2
23	2FAO	3.19	0.460		-11.0
24	4YLE	3.2	0.293		-10.6

25	3JQ1	3.21	0.711		-10.2
26	5JUF	3.21	0.544		-10.2
27	2QRU	3.25	0.586		-8.6
28	1M6S	3.36	0.586		-5.3

Table S2|. The chemical shift (δ) values of $^{13}C=O$ of NNDMA and ^{13}CN of ACN at different concentrations of ACN and NNDMA, respectively. [NNDMA] and [ACN] are the concentrations of N,N-dimethylacetamide and acetonitrile, respectively. Titration 1: [NNDMA] kept constant varying the [ACN]. Titration 2: [ACN] kept constant varying the [NNDMA].

S.No.	**[ACN] (M)**	**[NNDMA] (M)**	**^{13}C δ(C=O) (ppm)**	**[NNDMA] (M)**	**[ACN] (M)**	**^{13}C δ(CN) (ppm)**	**^{13}C δ(CH_3) (ppm)**
Titration-1				**Titration-2**			
1	0.00	0.21	168.87	0.00	1.70	115.85	1.94
2	0.08	0.21	168.91	0.08	1.70	115.90	1.93
3	0.15	0.21	168.94	0.17	1.70	115.93	1.92
4	0.22	0.21	168.98	0.25	1.70	115.96	1.90
5	0.30	0.21	169.01	0.33	1.70	116.00	1.89
6	0.37	0.21	169.06	0.41	1.70	116.03	1.89
7	0.44	0.21	169.11	0.51	1.70	116.07	1.88
8	0.51	0.21	169.13	0.60	1.70	116.11	1.86
9	0.59	0.21	169.16	0.79	1.70	116.16	1.84
10	0.73	0.21	169.22	0.97	1.70	116.22	1.82
11	0.90	0.21	169.27	1.14	1.70	116.29	1.81
12	1.24	0.21	169.36	1.31	1.70	116.34	1.79
13	1.40	0.21	169.41	1.47	1.70	116.40	1.78
14	1.56	0.21	169.45	1.78	1.70	116.49	1.76
15	2.03	0.21	169.52	2.06	1.70	116.56	1.73
16	2.61	0.21	169.63	2.46	1.70	116.66	1.70
17	3.15	0.21	169.73	3.04	1.70	116.78	1.66

References

1 Berman, H. M. et al. The Protein Data Bank. Nucleic Acids Res. 28, 235-242, (2000).

2 Frishman, D. & Argos, P. Knowledge-based protein secondary structure assignment. Proteins 23, 566-579, (1995).

3 Hristova, Y. R., Smulders, M. M. J., Clegg, J. K., Breiner, B. & Nitschke, J. R. Selective anion binding by a "Chameleon" capsule with a dynamically reconfigurable exterior. Chem. Sci. 2, 638-641, (2011).

4 Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J. Chem. Phys. 132, 154104, (2010).

5 Furche, F. et al. Turbomole. Wiley Interdiscip. Rev. Comput. Mol. Sci. 4, 91-100, (2014).

6 Glendening, E. D., Landis, C. R. & Weinhold, F. NBO 6.0: Natural bond orbital analysis program. J. Comput. Chem. 34, 1429-1437, (2013).

7 Frisch, M. J. et al. Gaussian 09, Revision B.01. (2009).

Optimized Structures coordinates:

NNDMA

C	0.049620400	0.000000000	4.255095300
O	0.020136900	0.000000000	3.018354400
N	1.240912100	0.000000000	4.957639600
C	2.510319500	0.000000000	4.232540800
H	3.100223000	-0.896500000	4.500223300
H	2.306743700	0.000000000	3.154689000
H	3.100223000	0.896500000	4.500223300
C	-1.248597900	0.000000000	5.064728300
H	-2.078369300	0.000000000	4.346694900
H	-1.328889800	-0.893039600	5.708396100
H	-1.328889800	0.893039600	5.708396100
C	1.368406900	0.000000000	6.410873500
H	1.927544400	-0.894691100	6.746364500
H	1.927544400	0.894691100	6.746364500
H	0.390391000	0.000000000	6.904241500

ACN

C	2.904718600	0.751166800	0.359372200
H	2.430719900	0.226671800	1.205460200
H	2.513332600	1.781017900	0.315231600
H	2.640205900	0.229825500	-0.575457400

C	4.360553200	0.771095400	0.530031800
N	5.522712100	0.785843000	0.665299900

ALA

C	3.415630400	2.462165100	2.075955300
H	2.792438900	1.972704500	2.842650700
H	2.817107200	3.253109500	1.597314000
H	3.704223300	1.712889500	1.324022300
C	4.657108400	3.076439100	2.750290300
H	5.281813300	2.238163000	3.140496400
N	4.249846300	4.031748300	3.788409600
H	3.923906200	3.528106900	4.613122100
H	5.043382000	4.610834800	4.066904100
C	5.586071200	3.784980200	1.763332900
O	6.150056700	4.846081700	1.963501300
O	5.771703300	3.053608900	0.620168700
H	6.412278300	3.554847500	0.082457000

AcetylCl

C	-1.866705100	1.704539700	0.018420500
H	-1.370782400	2.317023300	0.788577000
H	-1.736107100	0.632357200	0.224546700
H	-1.424473000	1.975458600	-0.953818000

C	-3.336815500	2.031321300	-0.000262700
O	-4.265807800	1.303241800	0.171700100
Cl	-3.614745600	3.836948600	-0.346704300

AA

C	-0.043124200	-1.928461200	2.196444300
H	0.580233200	-2.837269600	2.151995300
H	-0.425839000	-1.741443300	1.179953400
H	0.556169300	-1.077884400	2.545119700
C	-1.194397500	-2.152447700	3.150920500
O	-1.392678000	-1.552865300	4.192002200
O	-2.025478000	-3.150716100	2.702931900
H	-2.729688800	-3.235636300	3.372280200

NM

C	-1.217404900	-2.014169900	0.002650600
H	-2.022793700	-1.643757300	-0.644927400
H	-1.381682100	-1.676901100	1.036131900
H	-0.219101600	-1.742105400	-0.358060400
N	-1.317137200	-3.527586200	0.024271500
O	-0.276532400	-4.169902200	-0.117089800
O	-2.439315400	-4.002582100	0.200270300

TFACE

C	0.941871100	-0.853887900	0.001122900
H	0.607394600	-0.275533700	-0.877386500
H	0.502614700	-1.860192100	-0.009017800
H	0.612715500	-0.296804600	0.895352800
C	2.448376700	-0.964255000	-0.003063700
O	3.083470400	-1.996953100	-0.007399200
C	3.208608300	0.415480400	-0.000463200
F	2.854339600	1.145973200	-1.105043000
F	2.852105300	1.143204800	1.105234400
F	4.545876700	0.271651000	0.000663400

NNDMA-ACN

C	2.794177600	0.590869300	0.003551500
H	2.401374100	0.213733900	0.960428100
H	2.655883300	1.682785200	-0.016080300
H	2.191589800	0.156288100	-0.808823700
C	4.208885200	0.241210600	-0.151830100
N	5.339877000	-0.035808100	-0.273558200
C	-1.353804200	1.607513600	0.452130000
O	-0.155853500	1.313911500	0.322139100
C	-1.793585100	3.070493700	0.508778600
H	-2.270491200	3.314748100	1.473778800

H	-0.894479600	3.687924400	0.389478400
H	-2.509929400	3.313474000	-0.294595400
N	-2.333213400	0.642287800	0.561848100
C	-3.767829200	0.913614400	0.615720400
H	-4.239456500	0.234310200	1.347348000
H	-3.973166800	1.943591600	0.929131400
H	-4.244659800	0.739513400	-0.368944600
C	-1.978000700	-0.771488000	0.434355600
H	-2.418834600	-1.196775300	-0.486959600
H	-0.886237200	-0.864649100	0.393560200
H	-2.368339800	-1.335562400	1.300120300

NNDMA-NM

C	-1.209924600	2.162094800	0.122940500
O	-1.185410600	0.922143700	0.102931200
N	-0.053098800	2.911529600	0.125205000
C	1.249025500	2.244516300	0.105650700
H	1.826742100	2.513163300	1.009044700
H	1.096445600	1.159128400	0.075823700
H	1.824001200	2.563641500	-0.782854400
C	-2.543863800	2.908668700	0.141367700
H	-3.340507600	2.154356500	0.127378200
H	-2.652978300	3.529222100	1.047204400

H	-2.655724500	3.566209400	-0.737701300
C	0.009346100	4.369697600	0.168663800
H	0.545327900	4.701520100	1.078059600
H	0.559320700	4.753523700	-0.711041100
H	-0.989880300	4.817739500	0.173599000
C	-1.217407400	-2.018425600	0.002561800
H	-2.019559900	-1.637584100	-0.639453200
H	-1.382627100	-1.669414900	1.030194600
H	-0.222675300	-1.738727500	-0.356582700
N	-1.316340900	-3.527473400	0.024050700
O	-0.277271600	-4.176869700	-0.117255200
O	-2.438087400	-4.008514300	0.199730800

NNDMA-AA

C	3.652399200	-1.487389900	-0.065921700
O	2.564386700	-1.546083300	0.522768300
N	4.689169900	-0.696595900	0.388831700
C	4.551889900	0.011475900	1.661971800
H	3.514222200	-0.072682800	2.006613400
H	5.228053400	-0.425690400	2.420996700
H	4.816628100	1.075450700	1.529200300
C	3.889620100	-2.287195700	-1.346493500
H	2.977104400	-2.862255700	-1.547563700

H	4.093780500	-1.625530700	-2.205833300
H	4.741670900	-2.980480400	-1.242412100
C	6.034045000	-0.692354800	-0.181986000
H	6.023836400	-0.974719000	-1.241647800
H	6.454654100	0.324941200	-0.105694200
H	6.707973300	-1.382144100	0.364353400
C	-0.045883700	-1.927501700	2.200092900
H	0.584200000	-2.829309800	2.145889300
H	-0.417042700	-1.743115700	1.180015800
H	0.561639500	-1.080584200	2.537946200
C	-1.194776200	-2.152325400	3.150613800
O	-1.400791700	-1.557797900	4.195284000
O	-2.030334600	-3.152620300	2.704683900
H	-2.731815400	-3.233473700	3.377127700

NNDMA-ALA

C	3.416043100	2.464479400	2.081413100
H	2.792653200	1.975803300	2.847087300
H	2.809082100	3.245046600	1.600255100
H	3.690324900	1.711203500	1.330319000
C	-0.255595600	0.725667000	0.034208500
O	0.743308300	1.207855500	0.582516800
C	-1.238301200	1.618776500	-0.724559100

H	-2.257235700	1.551513300	-0.306107300
H	-0.878199600	2.651440300	-0.635607900
H	-1.292483600	1.347831000	-1.793245900
N	-0.524376300	-0.630590800	0.065740600
C	-1.746420700	-1.244666300	-0.446098000
H	-2.484635900	-1.412814200	0.363602100
H	-2.212339000	-0.628531200	-1.224393800
H	-1.501117400	-2.224869300	-0.890729200
C	0.329603700	-1.516071000	0.857205700
H	0.648402800	-2.376744500	0.242468600
H	1.209369900	-0.955075700	1.195408000
H	-0.221669200	-1.900672000	1.736751700
C	4.660161700	3.079683000	2.746743500
H	5.289308900	2.242760800	3.134647800
N	4.259516000	4.036785200	3.787463500
H	3.895976700	3.531370600	4.595304300
H	5.063845500	4.586185400	4.093533500
C	5.583678500	3.789381200	1.757424700
O	6.135408200	4.860006200	1.946264500
O	5.786838300	3.047962900	0.622850100
H	6.422745100	3.555022800	0.085326900

NNDMA-AcetylCl

C	2.202708100	0.401846600	0.041468300
O	1.028945900	0.800126000	0.062233000
N	2.520685800	-0.936230400	0.149479900
C	1.452857700	-1.935149400	0.201183900
H	1.591189200	-2.587554900	1.081775700
H	0.485912600	-1.422433000	0.268846100
H	1.473040500	-2.565765500	-0.707569700
C	3.357769900	1.393762300	-0.094718600
H	2.919567400	2.395083100	-0.190357600
H	4.016731400	1.374038200	0.790531400
H	3.976570100	1.181241500	-0.983209600
C	3.870334300	-1.482917400	0.034082100
H	4.001348600	-2.287900200	0.778334600
H	4.041764500	-1.915575100	-0.971349100
H	4.633183700	-0.719001000	0.223747100
C	-1.872181800	1.703229900	0.018645700
H	-1.366121500	2.312059100	0.782951800
H	-1.725094600	0.635097000	0.223736200
H	-1.418965400	1.973069200	-0.947103000
C	-3.337602300	2.025357000	0.000927300
O	-4.274496100	1.306943600	0.171410600
Cl	-3.620979700	3.845138600	-0.348109500

NNDMA-TFACE

C	-3.230964300	-0.307185200	-0.015030600
O	-2.011314800	-0.527445400	-0.002376700
N	-4.158574600	-1.327572400	-0.068337100
C	-3.711125800	-2.718891500	0.010157300
H	-4.159221900	-3.303967400	-0.812301900
H	-2.617746800	-2.747187500	-0.069008200
H	-4.020616900	-3.171240200	0.971350000
C	-3.765229600	1.124847900	0.013027100
H	-2.901746200	1.796353900	0.099112200
H	-4.317593900	1.367877400	-0.911327000
H	-4.445033800	1.292488100	0.865854900
C	-5.602201300	-1.126252900	0.043164200
H	-6.119193300	-1.911755400	-0.532988900
H	-5.943956000	-1.191567800	1.095291500
H	-5.904810000	-0.154644500	-0.366771400
C	1.050021500	-0.907725700	0.002130900
H	0.683958800	-0.339693900	-0.868722000
H	0.629125300	-1.920285400	-0.006939400
H	0.690459600	-0.358610300	0.887780500
C	2.552682000	-0.971967700	-0.003281800
O	3.228116600	-1.981151600	-0.010548900
C	3.274207200	0.430746600	0.000405400

F	2.906799300	1.154253200	-1.105082300
F	2.901212200	1.152702300	1.105035800
F	4.617756700	0.324384700	0.003764000

PDB structures:

1M6S

C	17.724995000	62.030982600	10.551997000
C	19.043994700	61.397982800	10.118997200
O	19.800994400	61.991982600	9.348997400
H	17.944911100	62.995899800	11.036054400
H	17.130256000	61.409030400	11.242875300
H	17.126982800	62.239419900	9.650540300
N	19.330994600	60.196983100	10.611997000
C	20.568994200	59.533983300	10.232997100
H	18.685100400	59.731969100	11.236441200
H	20.617011500	59.401071100	9.139536200
H	20.615425500	58.551781700	10.725626400
H	21.439873000	60.138826500	10.536076400
C	22.137993800	63.255982200	7.291998000
C	23.088993500	63.737982100	6.197998300
O	23.514993400	62.958982300	5.347998500
H	21.819254600	64.035557600	8.003631200
H	22.625300500	62.444206800	7.853845300

H	21.239848200	62.826165400	6.823562500
N	23.402993400	65.029981700	6.212998300
C	24.340993200	65.589981600	5.250998500
H	23.065512400	65.615891900	6.965583600
H	24.309107500	66.688202300	5.313361300
H	24.063308700	65.268738000	4.234956000
H	25.371819300	65.242297700	5.444290200

2FAO

C	10.259997100	4.602998700	35.896989900
C	9.898997200	3.189999100	35.470990000
O	9.000997500	3.000999200	34.656990300
H	10.446748900	5.204818100	34.994246800
H	11.137096300	4.665022600	36.563345100
H	9.389812000	5.041312400	36.412582900
N	10.560997000	2.193999400	36.042989900
C	10.215997100	0.823999800	35.708990000
H	11.298826400	2.381958500	36.708786500
H	9.161135600	0.620259400	35.957877100
H	10.864102800	0.139793700	36.275420900
H	10.347153400	0.643479100	34.629311100
N	6.073998300	2.532999300	34.664990300
C	6.178998300	2.655999300	33.213990700

C	4.809998700	2.643999300	32.532990900
O	4.571998700	3.433999000	31.618991100
C	7.049998000	1.537999600	32.643990800
H	6.117156200	1.544337800	34.928550600
H	6.923224200	2.928568300	35.069134800
H	6.614288000	3.636631700	32.975350800
H	6.593611000	0.557114000	32.874276900
H	8.051002600	1.584461800	33.093357300
H	7.144045100	1.608966900	31.545491500
N	3.907998900	1.766999500	32.970990700
C	2.594999300	1.706999500	32.342990900
H	4.046189500	1.331815600	33.874719600
H	1.993595200	2.606771200	32.567883300
H	2.063796400	0.812485500	32.703222400
H	2.716967000	1.649955700	31.250925700

4IVN

C	18.018994900	-6.065998300	4.009998900
C	17.969995000	-7.389997900	3.240999100
O	19.025994700	-7.975997800	2.939999200
H	18.644731900	-5.356550300	3.446149800
H	18.514850100	-6.253480600	4.975834600
H	17.030503800	-5.610976200	4.193742700
N	16.759995300	-7.891997800	2.982999200

C	16.556995400	-9.176997400	2.296999400
H	15.941538100	-7.379479600	3.286440800
H	15.519143200	-9.227846200	1.937046200
H	16.754273700	-10.024025800	2.976942400
H	17.253734300	-9.248136800	1.448133600
C	22.094993800	-8.659997600	2.782999200
C	23.590993400	-8.807997500	2.846999200
O	24.203993200	-8.506997600	3.885998900
H	21.833936000	-7.590801500	2.818535300
H	21.643124800	-9.120454100	3.673070600
H	21.626037600	-9.093752600	1.883382400
N	24.201993200	-9.251997400	1.756999500
C	25.655992800	-9.312997400	1.716999500
H	23.668088500	-9.468106300	0.925381700
H	26.011136900	-10.356303200	1.637287400
H	26.035922800	-8.876460900	2.649367300
H	26.044797300	-8.743101700	0.854807100

3DNX

C	32.747990800	54.889984600	7.205998000
C	33.947990500	54.141984800	6.626998100
O	34.720990300	54.702984600	5.852998400
H	33.124013100	55.745962900	7.789154500
H	32.098352400	54.274114000	7.850640600

H	32.151918300	55.291351600	6.371468700
N	34.113990400	52.885985200	7.025998000
C	35.206990100	52.048985400	6.581998200
H	33.442122000	52.485551900	7.669139800
H	34.827991400	51.126834200	6.109078100
H	35.863297900	51.771037300	7.425240100
H	35.792760200	52.616903300	5.846495100
C	37.269989500	53.286985000	1.766999500
C	37.367989500	54.763984600	1.358999600
O	37.846989400	55.098984500	0.266999900
H	36.237667300	53.013120300	2.048619500
H	37.914111400	53.086247100	2.641965000
H	37.600591400	52.673875500	0.918916900
N	36.903989600	55.675984400	2.232999400
C	37.152989600	57.096984000	1.968999400
C	36.360989800	57.803983800	3.070999100
C	36.245989800	56.844984000	4.163998800
C	36.174989800	55.469984400	3.500999000
H	38.239869900	57.295745600	2.030067200
H	36.834963400	57.364402800	0.949463500
H	35.359885100	58.060866700	2.674431500
H	36.832197200	58.757113000	3.363546200
H	37.146832300	56.872922700	4.806103500

H	35.384762700	57.000231300	4.828856200
H	36.629355700	54.676148700	4.109504700
H	35.131352000	55.170459100	3.297547600

5JUF

C	26.981992400	22.387993700	71.496979900
C	27.792992200	22.504993700	72.781979600
O	27.311992300	22.159993800	73.866979300
H	26.062923300	22.984632500	71.610410000
H	26.679627100	21.336108900	71.372354000
H	27.520848400	22.719782500	70.594274000
N	29.028991900	23.004993500	72.671979600
C	29.901991600	23.128993500	73.832979300
H	29.378893900	23.274867800	71.761743900
H	29.524214600	23.896750900	74.530341300
H	30.909179500	23.411535900	73.495898200
H	29.948596300	22.170544700	74.374254400
C	25.210992900	25.090993000	76.095978600
C	25.034993000	23.646993400	76.538978500
O	23.971993300	23.304993500	77.075978400
H	25.141820100	25.735227400	76.987066100
H	24.374702000	25.354692900	75.430323600
H	26.166586400	25.292728200	75.580697300

N	26.073992700	22.817993600	76.398978600
C	25.942992700	21.385994000	76.661978500
C	25.219992900	21.090994100	77.981978100
O	24.315993200	20.245994300	78.041978100
C	27.325992300	20.730994200	76.671978500
H	26.796119900	23.057371300	75.724124700
H	25.313341900	20.922213700	75.884873700
H	27.218780500	19.657806500	76.895381000
H	27.962358200	21.184169500	77.451636200
H	27.815393400	20.843442300	75.695186800
N	25.618992800	21.768993900	79.053977800
C	25.146992900	21.383994000	80.379977400
H	26.323631200	22.489493300	78.957332200
H	24.062754100	21.564704200	80.471488200
H	25.682376400	21.978800300	81.135535500
H	25.327829700	20.310267200	80.554130400

4Q6U

C	42.416988100	9.042997500	21.201994000
C	43.793987700	8.412997600	21.385994000
O	43.922987700	7.387997900	22.048993800
H	42.421870400	9.977093000	20.614569000
H	41.993343800	9.245832400	22.198165000

H	41.762803500	8.309097300	20.704759700
N	44.830987400	9.018997500	20.826994200
C	46.143987000	8.407997600	20.939994100
H	44.702633400	9.840226000	20.250208100
H	46.164567500	7.423359400	20.442440300
H	46.398668800	8.252724400	22.000181800
H	46.890437400	9.068511000	20.474909100
C	42.389988100	4.002998900	25.526992800
C	43.490987800	3.014999200	25.147992900
O	43.579987800	1.932999500	25.715992800
H	42.821823400	4.914097100	25.978815700
H	41.729393400	3.514604000	26.254673200
H	41.804530800	4.325544700	24.648086700
N	44.320987600	3.383999000	24.170993200
C	45.492987200	2.577999300	23.831993300
C	45.955987100	3.179999100	22.505993700
C	45.529987200	4.618998700	22.594993700
C	44.255987600	4.634998700	23.397993400
H	45.218171200	1.513974600	23.777935500
H	46.260721900	2.678266700	24.624392800
H	47.037820800	3.053201600	22.341591300
H	45.433969800	2.672196400	21.674341700
H	45.396445200	5.109087800	21.619717900

H	46.302789600	5.201055900	23.130591900
H	44.175684700	5.517646200	24.049562500
H	43.361683100	4.634508200	22.748812600

3LGI

C	-23.239993500	-4.788998700	-38.573989200
C	-23.312993500	-3.697999000	-37.500989500
O	-23.318993500	-2.521999300	-37.835989400
H	-22.335045400	-4.616141900	-39.177153300
H	-24.112904500	-4.677685800	-39.236469700
H	-23.218501100	-5.817617400	-38.174853500
N	-23.366993400	-4.066998900	-36.225989800
C	-23.444993400	-3.044999100	-35.173990100
H	-23.368694600	-5.046487800	-35.972285700
H	-24.367792800	-2.451301500	-35.282498900
H	-22.585904000	-2.359651200	-35.248066600
H	-23.440354400	-3.544166800	-34.194563500
N	-25.094993000	1.593999600	-37.988989300
C	-24.299993200	2.497999300	-38.831989100
C	-23.571993400	1.541999600	-39.779988800
C	-23.535993400	0.245999900	-39.060989000
C	-24.794993000	0.176000000	-38.253989300
H	-23.606133100	3.078014000	-38.195670500

H	-24.152505100	1.450349600	-40.717353900
H	-22.573757000	1.915300400	-40.059496700
H	-23.424001300	-0.630369300	-39.715371800
H	-22.674242900	0.213599400	-38.370227500
H	-25.627189200	-0.287237400	-38.821925000
H	-24.672603700	-0.394653500	-37.323144400
H	-24.953905900	3.225624600	-39.338155900
C	-26.703992500	1.161999700	-36.202989800
C	-25.927992700	2.124999400	-37.089989600
O	-26.063992700	3.328999100	-36.964989600
H	-27.442982400	1.738756700	-35.632227300
H	-26.022847400	0.649432500	-35.499035400
H	-27.213612500	0.377406400	-36.790339300

3A57

C	25.134992900	21.156994100	35.458990100
C	26.506992600	21.818993900	35.016990200
N	27.361992300	20.865994100	34.328990400
C	28.221992100	21.162994100	33.360990600
C	22.375993700	23.858993300	36.475989800
C	23.768993300	24.366993200	36.749989700
O	24.799993000	23.753993300	36.367989800
N	23.811993300	25.546992800	37.325989500
H	24.329161500	21.412582600	34.752897200

H	24.836362200	21.477770000	36.463030600
H	25.253217900	20.058390900	35.449566800
H	26.307625400	22.656987800	34.328313400
H	26.999027000	22.270588000	35.904517900
H	27.673355600	20.136106100	34.963310000
H	21.668269800	24.084427900	37.292650500
H	22.411574300	22.773741200	36.309845200
H	21.996287100	24.335327600	35.555031000
H	24.714576100	25.973314500	37.505301100
H	22.984561600	26.027188500	37.655996000
H	27.731479300	21.726447300	32.536435500
H	28.636488200	20.250157000	32.888966800
H	29.121952300	21.793177600	33.639254900

3DXY

C	18.182994900	-7.923997800	21.559993900
C	17.378995100	-6.903998100	20.761994200
C	20.068994400	-10.988996900	23.625993400
O	19.745994500	-9.933997200	23.083993500
C	16.325995400	-6.249998200	21.670993900
C	16.720995300	-7.563997900	19.547994500
H	15.733437700	-5.490559700	21.129850800
H	16.798587700	-5.760188200	22.540089600
H	15.625202000	-7.014314300	22.054734900

H	18.063224100	-6.111676600	20.399386000
H	16.135129800	-6.838562200	18.955522600
H	16.033867400	-8.367150100	19.873301200
H	17.475815100	-8.018422300	18.882251900
H	17.529906600	-8.727239900	21.942165300
H	18.680057100	-7.469424800	22.431943000
H	18.965957000	-8.405347000	20.951365700
C	19.033994700	-11.798996700	24.369993200
H	18.569156300	-11.156501700	25.134414400
H	19.424833700	-12.709490800	24.856853400
H	18.242609300	-12.088343300	23.659784100
N	21.304994000	-11.480996800	23.602993400
C	22.389993700	-10.832997000	22.870993600
H	21.490397200	-12.374508400	24.040988500
H	22.055519600	-9.821961800	22.601797700
H	22.636784900	-11.386570700	21.947402800
H	23.291682000	-10.768480500	23.502000100

3OAJ

C	30.572991400	12.379996500	32.942990800
C	30.868991300	13.568996200	33.782990500
C	29.279991800	8.875997500	30.918991300
O	29.703991700	9.831997200	31.557991100

N	31.577991100	7.991997800	30.521991400
C	30.201991500	8.072997700	30.018991600
H	31.872020500	8.939717700	30.775400200
H	31.578187700	7.475806100	31.404946200
H	30.214024700	8.572180600	29.031355900
H	29.819429200	7.049295700	29.852407800
N	28.009992100	8.482997600	30.961991300
H	27.666172300	7.693432400	30.430270900
H	27.342993000	9.021784200	31.503871100
N	32.470990900	13.472996200	32.073991000
C	31.456991200	12.501996500	31.687991100
C	30.651991400	12.986996400	30.485991400
C	32.076991000	14.262996000	33.256990700
H	31.923034000	11.530511000	31.441483400
H	31.306637900	13.180902200	29.616062400
H	29.901481600	12.234491300	30.186558200
H	30.816379500	11.440889600	33.465397200
H	29.513973800	12.290747600	32.662070600
H	30.009510400	14.265066800	33.780486300
H	30.994697300	13.289382700	34.843954400
H	31.863348600	15.319686700	32.993288500
H	32.908960800	14.287270500	33.988390900
H	32.793577300	14.044047700	31.298528000

H	30.127008100	13.927540300	30.734498400

5JOV

C	49.794986000	109.327969300	10.793997000
H	50.111323800	110.334464800	11.193696800
H	48.779058500	109.498065600	10.374533800
H	50.447645000	109.143064500	9.917568400
N	49.798986000	108.302969600	11.636996700
C	48.893986300	108.261969600	12.757996400
C	49.084986200	109.466969300	13.680996200
C	49.121986200	106.980970000	13.537996200
H	50.729491400	108.007500700	11.927509300
H	47.860945300	108.277452600	12.352835200
H	48.905995000	110.422942100	13.160806600
H	50.120139700	109.478858500	14.068955300
H	48.398580500	109.426012700	14.545306000
H	48.421978900	106.884557300	14.380083900
H	50.141558700	106.958778300	13.963391800
H	49.005672600	106.092157900	12.899085100
C	48.188986500	102.320971300	15.004995800
C	49.057986200	103.351971000	15.684995600
O	49.131986200	104.499970700	15.253995700
H	47.142657800	102.494231400	15.311455300

H	48.452909400	101.278014500	15.252666600
H	48.246718900	102.465082100	13.916394600
N	49.719986000	102.958971100	16.778995300
C	50.241985900	103.929970800	17.739995000
H	49.526890000	102.029937200	17.137234600
H	50.665681600	104.777522900	17.184222900
H	51.028970800	103.458017300	18.347721400
H	49.443705700	104.309876700	18.403875700

2QRU

C	23.689993400	2.820999200	17.660995000
C	24.648993100	1.920999500	16.924995200
C	21.110994100	5.201998500	20.451994300
O	21.881993900	4.600998700	19.689994500
C	19.887994400	5.935998300	19.906994400
H	20.218578000	6.629057400	19.117514300
H	19.321201600	6.498605400	20.668011100
H	19.219616800	5.194660400	19.439358500
N	21.265994000	5.243998500	21.775993900
C	22.345993700	4.544998700	22.491993700
H	20.579713800	5.738936400	22.331862700
H	21.943120300	3.705674800	23.084430100
H	22.871970200	5.244184400	23.162242100

H	23.045354500	4.155248900	21.740235200
C	23.636993400	4.223998800	17.047995200
O	25.541992800	1.322999600	17.528995100
N	24.475993100	1.817999500	15.611995600
H	24.652870400	4.644164100	16.949534600
H	23.038743300	4.888834700	17.686062100
H	23.176869800	4.224527800	16.042870100
H	22.678970700	2.374274800	17.661276100
H	24.025813600	2.868412300	18.704717400
H	23.763873200	2.332692600	15.111672200
H	25.106910900	1.228631900	15.081953800

3JQ1

C	-1.395999600	64.355981900	61.628982700
C	-1.876999500	64.480981900	60.174983100
C	-0.016000000	64.201982000	65.858981500
O	-0.488999900	64.309982000	64.708981800
C	1.494999600	64.293982000	66.097981400
H	1.862285200	63.453989900	66.713500800
H	1.748017000	65.234069300	66.620230800
H	1.993270900	64.283904900	65.119985000
N	-0.767999800	63.966982000	66.930981200
C	-2.225999400	63.773982100	66.819981200

H	-0.331609000	63.886721700	67.841099000
H	-2.745822600	64.737662500	66.682418400
H	-2.589424200	63.286739600	67.736704600
H	-2.437445400	63.140409700	65.945518600
N	-3.252999100	64.933981800	60.133983100
C	-1.762999500	63.122982300	59.458983300
S	0.312999900	63.752982100	61.701982700
H	-3.345100000	65.860821300	60.553719300
H	-3.578319200	65.013186600	59.169893000
H	-1.183088700	65.192853200	59.666078600
H	-2.199049100	63.186018500	58.444901500
H	-2.323080200	62.356769000	60.021942200
H	-0.720138500	62.796245800	59.362757100
H	-1.476062100	65.321937300	62.156438400
H	-2.018743600	63.628216800	62.173071600
H	0.735009200	64.342724200	60.549330300

4YLE

N	1.038999700	0.682999800	24.818993000
C	2.132999400	0.025000000	24.035993300
C	2.660999300	-1.221999700	24.730993100
C	3.362999100	0.918999700	23.745993300
C	3.100999100	1.968999400	22.655993600

H	0.277296700	0.013507600	24.923470300
H	1.746415100	-0.303676100	23.041318900
H	3.674289600	1.406035700	24.688357200
H	4.192670100	0.268310600	23.421295200
H	3.181894700	-0.946173800	25.666523500
H	1.859387500	-1.927163200	24.984430000
H	4.001913100	2.571626800	22.444866900
H	2.790667400	1.481275500	21.714327900
H	3.383456200	-1.738177400	24.075476400
H	2.299840200	2.669335300	22.953235900
C	-2.331999300	-1.920999500	26.288992600
C	-1.286999600	-2.464999300	25.370992900
O	-0.344999900	-1.750999500	25.000993000
H	-2.919921800	-1.161786600	25.745615800
H	-1.837880200	-1.407415200	27.128162000
H	-3.026463600	-2.679040100	26.689209500
N	-1.447999600	-3.726999000	24.980993000
C	-0.542999800	-4.304998800	24.005993300
H	-2.268461600	-4.243548200	25.269285100
H	0.493347800	-4.053822200	24.276013200
H	-0.735234600	-3.909706700	22.992658700
H	-0.664839000	-5.398205900	23.997704000
H	0.682085600	1.478500900	24.287392500

3HZ6

N	32.977990700	26.838992500	5.067998600
C	32.131991000	26.128992700	6.035998300
C	32.317990900	24.647993100	5.801998400
O	33.020990700	24.230993200	4.892998600
C	30.656991400	26.510992600	5.813998400
O	30.240991500	26.148992700	4.494998700
H	33.442987800	27.636134400	5.495262800
H	33.673745500	26.200563400	4.681030700
H	32.385572900	26.362406300	7.083730800
H	29.994975000	25.990477900	6.522449700
H	30.548768900	27.599628600	5.992400700
H	30.960255800	26.456779100	3.913051500
N	31.674991100	23.823993300	6.623998100
H	31.066099900	24.165868800	7.356924600
H	31.712114100	22.829111100	6.441023000
C	31.805991100	28.503992000	10.487997100
C	31.175991200	27.230992400	9.985997200
O	31.104991300	26.992992400	8.792997500
H	32.864914600	28.521221200	10.181040800
H	31.754360300	28.640921200	11.581024300
H	31.313812000	29.360052400	9.999108200

N	30.735991400	26.423992600	10.923996900
C	30.108991500	25.168992900	10.574997000
H	30.834871000	26.672513400	11.898929100
H	29.309599000	25.337137000	9.835213000
H	29.677757100	24.712012200	11.477221900
H	30.840647700	24.467922200	10.134691800

1RY9

C	21.443994000	63.516982200	21.020994100
O	22.542993700	63.190982300	20.576994200
C	25.215992900	62.357982500	19.132994600
C	26.616992500	62.011982600	18.618994800
C	20.807994200	62.728982400	22.157993800
H	19.805061100	63.083097700	22.447686800
H	21.475889700	62.782345100	23.032803800
H	20.745762300	61.671572700	21.856476800
N	20.752994200	64.562981900	20.568994200
C	21.238994000	65.401981600	19.482994500
H	19.854621600	64.776319300	20.982564200
H	20.481624800	66.163808600	19.252372000
H	21.433460400	64.794760900	18.583694300
H	22.181079100	65.901947700	19.766653900
C	25.663992800	64.255982000	22.428993700

C	25.869992700	63.202982300	21.333994000
O	26.674992500	62.283982500	21.500994000
H	26.610362600	64.804631500	22.562136800
H	25.444106500	63.734407100	23.373206800
H	24.858186500	64.977546400	22.209938100
N	25.182992900	63.357982200	20.205994300
O	26.938992400	60.845982900	18.420994800
C	24.337993200	62.802982400	17.978995000
H	24.338432600	63.919925900	20.235094400
H	24.841002800	61.400712300	19.532758600
H	24.690739000	63.762807500	17.562695700
H	23.295432200	62.916615400	18.309491200
H	24.374927200	62.048367100	17.178059600
N	27.431992300	63.023982300	18.368994800
H	27.236363100	63.940598400	18.756480900
H	28.397272900	62.801830600	18.155082600

3ZK4

C	-42.937987900	52.852985200	78.361978000
C	-42.433988100	51.495985500	77.900978100
O	-43.078987900	50.881985700	77.039978400
H	-42.923300200	53.536190800	77.497542900
H	-42.355778000	53.303812400	79.184864900

H	-43.986362200	52.738013500	78.677670200
N	-41.302988400	51.008985700	78.447978000
C	-40.705988600	49.762986000	78.015978100
H	-40.801179300	51.580840300	79.116020300
H	-41.458939800	48.958480600	78.034015100
H	-39.880852100	49.504226500	78.695323000
H	-40.319967900	49.837597100	76.983601700
N	-45.678987200	50.104985900	75.733978700
C	-44.766987400	49.673986100	74.693979000
C	-45.420987300	49.182986200	73.426979400
C	-43.859987700	48.555986400	75.212978900
O	-44.751987400	49.020986200	72.423979700
H	-44.133792000	50.531502600	74.406922100
H	-44.467118700	47.668872900	75.476829800
H	-43.318799800	48.905088000	76.099818700
H	-43.150000300	48.259294400	74.427363400
H	-46.106712200	49.326410300	76.233584100
H	-46.406908900	50.730284500	75.394027000
N	-46.773556400	48.908455700	73.452724600
H	-47.300151200	48.912830900	74.316099500
H	-47.170391800	48.457455000	72.637510700

4WDC

C	18.962994700	-14.407996000	67.481981100
C	18.291994900	-15.757995600	67.691981000
N	19.150994600	-16.708995300	68.461980800
C	18.985994700	-13.487996200	68.680980700
H	19.520469400	-12.547376900	68.460382200
H	19.482164900	-13.971508000	69.539817400
H	17.959139700	-13.224906400	68.995554600
H	18.445593100	-13.896803700	66.646539900
H	19.997734800	-14.581845900	67.123440400
H	17.997594600	-16.174778400	66.704475700
H	17.355608000	-15.629423100	68.265083800
H	19.995478500	-16.893627500	67.914668200
H	18.651637500	-17.596249000	68.526519000
C	16.931995200	-20.658994200	69.184980600
C	16.512995400	-19.684994500	68.091980900
O	17.084995200	-18.606994800	67.931980900
H	16.239145200	-21.505752700	69.322619200
H	17.023576000	-20.107384500	70.133482000
H	17.930975500	-21.054379500	68.935584600
N	15.520995600	-20.097994400	67.315981100
C	15.029995800	-19.343994600	66.172981400
H	15.121121300	-21.013285600	67.481114000
H	15.296012400	-18.287363100	66.315775800

H	13.935365000	-19.443492200	66.102129000
H	15.485689300	-19.696263000	65.229943500

3EEA

C	-11.172996900	-2.004999400	-7.967997800
H	-10.974996400	-3.069266500	-7.735240500
H	-11.711864500	-2.010370000	-8.958309900
H	-11.939872800	-1.663397600	-7.224686100
N	-10.046997200	-1.277999600	-7.928997800
C	-10.135997200	0.177000000	-7.924997800
C	-8.669997600	0.632999800	-7.791997800
C	-8.654997600	-1.748999500	-7.833997800
C	-7.924997800	-0.551999800	-7.305997900
H	-10.612659800	0.571983300	-8.850858000
H	-10.769704100	0.523716700	-7.073916000
H	-8.568380100	1.514428500	-7.136911700
H	-8.299111400	0.939075000	-8.788297900
H	-6.865934800	-0.557647800	-7.602776300
H	-7.942716000	-0.583250400	-6.199357900
H	-8.255646100	-2.060822000	-8.821917500
H	-8.584697700	-2.631496400	-7.175102200
C	-4.017998900	-2.929999200	-9.026997500
C	-4.966998600	-3.284999100	-7.876997800

O	-6.130998300	-2.930999200	-7.911997800
H	-4.491222500	-2.145967000	-9.631667700
H	-3.849691300	-3.816404300	-9.663855700
H	-3.033070600	-2.580671100	-8.669083900
N	-4.469998700	-4.069998900	-6.916998100
C	-5.086998600	-4.256998800	-5.617998400
H	-3.472602200	-4.253620100	-6.957637700
H	-6.173548300	-4.132888600	-5.730824800
H	-4.721440200	-3.516737800	-4.882209500
H	-4.877380900	-5.271154100	-5.242526900

4PGM

C	-4.141998800	0.929999700	17.167995200
C	-2.994999200	1.050999700	18.145994900
O	-2.374999300	0.061000000	18.519994800
H	-4.665262500	1.871990800	16.970913900
H	-4.853884700	0.195421500	17.555999600
H	-3.752677900	0.538005500	16.222408000
N	-2.720999200	2.277999400	18.578994800
C	-1.627999500	2.537999300	19.527994500
C	-2.112999400	3.549999000	20.555994200
O	-2.996999200	4.339998800	20.260994300
C	-0.351999900	3.079999100	18.794994700

C	0.784999800	3.319999100	19.798994400
C	0.116000000	2.055999400	17.754995000
H	-3.320036900	3.067039300	18.379280200
H	-1.371843400	1.579503500	19.991403100
H	-0.626078700	4.022202900	18.302160000
H	0.593332900	4.145663400	20.485997400
H	1.716778400	3.554722600	19.276845600
H	0.982408200	2.417986200	20.391139400
H	-0.651989800	1.848132100	17.004792100
H	0.367193300	1.107681600	18.241522700
H	1.007294100	2.426794800	17.235403800
N	-1.554999600	3.504999000	21.758993900
C	-1.934999500	4.431998800	22.807993600
H	-0.792450600	2.867755400	21.924569900
H	-1.423461600	5.397931800	22.704016700
H	-1.682932000	3.999365000	23.780370200
H	-3.011337700	4.610981900	22.757225700
C	1.238999700	1.656999500	25.180992900
C	2.590999300	2.383999300	25.182992900
O	3.610999000	1.776999500	25.460992900
H	0.386549400	2.301144100	24.938363000
H	1.090290400	1.218447400	26.171644800
H	1.291971400	0.839169800	24.456625400

N	2.577999300	3.690999000	24.938993000
C	3.799998900	4.498998700	24.895993000
C	3.544999000	5.710998400	25.770992800
O	2.492999300	6.336998200	25.639992800
C	4.090998900	5.004998600	23.467993400
C	4.047998900	3.903998900	22.420993700
O	3.048999100	3.735999000	21.675993900
N	5.122998600	3.149999100	22.346993700
H	1.722008900	4.185777600	24.733819700
H	4.606290400	3.859070800	25.264940500
H	3.323764700	5.734637200	23.199037700
H	5.064919800	5.512160600	23.463689800
H	5.852301600	3.190365700	23.041493700
H	5.098157700	2.336688200	21.748421200
N	4.482998700	6.037998300	26.665992500
C	4.325998800	7.196998000	27.553992300
H	5.278308900	5.432019400	26.795674100
H	3.972441600	8.051430900	26.972094000
H	5.292624200	7.429921000	28.005985000
H	3.592457300	6.990970900	28.341589700

2O7C

C	43.145987900	194.420945400	50.609985800

C	43.023987900	195.943945000	50.630985800
O	42.411988100	196.510944800	51.540985500
H	42.593917400	193.993333400	49.764270100
H	42.720910200	194.042597300	51.539444700
H	44.188387600	194.094334400	50.520672700
N	43.607987800	196.630944800	49.633986100
C	44.362987500	196.067945000	48.507986400
C	45.860987100	196.002945000	48.778986300
O	46.598986900	195.315945200	48.071986500
C	44.053987600	197.028944700	47.365986700
C	43.779987700	198.370944300	48.063986500
C	43.522987800	198.099944400	49.537986100
H	44.040940800	195.050746800	48.283175200
H	44.870296600	197.074588400	46.641561900
H	43.160429300	196.679813800	46.838414500
H	42.929895700	198.882669300	47.603427500
H	44.639466700	199.039831400	47.947180900
H	44.252882700	198.591327400	50.194133600
H	42.534064700	198.409960400	49.884892700
N	46.297987000	196.720944800	49.805986000
C	47.709986600	196.783944800	50.145985900
H	45.633275900	197.110502300	50.453677900
H	48.040282600	195.880507400	50.674014900

H	47.887725000	197.655735300	50.780700600
H	48.297854100	196.869552000	49.228908400
C	50.546985800	189.407946800	46.655986900
C	49.769986000	190.440946500	45.849987100
O	48.999986200	190.097946600	44.958987400
H	49.830514400	188.752595300	47.161209500
H	51.222201300	189.841824800	47.401570300
H	51.124492700	188.789873200	45.961856400
N	49.976986000	191.710946200	46.164987000
C	49.348986100	192.771945900	45.402987300
C	50.542985800	193.589945700	44.936987400
O	51.518985500	193.712945600	45.658987200
C	48.438986400	193.603945700	46.288987000
H	50.729762900	191.994994600	46.775614700
H	48.805016200	192.311939000	44.575039900
H	47.984478200	194.434005300	45.737932200
H	48.997320400	194.027561200	47.128392200
H	47.631161300	192.987580100	46.691649800
N	50.497985800	194.111945500	43.721987700
C	51.597985500	194.927945300	43.237987900
H	49.689670400	193.964577200	43.138413000
H	51.740834400	195.805550400	43.877695100
H	51.380588900	195.253947900	42.217815800

H 52.530233400 194.353903700 43.244901000

1DDX

C -5.652998400 45.283987300 59.097983400

C -5.574998400 46.452987000 60.088983100

O -6.426998200 47.335986700 60.100983100

H -6.636066200 44.817542500 59.197483500

H -4.876183500 44.527391000 59.244160800

H -5.577701000 45.685635400 58.083898100

N -4.546998700 46.450987000 60.925982900

C -4.363998800 47.527986700 61.883982600

C -3.254999100 48.426986400 61.375982800

O -2.125999400 47.986986500 61.187982800

C -3.953998900 47.002986800 63.252982200

C -4.916998600 45.897987100 63.686982100

C -3.933998900 48.151986500 64.249982000

C -4.292998800 44.866987400 64.609981900

H -3.780215200 45.805445100 60.808828400

H -5.325505500 48.035607900 61.955137600

H -2.942620700 46.585481300 63.156375400

H -5.309677600 45.376429100 62.808728500

H -5.790774300 46.358630000 64.171039300

H -3.193595500 48.915740500 63.986374400

H	-3.690271200	47.790211700	65.254848300
H	-4.919591200	48.631747800	64.300997500
H	-5.015202800	44.107979300	64.926001900
H	-3.886776700	45.316866700	65.523480300
H	-3.465545500	44.342875300	64.117435800
N	-3.555999000	49.705986000	61.153982800
C	-4.813998600	50.438985800	61.362982800
H	-2.759578800	50.253688600	60.852982700
H	-4.713080800	51.422135500	60.898301400
H	-5.029862800	50.575135100	62.429508000
H	-5.649296200	49.914052800	60.889602200
C	-1.160999700	45.950987100	69.094980600
C	-1.695999500	44.547987500	69.359980500
O	-2.288999400	44.288987600	70.412980200
H	-0.489836900	46.221330300	69.915215500
H	-2.001711000	46.651093300	69.112083700
H	-0.625830100	46.051974600	68.144585800
N	-1.503999600	43.644987700	68.404980800
C	-1.978999400	42.289988100	68.579980800
C	-2.650999300	41.815988300	67.318981100
O	-3.088999100	42.634988000	66.495981300
H	-1.231275200	43.928267700	67.476020100
H	-1.156978500	41.620890200	68.876751200

H	-2.701476900	42.298824400	69.405616900
N	-2.703999200	40.493988600	67.168981100
C	-3.321999100	39.844988800	66.016981500
H	-2.289461500	39.907771400	67.877924100
H	-3.850065700	40.609367700	65.445205700
H	-4.031987500	39.078929500	66.345015000
H	-2.564713900	39.379579800	65.375060800