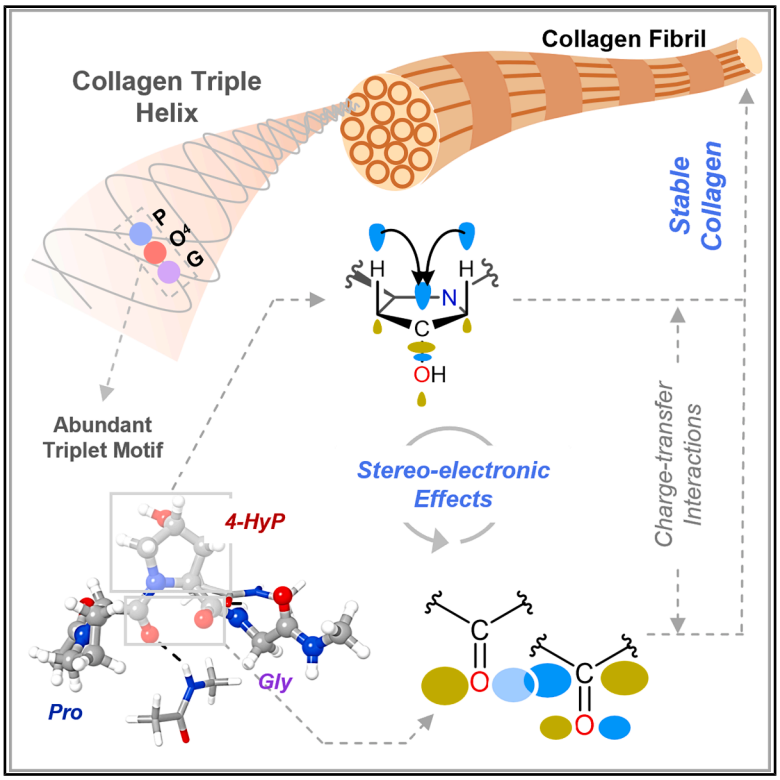
iScience

Article

Decoding elegant interplay among different stereo-electronic effects due to the ancient prolyl-4-hydroxylation stabilizing collagenous helicity

Graphical abstract



Authors

Ashutosh Joshi, Trayambak Basak, Bhaskar Mondal

Correspondence

trayambak@iitmandi.ac.in (T.B.), bhaskarmondal@iitmandi.ac.in (B.M.)

In brief

Physical chemistry; Quantum chemistry; Quantum chemical calculations

Highlights

- Deciphered the stereoelectronic effects of prolyl-4-hydroxylation in collagen
- Prolyl-4-hydroxylation-mediated *exo* ring pucker tunes Bürgi-Dunitz trajectory
- 4-HyP enhances the peptide backbone stabilization through n→π* interaction
- Decoded elegant interplay between geometry, bond isomerization, and charge-transfer

Joshi et al., 2025, iScience *28*, 112393
May 16, 2025 © 2025 The Author(s). Published by Elsevier Inc.
https://doi.org/10.1016/j.isci.2025.112393

CellPress

iScience

CellPress
OPEN ACCESS

Article

Decoding elegant interplay among different stereo-electronic effects due to the ancient prolyl-4-hydroxylation stabilizing collagenous helicity

Ashutosh Joshi,[1] Trayambak Basak,[1,*] and Bhaskar Mondal[2,3,*]
[1]School of Biosciences and Bioengineering, Indian Institute of Technology Mandi, Mandi, Himachal Pradesh 175075, India
[2]School of Chemical Sciences, Indian Institute of Technology Mandi, Mandi, Himachal Pradesh 175075, India
[3]Lead contact
*Correspondence: trayambak@iitmandi.ac.in (T.B.), bhaskarmondal@iitmandi.ac.in (B.M.)
https://doi.org/10.1016/j.isci.2025.112393

SUMMARY

Prolyl-4-hydroxylation is an ancient evolutionarily conserved post-translational modification (PTM) critical for both structural and regulatory functions in multicellular life forms. This PTM plays a pivotal role in stabilizing collagen's triple helix by influencing the puckering of the pyrrolidine ring. The elegant interplay between ring pucker, torsional angles, peptide bond isomerization, and charge-transfer interactions (O···C=O n→π* and σ→σ*) attaining the helical stability remains underappreciated. Using density functional theory calibrated against gold standard *ab initio* methods, we analyzed a physiologically relevant collagenous peptide proline-4-hydroxyproline-glycine (PO^{4}G) to establish the correlation between stereo-electronic effects due to prolyl-4-hydroxylation. Our results show that 4(*R*)-hydroxylation promotes an *exo* ring pucker, optimizing main-chain torsional angles for a stable *trans* peptide bond and maximizing the n→π* interaction ($E_{n\rightarrow\pi^*}$ = 0.9 kcal/mol) by tuning Bürgi-Dunitz trajectory, and maximizes σ→σ* interactions between axial C–H σ-electrons and C–OH* orbitals of the pyrrolidine ring. This study reveals the intricate stereo-electronic effects driving collagen's structural stability.

INTRODUCTION

Evolutionarily conserved post-translational modifications (PTMs) of collagens have been instrumental in regulating the cell-extracellular matrix (ECM) interactions and maintaining the structural integrity of the ECM.[1,2] Type I collagen is the most abundant ECM protein of multicellular vertebrate life forms, including humans.[3] The trimeric collagen I protein comprises three polypeptide chains, which can be homotrimer and/or heterotrimer, each conforming to the left-handed helical structure similar to the polyproline type II (PPII) helix.[3,4] The three polypeptide chains, arranged in one amino acid staggered manner, wrap around the central fibril axis to construct a right-handed triple helical structure (Figure 1A). Each helical chain of type I collagen is predominantly arranged by the repetitive triplet –Xaa–Yaa–Gly (XYG).[3–5] The highly abundant prolyl-4-hydroxylation PTM, catalyzed by prolyl-4-hydroxylases, is prevalent in the Yaa position of the collagenous triplet.[5] Specifically, the amino acid hydrolysis of fibrillar and non-fibrillar collagen chains from human tissue sources showed that approximately 38% of Yaa amino acids are 4-hydroxyproline (4-HyP).[6] The occurrence of 4-HyP in the repetitive triplet in the form of –Xaa–4-HyP–Gly– (XO^{4}G) provides the foundational framework to achieve the typical helical structural attainment of type I collagen and its thermal stability (Figure 1A).[7] Commonly, proline (Pro) occupies the Xaa position to construct the most abundant triplet in collagen I, –Pro–4-HyP–Gly– (PO^{4}G).[5,6]

The structural integrity of the triplet PO^{4}G repeats is maintained by different highly synchronized stereoelectronic parameters that involve pyrrolidine ring puckers in prolines, main-chain torsional angles (ϕ, ψ, and ω), charge transfer interactions, and peptide bond *cis*/*trans* conformation (Figure 1B).[8] The post-translationally modified 4-HyP in PO^{4}G attains a C$^\gamma$-*exo* ring pucker owing to the hydroxylation at the C4 or γ-carbon of the pyrrolidine ring. The resulting *exo* ring pucker is 0.6 kcal/mol ($\Delta E_{endo-exo}$) more stable over the *endo* ring pucker.[9] This stabilization effect has been attributed to the "gauche" effect facilitated by a charge-transfer interaction (σ→σ*) between the σ-bonding electrons of C$^\beta$–H$_{ax}$/C$^\delta$–H$_{ax}$ bonds and σ-antibonding (σ*) orbital of C$^\gamma$–OH (Figure 1B).[9] To gain a deeper understanding of this effect, more electronegative groups like –F, –Cl, and –SH were incorporated in synthetic peptidomimetics at the C$^\gamma$ position. The electronegative groups were found to enhance the charge-transfer interaction, further stabilizing the *exo* ring pucker.[8,10] Interestingly, the fluoroproline 4(*R*)-FlP-*exo* single amino acid peptidomimetics was shown to possess the highest $E_{\sigma\rightarrow\sigma^*}$ value among all the four conformers.[8] However, electronic structure analysis for the four conformers due to the naturally occurring C$^\gamma$–OH hydroxylation in 4-HyP in collagen chains, i.e., 4(*R*)-HyP-*exo*, 4(*R*)-HyP-*endo*, 4(*S*)-HyP-*exo*, and 4(*S*)-HyP-*endo* (Figure 1B) has not been reported thus far. Taken together, the positional preferences of prolines in collagen triple helix have led to the propensity-based hypothesis suggesting

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

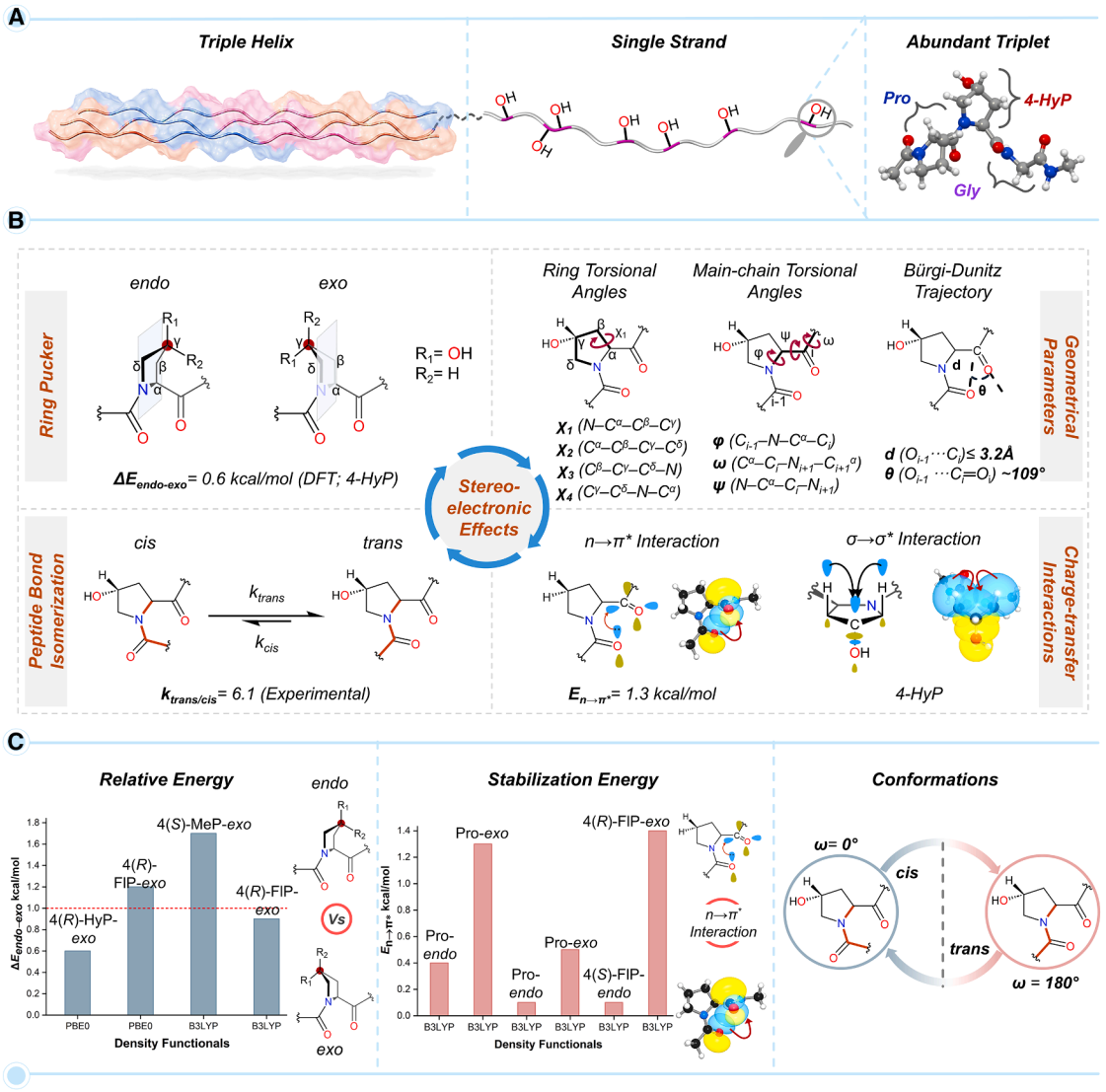


Figure 1. Stereoelectronic effects in collagen due to 4-HyP and quantum chemical studies of stabilization energies

(A) The triple helical structure of collagen, a single strand of collagen, and the abundant triplet (–Xaa–4-HyP–Gly–). The sections in pink represent the –Xaa–4-HyP–Gly– triplet motif in a single strand in the triple helical domain.

(B) The stereoelectronic parameters associated with 4-HyP responsible for collagenous helical stability; pyrrolidine ring pucker, key geometrical parameters (ring (χ) and main-chain (ϕ, ψ, and ω) torsional angles and Bürgi-Dunitz trajectory), peptide bond conformations, and charge-transfer (n→π* and σ→σ*) interactions.

(C) An illustration of the reported $\Delta E_{endo-exo}$ (kcal/mol) and $E_{n\rightarrow\pi^*}$ (kcal/mol) values for single amino acid conformers using DFT methods.

that prolyl-4-hydroxylation at the Yaa position restricts the conformational flexibility of proline adopting the C^{γ}-*exo* pucker Consequently, Pro and 4-HyP pre-organize a single strand to attain a folded triple helix.[7,11] The restriction in conformational flexibility of the pyrrolidine ring would certainly imply that the ring torsional angles (χ_1, χ_2, χ_3, χ_4) would be restricted to a particular value, which in turn would dictate the main-chain torsional angles (ϕ, ψ, and ω) (Figure 1B).[12] Indeed, as revealed by crystallography studies, the angles χ_1 and ϕ are required to attain approximately −20° and −60°, respectively, for the C^{γ}-*exo* pucker for 4-hydroxyproline in the PO^4G triplet.[7,13] More importantly, the main-chain torsional angle ω with a value of ∼180° in PO^4G triplet significantly contributes to the collagenous helical stability by enforcing all peptide bonds to their *trans* conformation (Figure 1B).[3,14] In this *trans* peptide bond conformation, two carbonyl groups are orientated in such a way that the Bürgi-Dunitz trajectory (Figure 1B), distance from O_{i-1} to C_i (*d*), and the angle formed between O_{i-1} … C_i = O (*θ*), attain optimal values of ∼3.2 Å and ∼109°, respectively.[15–17] This specific orientation of the carbonyl groups promotes a crucial charge-transfer interaction (n→π*) between the donor lone-pair electron (n) of one carbonyl oxygen (–C=O_{i-1}) and the acceptor antibonding C=O π-orbital (π*) of another carbonyl moiety (–C=O_i) (Figure 1B).[17] Such n→π* interactions are associated with a

typical stabilization energy of $E_{n\rightarrow\pi^*} \sim$ 0.3–0.7 kcal/mol in proteins.[17] It is important to mention here that the n→π* interaction has made the acyl groups of the helical backbone resistant to acid hydrolysis, which has preserved collagen through evolution.[2] Remarkably, the n→π* interaction is larger in 4(*R*)-HyP-*exo* as compared to the 4(*R*)-HyP-*endo* conformer with an $E_{n\rightarrow\pi^*}$ value of 1.3 kcal/mol.[18] In this line, an estimation of the $E_{n\rightarrow\pi^*}$ values for 4(*S*)-HyP-*exo* and 4(*S*)-HyP-*endo*, along with their 4(*R*) diastereoisomers, would provide valuable insights into explaining the naturally occurring conformer −4(*R*)-HyP– in collagen chains.

The above discussion suggests that the crucial stereoelectronic effects involving pyrrolidine ring puckers, torsional angles (ϕ, ψ, and ω), peptide bond *cis/trans* conformation, and charge-transfer interactions (n→π* and σ→σ*) operate in a highly orchestrated manner to provide the overall stability of the collagen chain and triple helix. However, the role of the elegant interplay between these stereoelectronic effects in dictating the collagen helical stability has remained underappreciated since all previous reports involve a single amino acid model for quantum chemical investigations.[8,9,19–22] To appropriately decipher the correlation between different stereoelectronic effects due to prolyl-4-hydroxylation stabilizing the collagenous helicity, a physiologically relevant collagenous tripeptide would be essential. Moreover, one of the most critical stabilizing factors to helical stability originating from the C^{γ}-*exo* ring pucker, $\Delta E_{endo-exo}$ ranges between ∼0.6 and ∼1.7 kcal/mol (Figure 1C),[5,9,23] which is close to the accuracy limit (∼2–3 kcal/mol)[24] of modern density functional theory (DFT) methods. This necessitates rigorous calibration of the chosen DFT methods against the gold-standard *ab initio* quantum chemical methods, which offer significantly higher chemical accuracy, typically within ∼1 kcal/mol.[25,26] In this pursuit of understanding the critical role of stereoelectronic effects in collagenous helical stability, we employed DFT methods–calibrated against *ab initio* quantum chemical methods–on a physiologically relevant collagenous peptide (PO^4G). Specifically, a wide range of DFT methods involving 24 DFT functionals have been calibrated against the gold standard *ab initio* method, domain-based pair natural orbital coupled cluster methods with single, double, and perturbative triples corrections (DLPNO-CCSD(T)). Subsequently, the most optimal DFT method has been used to obtain the stereoelectronic parameters, $\Delta E_{endo-exo}$, $E_{n\rightarrow\pi^*}$, and $E_{\sigma\rightarrow\sigma^*}$ with the physiologically relevant collagenous peptide. Overall, this study has comprehensively delineated the electronic level understanding of the role of prolyl-4-hydroxylation in the helical collagenous structural attainment that is the fundamental requirement for all vertebrate life forms.

RESULTS AND DISCUSSION

Calibration of DFT functionals against *ab initio* methods for relative energy to study the stereoelectronic effects of collagen prolyl-4-hydroxylation

Collagen helical structure and stability are critically dependent on the 4-HyP ring conformer, which makes the relative energy between *endo* and *exo* ring pucker ($\Delta E_{endo-exo}$) one of the most crucial deterministic parameters in structural stability. Considering the low magnitude of $\Delta E_{endo-exo}$, typically between 1 and 2 kcal/mol, it is of utmost importance to ensure the accuracy of the chosen DFT method for the investigation of the stereoelectronic effects of prolyl-4-hydroxylation. In this quest, we begin with the direct comparison of the relative energy due to pyrrolidine ring conformations, $\Delta E_{endo-exo}$ for the 4(*R*)-HyP-*endo* relative to the natural 4(*R*)-HyP-*exo* conformer, estimated by DLPNO-CCSD(T) and 2nd-order Møller–Plesset perturbation theory (MP2) methods (Figure S1). Pleasingly, the $\Delta E_{endo-exo}$ value of 1.7 kcal/mol predicted at the MP2 level was found to be nearly identical to that of the gold-standard *ab initio* method, DLPNO-CCSD(T) producing a $\Delta E_{endo-exo}$ value of 1.8 kcal/mol. This result suggests that the geometry and relative energy at the MP2 level can be used as the benchmark for further calibration of the DFT methods. Figure 2B presents a direct comparison of the MP2-predicted $\Delta E_{endo-exo}$ values with twenty-four DFT functionals. As it appears, the DFT functionals, CAM-B3LYP, revTPSS, M06L, MN15L, M052X, M06, M062X, MN15, B2PLYP, DSD-PBEP86, and mPW2PLYP, show closest match with the MP2-estimated relative energies. Moreover, these DFT functionals also predict the $\Delta E_{endo-exo}$ relative energy around or slightly above the chemical accuracy limit of ∼1 kcal/mol (Figure 2B). Thus, this result provides an initial screening of the DFT methods against the *ab initio* MP2, and thereby, the DLPNO-CCSD(T) method.

The DFT functional calibration study was further extended to the diastereomer of the 4(*R*)-HyP, i.e., 4(*S*)-HyP. Both the *endo* and *exo* conformer of the 4(*S*)-HyP, i.e., 4(*S*)-HyP-*endo* and 4(*S*)-HyP-*exo* were subjected to the DFT functional calibration study. The relative electronic energies (ΔE) of all the non-natural HyP conformers, i.e., 4(*S*)-HyP-*endo* ($\Delta E_{endo-exo} = E$[4(*S*)-*endo*] – *E*[4(*R*)-*exo*]), 4(*S*)-HyP-*exo* ($\Delta E_{exo-exo} = E$[4(*S*)-*exo*] – *E*[4(*R*)-*exo*]), and 4(*R*)-HyP-*endo* ($\Delta E_{endo-exo}$ = *E*[4(*R*)-*endo*] – *E*[4(*R*)-*exo*]) together relative to the natural 4(*R*)-HyP-*exo* conformer at the DFT and MP2 levels are presented in the supplemental information (Figure S2; Table S1). A close inspection of the relative energy comparison reveals DFT functionals, revTPSS, M06L, M052X, M06, M062X, B2PLYP, DSD-PBEP86, and mPW2PLYP, unanimously predict $\Delta E_{endo-exo}$ close to the *ab initio* MP2 values and above the chemical accuracy limit of ∼1 kcal/mol. The DFT-M06 functional predicts a $\Delta E_{endo-exo}$ value of 1.4 kcal/mol for the 4(*R*)-HyP conformer, which is the closest to the corresponding MP2 value (Figures 2B and S2). Based on the data presented in Figures 2B and S2, the calibrated DFT functionals having closer $\Delta E_{endo-exo}$ values to MP2 can be arranged in the decreasing order of their accuracy in predicting the relative energies as M06 > M052X ≈ M06L > M062X > DSD-PBEP86 ≈ mPW2PLYP > B2PLYP ≈ revTPSS. Mostly, the hybrid DFT functionals predict relative energy values that are closer to the *ab initio* methods. In this context, it should be noted that the results obtained from the hybrid DFT methods significantly depend on the amount of the Hartree-Fock exchange (%HF-X) correlation that constitutes the functional. Keeping this in mind, we evaluated the role of %HF-X in predicting the ΔE values for all non-natural conformers, i.e., 4(*S*)-HyP-*endo* ($\Delta E_{endo-exo}$), 4(*S*)-HyP-*exo* ($\Delta E_{exo-exo}$), and 4(*R*)-HyP-*endo* ($\Delta E_{endo-exo}$) relative to the natural 4(*R*)-HyP-*exo* conformer (Figure S3; Table S2). Specifically, twelve DFT functionals with varying %HF-X between 0 and 56% were calibrated against the MP2 level. Among the previously screened functionals, only five DFT functionals satisfy the chemical accuracy mark, as shown with the

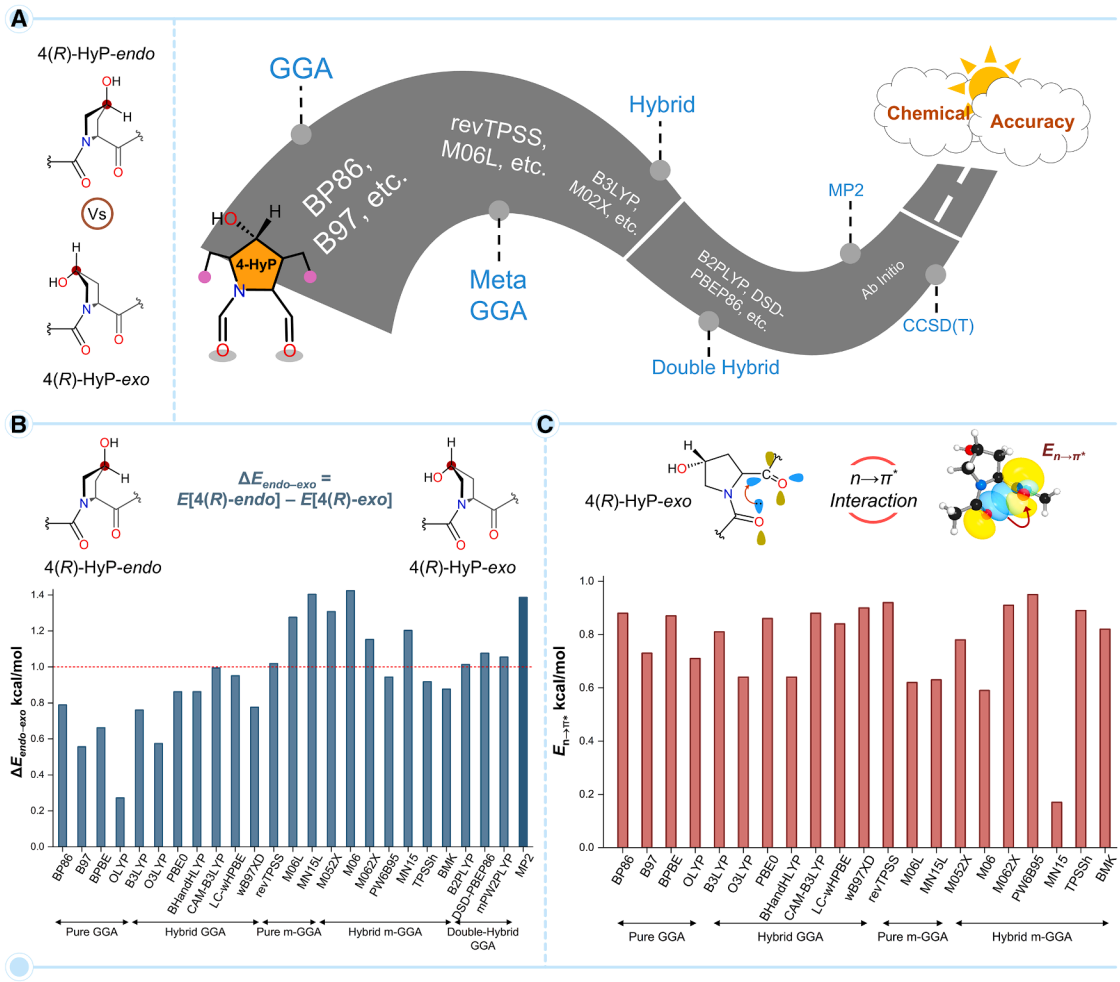


Figure 2. Calibration of DFT methods against *ab initio* methods

(A) An illustration of the category of DFT methods that are calibrated against *ab initio* MP2 and the DLPNO-CCSD(T) methods using the relative energy $\Delta E_{endo-exo}$ = E[4(*R*)-*endo*] – E[4(*R*)-*exo*] (kcal/mol) for 4-HyP conformers as the parameter.

(B) Calibration of twenty-four DFT functionals including, pure GGA, hybrid GGA, pure *meta*-GGA, hybrid *meta*-GGA, and double hybrid functionals against the *ab initio* MP2 method. See also Figure S2 and Table S1.

(C) Estimation of the interaction energy due to the n→π* charge-transfer ($E_{n\rightarrow\pi^*}$) for the natural conformer of 4-HyP, i.e., 4(*R*)-HyP-*exo* at different DFT levels. See also Figure S5 and Table S4.

$\Delta E_{endo-exo}$ values of 4(*R*)-HyP-*endo* (Figure S3). Also, in this case, the $\Delta E_{endo-exo}$ value of 1.4 kcal/mol predicted by the M06 functional (27% HF-X) is the closest to the MP2 value. In comparison, the M052X functional with the highest amount of %HF-X of 56% predicts a close value of 1.3 kcal/mol. Therefore, the DFT functionals that predict the $\Delta E_{endo-exo}$ values to the chemical accuracy mark can be further screened and arranged in the decreasing order of accuracy as, M06 > M052X > M062X > DSD-PBEP86 > B2PLYP.

Evaluation of n→π* charge-transfer interaction energy, $E_{n\rightarrow\pi^*}$ at different DFT levels to select the optimal DFT method for studying the electronic effect of collagen prolyl-4-hydroxylation

In this section, we analyzed the second-order perturbation energy ($E^{(2)}$) obtained from the natural bond orbital (NBO) analysis. The carbonyl groups (–C=O) in collagen, as in other proteins, play a vital role in maintaining overall structural stability.[27,28] The vacant π* orbital of the carbonyl group readily interacts with nucleophiles, such as the lone pair (n) of adjacent oxygen atoms. The underlying mechanism of this interaction can be understood as a charge-transfer process between the "donor" oxygen lone-pair orbital (n) and the "acceptor" C=O π* orbital. The resulting orbital mixing between n and π* exerts a stabilizing effect on the overall molecular framework. In the context of protein structure, the stabilizing n→π* charge-transfer interaction is significantly influenced by the distance between the donor oxygen of one C=O and the acceptor carbon of the other C=O moiety. The energy ($\Delta E^{(2)}_{i\rightarrow j}$) associated with the donor (*i*) → acceptor (*j*) charge-transfer interaction can be estimated using the following equation and obtained through NBO analysis.[29,30]

$$\Delta E_{i\rightarrow j}^{(2)} = -2\frac{\langle \sigma_i|\widehat{F}|\sigma_j^*\rangle^2}{\epsilon_j^* - \epsilon_i}$$

Here, $\widehat{F}$ is the effective orbital Hamiltonian (Kohn-Sham operator in DFT) and $\epsilon_i = \langle \sigma_i|\widehat{F}|\sigma_i\rangle$, $\epsilon_j^* = \langle \sigma_j^*|\widehat{F}|\sigma_j^*\rangle$ are the respective orbital energies of "donor" and "acceptor" NBOs. Using perturbation theory, the strength of the charge-transfer interaction $q_{i\rightarrow j}$ can be approximately estimated as

$$q_{i\rightarrow j} \propto \Delta E_{i\rightarrow j}^{(2)}$$

Therefore, we used the calculated energy $\Delta E_{i\rightarrow j}^{(2)}$, denoted as $E_{i\rightarrow j}$ in the manuscript, to estimate the strength of charge-transfer interactions in collagenous tripeptides. This stabilization energy due to n→π* charge-transfer interaction has been denoted as $E_{n\rightarrow\pi^*}$ throughout our discussion. The natural HyP conformer, i.e., 4(*R*)-HyP-*exo*, was subjected to the $E_{n\rightarrow\pi^*}$ analysis at different DFT levels (Figure 2C). Since no *ab inito* method could be used as a reference for the calculation of the $E_{n\rightarrow\pi^*}$ value, we focused on the primarily screened density functionals that were calibrated against *ab initio* methods based on the $\Delta E_{endo-exo}$ values. For completion, the $E_{n\rightarrow\pi^*}$ values obtained at the other pure-GGA, hybrid-GGA, pure *meta*-GGA, and hybrid *meta*-GGA DFT functionals are also presented together. However, the double-hybrid GGA functionals could not be used as they are not yet implemented in the NBO 7.0 program. Because $E_{n\rightarrow\pi^*}$ quantifies the stabilization effect due to charge-transfer interaction and a higher value signifies greater stabilization, the DFT functional that gives a higher $E_{n\rightarrow\pi^*}$ can be selected as the most suitable one among the selected functionals. Figure 2C presents the $E_{n\rightarrow\pi^*}$ values for the 4(*R*)-HyP-*exo* conformer obtained from the NBO analysis with different DFT functionals. The result shows that PW6B95 predicts the highest $E_{n\rightarrow\pi^*}$ value of 1.0 kcal/mol as compared to the other functionals. The density functionals that estimate the second (0.9 kcal/mol) highest $E_{n\rightarrow\pi^*}$ values are revTPSS and M062X, respectively. It should be noted that the PW6B95 method predicted a $\Delta E_{endo-exo}$ value lower than the chemical accuracy limit for the 4(*R*)-HyP-*endo* conformer in our primary calibration study, and therefore, we discarded this DFT functional for the present analysis of stereoelectronic effect together. Similarly, the revTPSS functional, although predicted an appreciably high $E_{n\rightarrow\pi^*}$ value of 0.9 kcal/mol, failed to estimate an accurate relative energy ($\Delta E_{endo-exo}$) value close to the MP2 value (*vide supra*). Therefore, the revTPSS functional is also not a good choice for accurately predicting both relative energy and $E_{n\rightarrow\pi^*}$ values together. This leaves us with the M062X hybrid *meta*-GGA functional among the eight primarily short-listed DFT functionals, which can estimate a reasonably accurate $\Delta E_{endo-exo}$ value (1.1 kcal/mol) along with an appreciably high $E_{n\rightarrow\pi^*}$ value (0.9 kcal/mol). Therefore, M062X can be the method of choice for the electronic structure and energy calculation of the proline and 4-hydroxyproline conformers.

To provide a deeper insight into the charge-transfer interaction present in the 4(*R*)-HyP-*exo* moiety, we analyzed the occupation numbers (O.N.) of the donor (e.g., n_{CO}) and acceptor (π^*_{CO}) orbitals. The O.N. of the donor n_{CO} orbital significantly decreased to 1.89 from its ideal value of ∼2.0, whereas the same for the acceptor orbital appreciably increased to 0.28 from its ideal value of ∼0.0 (Table S3). However, it is rather challenging to directly correlate the occupation numbers of the donor and acceptor orbitals to the $E_{n\rightarrow\pi^*}$ values, as other auxiliary charge-transfer processes, such as $n_{CO}\rightarrow\sigma^*_{NC}$ and $n_{NH}\rightarrow\pi^*_{CO}$ associated with the amide bonds also contribute to the O.N. To verify this, we analyzed the O.N. of the associated orbitals in the single amino acid 4(*R*)-HyP-*exo* model and a structure (Structure-1) where the –NH(Me) moiety of one of the amide bonds is replaced with –H (Figure S4). This way, using Structure-1, we target to better estimate the O.N. associated with the primary $n_{CO}\rightarrow\pi^*_{CO}$ charge transfer. Indeed, in Structure-1, the O.N. of the donor oxygen lone pair orbital (1.90) and that of the acceptors π^*_{CO} (0.04) and σ^*_{NC} (0.06) is summed to ∼2.0 (Table S3).

With the single amino acid model used over the decades to understand the role of 4-hydroxylation in collagen structure and stability, we anticipated the $E_{n\rightarrow\pi^*}$ value to be the highest for the natural 4(*R*)-HyP-*exo* conformer. To validate this notion, we used our DFT functional of choice, i.e., M062X in conjunction with Pople's triple-ζ basis set 6-311+G(d,p) to perform geometry optimization followed by NBO analysis on all the conformational isomers of the single amino acid model (Figure 3A). Surprisingly, a higher $E_{n\rightarrow\pi^*}$ value was obtained for two out of the three non-natural conformers, 4(*R*)-HyP-*endo* ($E_{n\rightarrow\pi^*}$ = 1.1 kcal/mol) and 4(*S*)-HyP-*exo* ($E_{n\rightarrow\pi^*}$ = 1.0 kcal/mol), as compared to the natural conformer, 4(*R*)-HyP-*exo* ($E_{n\rightarrow\pi^*}$ = 0.9 kcal/mol). As such, the decreasing order of $E_{n\rightarrow\pi^*}$ values in 4-HyP conformers is 4(*R*)-HyP-*endo* > 4(*S*)-HyP-*exo* > 4(*R*)-HyP-*exo* > 4(*S*)-HyP-*endo*. To check whether this result is an artifact of the chosen DFT method, we calculated the $E_{n\rightarrow\pi^*}$ values with the complete set of twenty-one functionals as presented in Figure 2C. Interestingly, sixteen out of twenty-one functionals including M062X predicted a higher $E_{n\rightarrow\pi^*}$ value for a non-natural conformer than the natural one, and the rest five functionals predicted the $E_{n\rightarrow\pi^*}$ value of one of the three non-natural conformers equal to the natural one (Figure S5; Table S4). This unexpected observation led us to revisit the amino acid model used in the literature to investigate the role of the stereoelectronic effect in collagen structure.

Effect of solvent on geometry in understanding the role of prolyl-4-hydroxylation

Most quantum chemical studies in the literature on proline derivatives were performed with the single amino acid model and the geometry obtained in the gas-phase optimization.[8,9,19–22] The subsequent NBO analyses were also performed on the gas-phase geometries. To shed light on the effect of the solvent on the amino acid models used, we performed the geometry optimization, frequency calculation, and NBO analysis in both the gas- and solvent-phase (water) at the M062X/6-311+G(d,p) level of theory. The geometrical parameters, i.e., the main-chain torsional angles (ϕ,ψ) associated with the n→π* charge-transfer interaction, were compared for the gas- and solvent-phase geometries (Figure 3B). The ϕ and ψ values for the gas-phase optimized geometry, −85.2° and 78.9°, respectively, were found to be far away from the expected values, ϕ = −60 ± 7° and ψ = 150 ± 9°.[31] This results in a significant distortion in the Bürgi-Dunitz trajectory (d = 3.26 Å and θ = 158.7°), which is far off from the ideal value of θ ∼109°. As a consequence, the structure

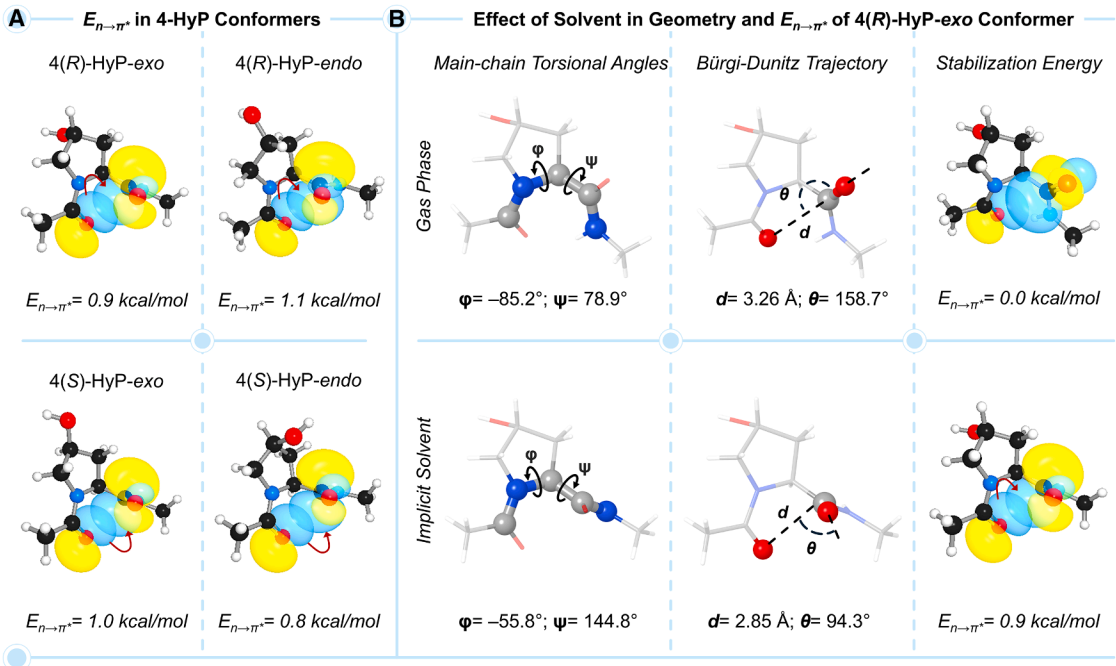


Figure 3. Geometry and natural bond orbital analysis of 4-HyP conformers

(A) Natural bond orbitals and $E_{n\to\pi^*}$ values in kcal/mol for four different conformers of 4-HyP.

(B) Gas- and solvent-phase geometries of the single amino acid model of the natural 4(*R*)-HyP-*exo* conformer along with the natural bond orbitals and $E_{n\to\pi^*}$ values. See also Figure S6.

displayed an $E_{n\to\pi^*}$ value of 0.0 kcal/mol, showing no n→π* interaction (Figure 3B). In contrast to the gas-phase geometry, the solvent-phase optimized geometry of the natural conformer, i.e., 4(*R*)-HyP-*exo* single amino acid model features the main-chain torsional angles of $\phi = -55.8^{\circ}$ and $\psi = 144.8^{\circ}$. In addition, the geometrical parameters associated with the Bürgi-Dunitz trajectory, d = 2.85 Å, $\theta = 94.3^{\circ}$ get closer to the ideal value and yield an $E_{n\to\pi^*}$ value of 0.9 kcal/mol (Figure 3B). This result nicely showcases the importance of the solvent-phase geometry in understanding the role of prolyl-4-hydroxylation. The solvent appears to have a more pronounced effect on the n→π* charge-transfer interaction than on the *endo/exo* puckering ($\Delta E_{endo-exo}$). This explains why the gas-phase geometries of the single amino acid model used in the literature to investigate the relative energy due to ring puckering were somewhat reasonable. However, for examining the correlation between the geometry and n→π* charge-transfer interaction, the inclusion of solvent becomes critical as the charge-transfer effects are best captured in the presence of a polar solvent dielectric environment.[32] To better understand the effect of solvent dielectric environment, we also performed the geometry optimization and NBO analysis of the single amino acid model, 4(*R*)-HyP-*exo* conformer in ethanol and ethanol:water (1:1) mixture, along with water solvent (Figure S6). However, no appreciable change in geometry was observed for the three solvents. As shown in Figure S6, the crucial C=O···C=O distance (d) associated with the Bürgi-Dunitz trajectory remains almost similar in all three geometries obtained in water, ethanol, and ethanol:water (1:1) mixture. As a consequence of the unaltered geometries, the calculated $E_{n\to\pi^*}$ values were also found to be similar (~1.0 kcal/mol) in all three cases. Therefore, we have performed all our calculations in the physiologically relevant water solvent along with the gas-phase calculations to understand the effect of the solvent. Our electronic structure analyses reveal that the solvent-phase optimized geometry of the single amino acid model is still not ideal for investigating the role of the stereoelectronic effect. Specifically, a higher $E_{n\to\pi^*}$ value was obtained for the non-natural conformer (*vide supra*), as compared to the natural conformer (Figure 3A). This prompted us to reassess the amino acid model that can correctly rationalize the interplay between different stereoelectronic effects of prolyl-4-hydroxylation on collagen stability.

The physiologically relevant collagenous peptide (PO^4G) to probe different stereoelectronic effects

In the pursuit to establish a physiologically relevant collagenous peptide model that can capture n→π* charge-transfer interaction more accurately as well as feature main-chain torsional angles (ϕ and ψ) and crucial geometrical parameters (d and θ) associated with Bürgi-Dunitz trajectory close to the crystal structure, we aimed at developing a "triplet" amino acid model. The model was constructed to represent the most abundant tripeptide motifs of a single strand of the collagen (Figure 4A). In humans, type I collagen has a higher count of PPG; 46 out of 127 Xaa–Pro–Gly triplets occur in one chain, as compared to any other triplets.[33] Initially, we extracted a tripeptide motif, PO^4G, featuring the 4(*R*)-HyP-*exo* at the Yaa position, from the crystal structure of collagen (PDB: 2G66), where the terminal amino acids were

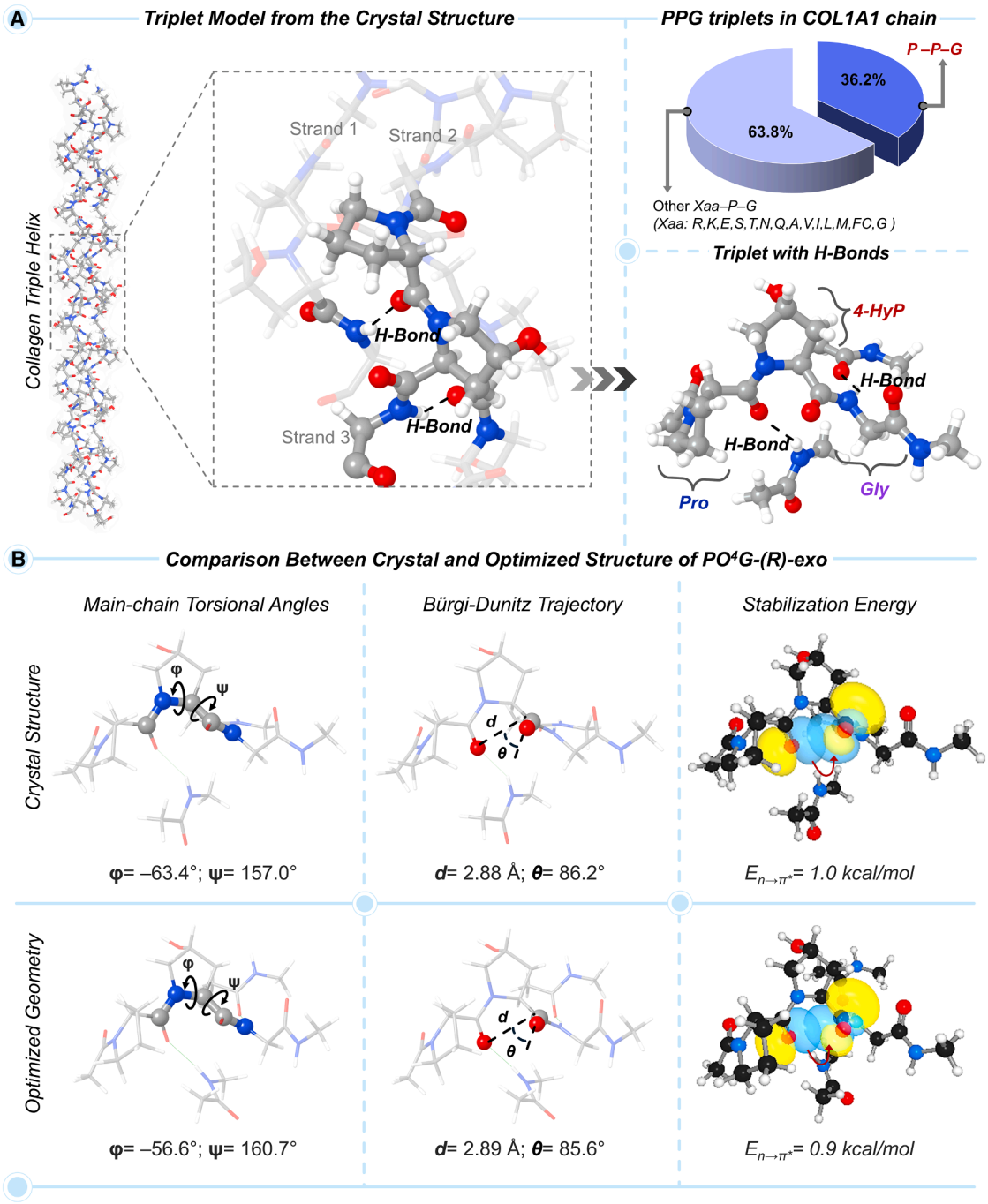


Figure 4. Occurence of PPG triplet in collagen chain and building the physiologically-relevant PO^{4}G peptide model

(A) Collagen triple-helix structure highlighting a tripeptide unit in a single strand with H-bonding interactions with the other strands. The occupancy of the PPG triplet in the COL1A1 chain is shown.

(B) Evaluation of the main-chain torsional angle and Bürgi-Dunitz trajectory of the DFT-M062X-optimized geometry and the crystal structure. The corresponding $E_{n\rightarrow\pi^*}$ value and natural bond orbitals are also shown. See also Figures S7 and S8, and Tables S5 and S6.

capped with methyl groups. The geometry optimization of the PO^{4}G triplet model yielded ϕ torsional angle and the geometrical parameters for n→π* interaction (*d*, *θ*) associated with Bürgi-Dunitz trajectory with significant deviation from the crystal structure (Figure S7A). Consequently, the $E_{n\rightarrow\pi^*}$ values were largely overestimated. Specifically, we noticed that the geometrical distortion in Pro at Xaa and Gly flanked to 4-HyP at Yaa was the cause of this deviation. Therefore, to rectify and make the model more

realistic, we included two additional amino acids at each terminal, 4-HyP and Gly at the N-terminal, and Pro and 4-HyP at the C-terminal to flank. However, this relatively large amino acid model also turned out to be poor at the full geometry optimization level (Figure S7B).

In our subsequent attempt to construct a physiologically relevant tripeptide model, we sought to incorporate the inter-helix H-bonding interaction in the PO^{4}G triplet. The –C=O···H–N– hydrogen bonds in collagen triple-helix involve the –C=O moiety of Pro and the –N–H moiety of Gly holding the three strands together, where each triplet-containing unit possesses two of such hydrogen bonds.[3,5,34] Using this crucial inter-strand structural feature as a guide, we extracted a triplet from one strand hydrogen-bonded with two peptide bonds from the other two strands (Figure 4A). This triplet PO^{4}G featuring the natural conformer of hydroxyproline, 4(*R*)-HyP-*exo* is labeled as PO^{4}G-(*R*)-*exo*. In this tripeptide model, the ϕ and ψ dihedral angles of 4-HyP at the Yaa position are calculated to be −56.6° and 160.7°, respectively, which are close to that of the crystal structure (Figure 4B). In addition, the geometrical parameters, d = 2.89 Å and θ = 85.6° in 4(*R*)-HyP-*exo* are in close agreement with the crystal structure values, 2.88 Å and 86.2°, respectively (Figure 4B). Such a close geometrical match in the optimized triplet model also yielded an $E_{n\rightarrow\pi^*}$ value of 0.9 kcal/mol, almost similar to that calculated with the crystal structure (1.0 kcal/mol, Figure 4B). Therefore, the established physiologically relevant collagenous peptide PO^{4}G can be used further to investigate the correlation between the stereoelectronic parameters due to the prolyl-4-hydroxylation.

As discussed above, the occupancy of 4(*R*)-HyP with an *exo* pyrrolidine ring pucker (O^4) in the Yaa position of the –Xaa–Yaa–G- triplet (PO^{4}G) directly governs the main-chain torsional angles (ϕ and ψ, Figure 4) to attain their optimal values, ϕ = −60 ± 7° and ψ = 150 ± 9° in the collagen PPII-helix. This conformational uniqueness enforces the –C=O groups of the amide bonds, preceding and succeeding to the pyrrolidine ring of the 4(*R*)-HyP, to orient at a C=O···C=O distance of 2.89 Å and an ∠OCO angle of 85.6° (Figure 4B). Notably, this angle closely aligns with the Bürgi-Dunitz trajectory, the preferred approach angle for a nucleophile attacking a carbonyl group. Table S5 clearly demonstrates that the C=O···C=O donor-acceptor distance is directly correlated with the strength of the n→π* charge-transfer interaction in collagenous tripeptides; as the distance increases, $E_{n\rightarrow\pi^*}$ decreases. The $E_{n\rightarrow\pi^*}$ value is maximum in the PO^{4}G collagenous tripeptide featuring the (*R*)-*exo* conformer of the 4-HyP.

To systematically compare other amino acids with 4-HyP at the Yaa position, we began by analyzing the sequence of the helical region in the human collagen-I alpha-1 (COL1A1) chain. The sequence analysis suggests that 35.7% of total tripeptides are PPG (Figure S8A). Melting temperature (T_m°C) experiments have well-established that the presence of 4-HyP at the Yaa position enhances the thermal stability of the collagen triple helix (Table S6).[35–39] Notably, replacing 4-HyP with any amino acid—except arginine (Arg)—leads to a significant decrease in the melting temperature. Given the predominant presence of proline in the collagen chain, we primarily focused our investigation on unraveling the stereoelectronic effects, specifically within the PO^{4}G tripeptide. To achieve the systematic comparison involving other amino acids in the Yaa position, we used the relative occupancies of the PYG motifs presented in Figure S8A as a guide. Specifically, we estimated the stabilization effect due to the crucial n→π* charge-transfer interaction ($E_{n\rightarrow\pi^*}$, kcal/mol) for alanine PAG (24.3%), arginine PRG (7.8%), lysine PKG (6.1%), serine PSG (6.1%), and aspartic acid PDG (0.9%) in the Yaa position of the PYG triplet (Figure S8B). The PAG, PDG, and PKG tripeptides exhibit lower stabilization energy ($E_{n\rightarrow\pi^*}$) values compared to PO^{4}G. While PRG and PSG display stabilization energies similar to PO^{4}G, the predominant presence of proline at the Yaa position and its hydroxylation likely exerts a significantly greater stabilizing effect on the collagen triple helix. In line with this, the T_m of PO^{4}G is the highest (Table S6), therefore establishing the crucial role of 4-HyP in collagen thermal stability.[39] In their work, Brodsky et al. mentioned that positively charged arginine could engage in side-chain interactions with the available backbone carbonyl groups, contributing to higher thermal stability than hydroxy-proline.[39] As the T_m of PRG containing triple helix is close to that of PO^{4}G, its role remains an open question. Thus, a systematic comparison of various amino acids at the Yaa position in the PYG triplet, alongside 4-HyP in the PO^{4}G motif, offers deeper insights into the role of the prevalent PPG sequence in collagen helical stability.

Positional preference of 4-HyP in pre-Organizing the main-chain torsional angles, Bürgi-Dunitz trajectory, and n→π* interaction that stabilizes collagen triple-helix is revealed using PO^{4}G triplets containing four different conformers of 4-HyP

To obtain insights into the natural preference of 4(*R*)-*exo* conformer, we first compared the geometrical parameters (d and θ), main-chain torsional angles (ϕ and ψ), and electronic parameters, $E_{n\rightarrow\pi^*}$ for four different tripeptides, PO^{4}G-(*R*)-*exo*, PO^{4}G-(*R*)-*endo*, PO^{4}G-(*S*)-*exo,* and PO^{4}G-(*S*)-*endo* featuring four different conformers of 4-HyP (Figure 5). To assess the geometrical parameters in the DFT-optimized structures, the crystal structure was used as the reference. As shown in Figures 5A and 5B, the ϕ and ψ angles and d and θ in PO^{4}G-(*R*)-*exo* are relatively closer to the crystal structure, as compared to the PO^{4}G-(*R*)-*endo*, PO^{4}G-(*S*)-*exo,* and PO^{4}G-(*S*)-*endo* containing unnatural conformers of 4-HyP (Figures 5A and 5B). Due to the optimal geometrical parameters in the PO^{4}G-(*R*)-*exo* tripeptide, the $E_{n\rightarrow\pi^*}$ value was estimated to be 0.9 kcal/mol (Figure 5C). On the other hand, in the other three triplets containing the non-natural conformers of 4-HyP, the calculated $E_{n\rightarrow\pi^*}$ value decreases as, PO^{4}G-(*R*)-*endo* (0.6 kcal/mol) > PO^{4}G-(*S*)-*exo* (0.2 kcal/mol) and PO^{4}G-(*S*)-*endo* (0.1 kcal/mol). Importantly, the $E_{n\rightarrow\pi^*}$ value appears to be nicely correlated with the main-chain torsional angles (ϕ and ψ) and geometrical parameters d and θ. As represented in Figure 5B, the highest $E_{n\rightarrow\pi^*}$ value (represented with the largest circle) corresponds to the crystal structure, and the next highest is associated with the DFT-optimized PO^{4}G-(*R*)-*exo* triplet. As $E_{n\rightarrow\pi^*}$ is a quantification of the stabilization energy of a conformer, the above results nicely portray the positional preference of 4-HyP at the Yaa with the *exo* ring pucker, which not only pre-organizes

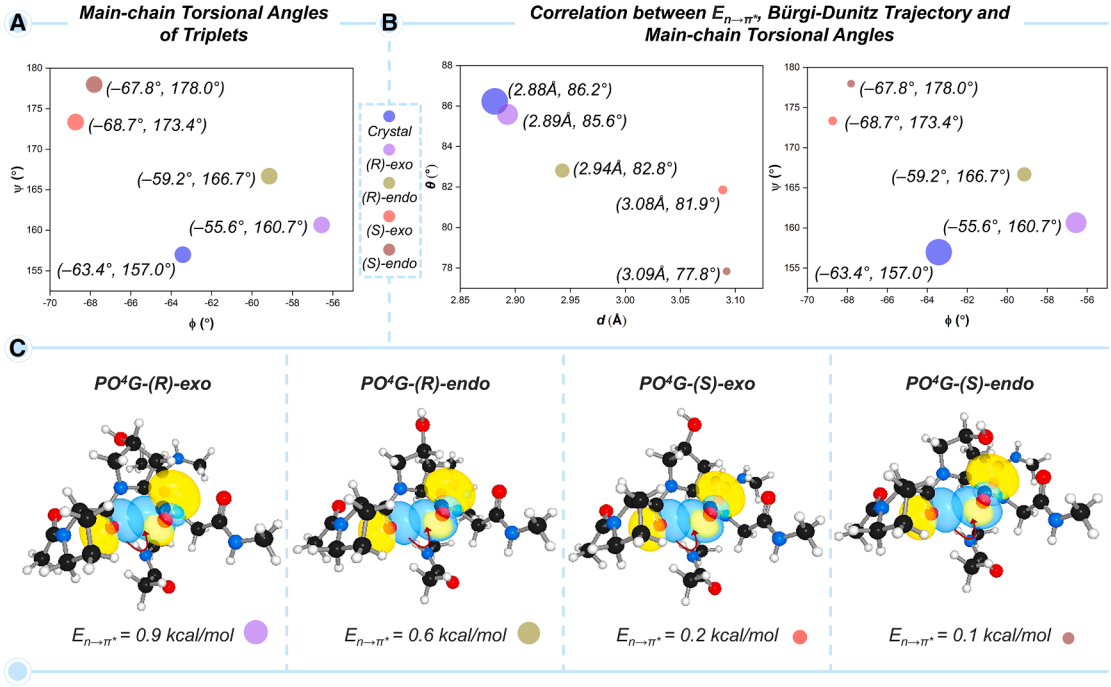


Figure 5. Correlation between geometrical parameters and n→π* charge-transfer interaction in PO^{4}G triplets

(A) Main-chain torsional angles (ϕ and ψ) of triplets along with the crystal structure.

(B) Correlations between $E_{n\to\pi^*}$ and Bürgi-Dunitz trajectory and $E_{n\to\pi^*}$ and main-chain torsional angles.

(C) Natural bond orbitals and $E_{n\to\pi^*}$ values for PO^{4}G triplets bearing the four different conformers of 4-HyP.

the main-chain torsional angles but also affects the n→π* interaction, stabilizing the overall structure of collagen.

In addition to the correlation between geometrical and electronic parameters of the PO^{4}G tripeptide, it would be interesting to see how the n→π* interaction gets affected by the hydroxylation of Pro at the Yaa position. To investigate this, we extracted a tripeptide, PPG-*exo* from the PDB file (ID: 1K6F). Our calculations reveal that the Bürgi-Dunitz trajectory (θ) associated with Pro at Yaa in the optimized geometry of the triplet PPG-*exo* decreased by 2.3°. Consequently, a lower $E_{n\to\pi^*}$ value of the triplet PPG-*exo* was obtained; 0.6 kcal/mol vs. 0.9 kcal/mol in the PO^{4}G-(*R*)-*exo* triplet (Figure 6). This result verifies that the 4-hydroxylation PTM indeed affects the structural parameters of a collagen single strand. This dictates the strength of the n→π* interaction, and thereby increases the collagenous triple-helical stability.

Moreover, the preference of proline for *exo* ring pucker increases as the σ→σ* interaction between C^{β}–H_{ax}/C^{δ}–H_{ax} and C^{γ}–X strengthens. The interaction becomes further stronger as the electronegativity of the X substituent increases.[9] Therefore, we anticipated a stronger σ→σ* interaction may exist in the abundant PO^{4}G-(*R*)-*exo* tripeptide, as compared to the unmodified PPG-*exo* tripeptide. Indeed, the calculated $E_{\sigma\to\sigma^*}$ value of ~9 kcal/mol for PO^{4}G-(*R*)-*exo* tripeptide is twice higher compared to that of the PPG-*exo* analog (Figure 6). This explains the crucial role of σ→σ* charge-transfer interaction in stabilizing the collagenous helicity.

To shed light on the positional preference of 4-HyP at the Yaa position, we examined the analogous O^{4}PG sequence, where 4-HyP is positioned at the Xaa site within the triplet. In this case, we aim to estimate the stabilization effect due to the crucial n→π* charge-transfer interaction through the $E_{n\to\pi^*}$ (in kcal/mol) energy calculation for the O^{4}PG triplet. However, the tripeptide model needed to extend to a tetrapeptide model to account for both the amide bonds connected to the 4-Hyp in the Xaa position. This resulted in a GO^{4}PG tetrapeptide sequence with the 4(*R*)-HyP-*exo* (O^4) conformer. Now, we compared the calculated $E_{n\to\pi^*}$ values for GO^{4}PG-(*R*)-*exo* and GPO^{4}G-(*R*)-exo, featuring the 4(*R*)-HyP-*exo* in the Xaa and Yaa positions, respectively. The same computational protocol, as described in the computational detail section, was followed to estimate the $E_{n\to\pi^*}$ values. The computed $E_{n\to\pi^*}$ values, along with the NBOs overlap between the donor and acceptor orbitals are presented in Figure S9.

As shown in Figure S9, the presence of 4(*R*)-HyP in the Xaa position, in GO^{4}PG-(*R*)-*exo*, increases the $E_{n\to\pi^*}$ stabilization energy by 0.4 kcal/mol around the Xaa position and by 1.9 kcal/mol around the Yaa position, compared to the GPO^{4}G-(*R*)-*exo* natural analog. This result is counterintuitive, given the natural preference of 4-HyP at the Yaa position, which is known to enhance collagen helical stability. However, the analysis of the main-chain torsional angles (ϕ and ψ) and Bürgi-Dunitz trajectory (d and θ) around Yaa reveals a significant deviation in the GO^{4}PG-(R)-*exo* sequence from the optimal values. Specifically, a $\phi = -45.1°$ for Pro in GO^{4}PG-(*R*)-*exo* and $\phi = -53.4°$ for 4-HyP in GPO^{4}G-(*R*)-*exo* was observed, where the former significantly deviates from the optimum range of $\phi = -60 \pm 7°$.[31]

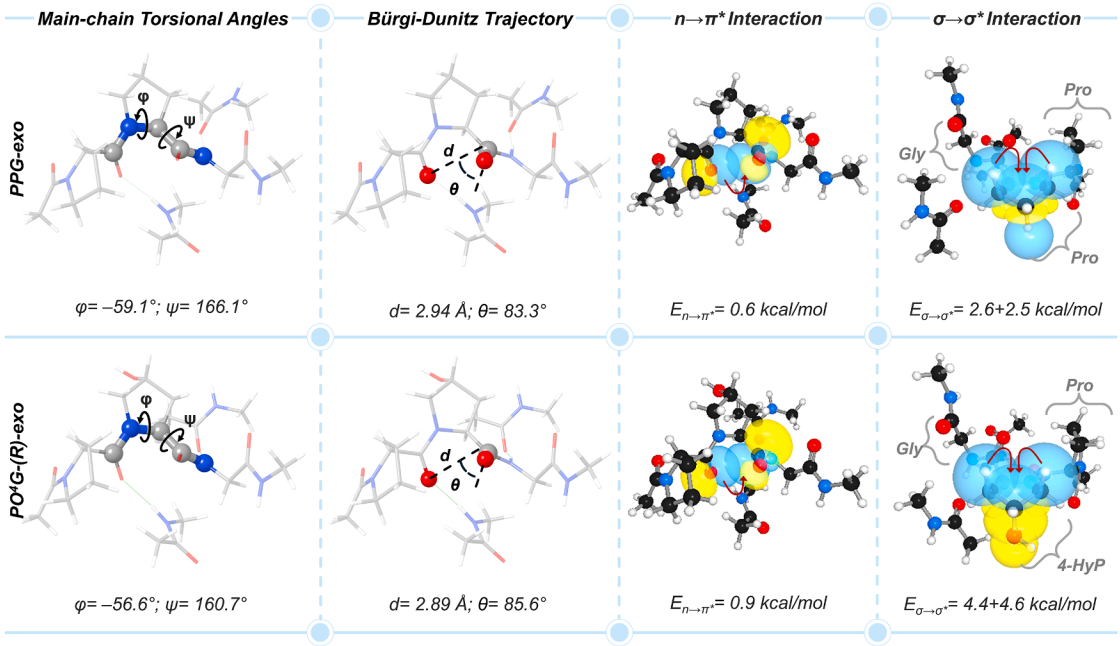


Figure 6. Comparative analysis between PPG and PO^{4}G triplets

Comparison of geometrical parameters and charge-transfer (n→π* and σ→σ*) interactions between the PPG-*exo* and PO^{4}G-(*R*)-*exo* tripeptides showcasing the effect of 4-hydroxylation at the Yaa position. See also Figure S9.

Furthermore, the Bürgi-Dunitz trajectory in GO^{4}PG-(R)-*exo*, d = 2.72 Å and θ = 93.1°, also appeared to be highly deviated compared to d = 2.78 Å and θ = 97.7° in GPO^{4}G-(*R*)-*exo* (Figure S9). Evidently, the Bürgi-Dunitz trajectory in GO^{4}PG-(*R*)-*exo* is far off from the optimal value of $\theta \sim 109°$.[17] This unnatural deviation in geometry in GO^{4}PG-*(R)*-*exo* overestimates the calculated $E_{n\rightarrow\pi^*}$ stabilization energy. Notably, the overall thermal stability (ΔH) GPO^{4}G-(*R*)-*exo* motif was calculated to be 0.3 kcal/mol higher compared to the GO^{4}PG-(*R*)-*exo* analog. In addition to the unnatural geometrical distortion due to the presence of 4-HyP in the Xaa position, it has been shown in the literature that the collagen model peptide (O^{4}PG)$_{10}$ with 4-HyP in the Xaa position does not form the triple helix.[40] This experimental observation also supports the positional preference of the 4-HyP in the Yaa position. Therefore, though the presence of 4-HyP shows an unnatural increment in $E_{n\rightarrow\pi^*}$ at the peptide backbone, the analysis of main-chain torsional angles and the thermochemistry of motifs suggest the opposite, that the presence of 4-HyP at Xaa may hinder the helicity of PPII-like structure of single strand, the thermal stability of triple helix, and the packing of collagen into fibrils. Consequently, 4(*R*)-HyP with *exo* ring pucker is likely positionally preferred at Yaa over Xaa in the collagen chain.

Correlation between the *trans*-peptide bond and charge-transfer (n→π* and σ→σ*) interactions

It is known that if a peptide bond attains a *cis* conformation in a single strand of collagen, the collagen triple helix does not form and transport out to the extracellular matrix.[41] An isomerase enzyme, called peptidyl-prolyl isomerase (PPIase) is responsible for the conformational change of the *cis* peptide bond to *trans* that leads to the formation of the collagen triple helix.[42,43] The *trans* peptide bond connected to the 4-HyP residue in the PO^{4}G triplet is likely to contribute to maximizing the n→π* interaction to provide greater stability to the collagen chains and subsequently to the triple helix. To investigate this, relaxed torsional angle (ω) scans of the preceding (ω_{prec}) and succeeding (ω_{succ}) peptide bonds were performed (Figure 7A; Table S7). Specifically, through the systematic variation of the torsional angle affecting the –C=O···O=C– distance (*d*), we monitored the $E_{n\rightarrow\pi^*}$ values (Figure 7C). This torsional angle vs. $E_{n\rightarrow\pi^*}$ analysis is set to provide insights into the interconnection between ring pucker, stereoelectronic effect, and peptide bond conformation.

The *trans* conformer possessing $\omega_{prec} \sim 180°$ and $\omega_{succ} \sim 180°$ (180°,180°) is calculated to be more stable than the *cis* isomer with $\omega_{prec} \sim 0°$ and $\omega_{succ} \sim 0°$ (0°,0°) by 4.5 kcal/mol (Figure 7B). This observation falls in line with the established notion. We anticipated that the ring pucker would bear the energy penalty of the *cis* ↔ *trans* transformation of the peptide bond. To probe the interplay between the peptide bond isomerization and the ring pucker, five different points were selected from the conformational landscape (Figures 7A and 7B). Interestingly, we observed that as one bond moves from *trans*/*trans* (180°,180°) to the highest energy point (60°,60°), the pyrrolidine ring becomes almost planar. Moving forward to the other end of the conformational landscape, as both the peptide bonds attain a *cis*/*cis* conformation (0°,0°), the ring puckering gets shifted from *exo* to *endo* leading to the 4(*R*)-HyP-*endo* conformer (Figure 7B). This result nicely showcases the interdependence of the ring puckering and the peptide bond conformation.

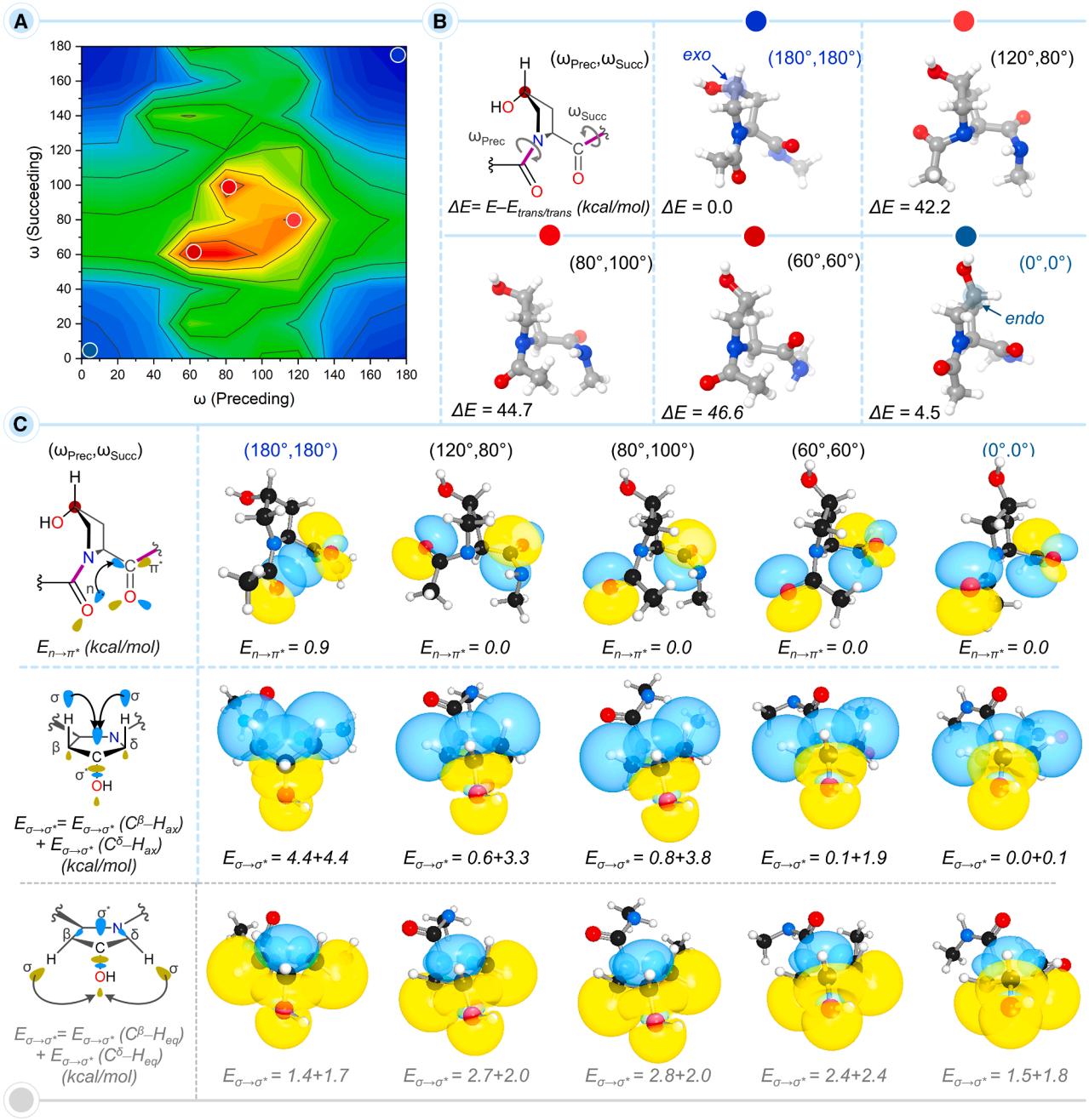


Figure 7. Relative energy and NBO analyses of conformers from two-dimensional energy landscape of 4(*R*)-HyP-*exo*

(A) A two-dimensional energy landscape of the relaxed torsional angle scans of the preceding (ω_{prec}) and succeeding (ω_{succ}) peptide bonds of the 4(*R*)-HyP-*exo* conformer at the M062X/6-31+G(d,p) level of theory.

(B) The highest, lowest, and a few intermediate energy conformers are presented along with their relative energies (Δ*E*, kcal/mol).

(C) Calculated $E_{n\rightarrow\pi^*}$ and $E_{\sigma\rightarrow\sigma^*}$ values (in kcal/mol) associated with the n→π* and σ→σ* interactions, respectively for five selected conformers. See also Tables S7 and S8.

Apart from the n→π*, another type of charge transfer interaction in amino acids can occur between an electron-rich σ-bond donor (e.g., C–H σ orbital) and a vacant electron-deficient σ* acceptor (e.g., C–OH σ* orbital) orbital. The nature of the σ→σ* interaction as the charge transfer interaction is evident from the change in occupation numbers of the associated orbitals (Table S8). With establishing the elegant correlation between the ring puckering and peptide bond conformation, it would be further intriguing to investigate how this correlation is connected to the electronic stabilization effects associated with the n→π*

and σ→σ* charge-transfer interactions. The NBO analysis evidence that the isomerization of both peptide bonds from *trans* to *cis* leads to destabilization due to the loss in the n→π* interaction. The latter is evident from the $E_{n\rightarrow\pi^*}$ value of the *trans* conformer, 0.9 kcal/mol, which completely diminishes to 0.0 for the *cis* peptide bond (Figure 7C). As previously discussed in the introduction section, the *exo* ring pucker is stabilized by the delocalization σ-bonding electrons of C^{β}–H_{ax}/C^{δ}–H_{ax} bonds into σ* orbital of C^{γ}–OH. Therefore, we analyzed the σ→σ* interactions to obtain further insights into the electronic stabilization effect (Figure 7C). In the *trans*/*trans* (180°,180°) conformer, the strength of σ→σ* interactions due to C^{β}–H_{ax} and C^{δ}–H_{ax} was calculated to be equal, 8.8 kcal/mol. During the transition of the peptide bonds from *trans*/*trans* (180°,180°) to *cis*/*cis* (0°,0°), the strength of the σ→σ* interactions gradually reduces. Interestingly, the σ→σ* interaction due to the C^{β}–H_{ax} reduced more drastically than the other, C^{δ}–H_{ax} (Figure 7C). The diminished σ→σ* charge transfer interaction due to the axial $C^{\beta/\delta}$–H_{ax} bonds of the pyrrolidine ring alters the ring torsional angle χ_1 (Figure 1B), which, in turn, dictates the ring pucker, *endo*, or *exo*. As such, the peptide bond isomerization is also interconnected with the nature of the ring pucker. This can be further evidenced through the preferred *endo* ring pucker of the *cis*/*cis* (0°,0°) conformer. As the peptide bonds attain the *cis*/*cis* (0°,0°) conformer, the σ orbitals of the equatorial C^{β}–H_{eq} and C^{δ}–H_{eq} bonds of the pyrrolidine ring and σ* orbital of the C^{γ}–OH experience an enhanced the σ→σ* charge-transfer interactions (Figure 7C). This stabilization effect eventually favors an *endo* ring pucker instead of the *exo*, which explains why the natural *trans*/*trans* peptide bond conformer favors an *exo* ring pucker for the pyrrolidine ring.

In addition to the electronic stabilization effect due to the peptide bond conformation, to qualitatively assess how the change in peptide bond conformation affects the pyrrolidine ring puckering, we analyzed the distortion experienced in the ring at four corner points on the two-dimensional energy landscape, i.e., (180°,180°), (0°,180°), (180°,0°), and (0°,0°) (Figure S10B). For this purpose, the conformers were fragmented into pyrrolidine ring and peptide bonds, and the fragments were capped with H atoms. Then, the relative energies of the rings with respect to that of the (180°,180°) ring were compared. The ring conformer at (0°,180°) experiences the highest distortion due to the peptide bond isomerization, as reflected in the relative energy value of 4.9 kcal/mol. A lower degree of destabilization, by 0.1 and 1.3 kcal/mol was observed for the ring conformers at (180°,0°), and (0°,0°), respectively. Although the highest degree of destabilization to the ring pucker is expected to be observed at (0°,0°), the change in ring pucker from *exo* to *endo* at this point likely compensates for the destabilization energy. Thus, the correlation between *cis* ↔ *trans* peptide bond conformation, pyrrolidine ring puckering, and electronic stabilization due to n→π* and σ→σ* charge-transfer interactions established in this section nicely rationalizes the natural preference of the *trans* peptide bond conformation.

Conclusions

Using a calibrated DFT method and a physiologically relevant collagenous PO^4G tripeptide, we have investigated the elegant interplay between different stereoelectronic effects, pyrrolidine ring puckers, torsional angles (ϕ, ψ, and ω), peptide bond *cis*/*trans* conformation, and n→π*/σ→σ* charge-transfer interaction, all governed by prolyl-4-hydroxylation in achieving collagen triple-helical stability. Our work employs the PO^4G tripeptide, incorporating four different conformers of 4-HyP–PO^4G-(*R*)-*exo*, PO^4G-(*R*)-*endo*, PO^4G-(*S*)-*exo*, and PO^4G-(*S*)-*endo*–alongside the non-hydroxylated PPG tripeptide to unveil the positional preference of 4-HyP with (*R*)-*exo* ring pucker. Specifically, the geometrical parameters and NBO analyses of the collagenous tripeptide at the DFT-M062X level show 4-HyP in the PO^4G-(*R*)-*exo* tripeptide enforces the *trans* peptide bond conformation through the optimal values of the main-chain torsional angles. This, in turn, tunes the Bürgi-Dunitz trajectory to achieve the highest $E_{n\rightarrow\pi^*}$ of 0.9 kcal/mol in the PO^4G-(*R*)-*exo* triplet. The NBO analyses also uncover that the (*R*)-*exo* ring pucker, due to 4-hydroxylation, gains additional stability due to the σ→σ* charge-transfer interaction between the axial $C^{\beta/\delta}$–H σ bonds and C^{γ}–OH σ* orbital in the pyrrolidine ring, contributing to the helical stability. Overall, this work offers a fundamental electronic-level understanding of collagenous helical stability driven by the intricate interplay of inter-connected stereoelectronic effects arising from prolyl-4-hydroxylation. Additionally, this study also paves the way for an accurate quantum chemical understanding of the role of other PTMs that contribute to collagen stability.

Limitations of the study

This study has not delved into the effect of other amino acids (positively and negatively charged and neutral) present at the Xaa position (non-Pro-4HyP-Gly motif) on the helical stabilization attained by prolyl-4-hydroxylation catalyzed by different prolyl-4-hydroxylases.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Bhaskar Mondal (bhaskarmondal@iitmandi.ac.in).

Materials availability

The study did not generate new unique reagents.

Data and code availability

- Data: Optimized Cartesian coordinates for all single amino acid and large amino acid models used in the study are available in the Mendeley data repository (Joshi, Ashutosh; Basak, Trayambak; Mondal, Bhaskar (2025), "Decoding Elegant Interplay Between Different Stereo-Electronic Effects Due to the Ancient Prolyl-4-Hydroxylation Stabilizing Collagenous Helicity", Mendeley Data, V1, https://doi.org/10.17632/ychk3jfdfk.1).
- Code: No code was generated in this study.
- Other: See the STAR Methods section for methodological details. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

A.J. thanks the Ministry of Education (MoE), Government of India, for the research fellowship. T.B. and B.M. are further grateful for the funding support

from the Department of Biotechnology, Govt. of India, for "Understanding the prolyl-3-hydroxylase 1 (P3H1) mediated perturbed post-translational modification of collagen network aggravating fibrotic remodeling during myocardial infarction", DBT – BT/PR/50004/CMD/150/148/2023. B.M. and T.B. are also grateful for the funding support from the Anusandhan National Research Foundation-Science and Engineering Research Board (ANRF-SERB) in the form of Core Research Grants, CRG/2023/002138 and CRG/2022/006204, respectively. The authors acknowledge the National Supercomputing Mission (NSM), for providing computing resources of 'PARAM Himalaya' at IIT Mandi, which is implemented by C-DAC and supported by the Ministry of Electronics and Information Technology (MeitY) and Department of Science and Technology (DST), Government of India. In addition, the High-Performance Computing (HPC) facility at IIT Mandi is also acknowledged for providing high-end computational resources. Further, we dedicate this work to Prof. Ronald T. Raines (MIT Chemistry), whose pioneering contributions to the field of prolyl-4-hydroxylation have been a profound source of inspiration for our research.

AUTHOR CONTRIBUTIONS

A.J., T.B., and B.M. designed the project. A.J. executed all the theoretical calculations. The manuscript was written with contributions from all the authors.

DECLARATION OF INTERESTS

There are no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- METHOD DETAILS
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2025.112393.

Received: December 13, 2024
Revised: February 26, 2025
Accepted: April 7, 2025
Published: April 10, 2025

REFERENCES

1. Pokidysheva, E., Boudko, S., Vranka, J., Zientek, K., Maddox, K., Moser, M., Fässler, R., Ware, J., and Bächinger, H.P. (2014). Biological role of prolyl 3-hydroxylation in type IV collagen. Proc. Natl. Acad. Sci. USA *111*, 161–166. https://doi.org/10.1073/pnas.1307597111.
2. Yang, J., Kojasoy, V., Porter, G.J., and Raines, R.T. (2024). Pauli Exclusion by n→π* Interactions: Implications for Paleobiology. ACS Cent. Sci. *10*, 1829–1834. https://doi.org/10.1021/acscentsci.4c00971.
3. Bella, J. (2016). Collagen structure: New tricks from a very old dog. Biochem. J. *473*, 1001–1025. https://doi.org/10.1042/BJ20151169.
4. Onursal, C., Dick, E., Angelidis, I., Schiller, H.B., and Staab-Weijnitz, C.A. (2021). Collagen Biosynthesis, Processing, and Maturation in Lung Ageing. Front. Med. *8*, 593874. https://doi.org/10.3389/fmed.2021.593874.
5. Shoulders, M.D., and Raines, R.T. (2009). Collagen structure and stability. Annu. Rev. Biochem. *78*, 929–958. https://doi.org/10.1146/annurev.biochem.77.032207.120833.
6. Ramshaw, J.A., Shah, N.K., and Brodsky, B. (1998). Gly-X-Y tripeptide frequencies in collagen: A context for host-guest triple-helical peptides. J. Struct. Biol. *122*, 86–91. https://doi.org/10.1006/jsbi.1998.3977.
7. Vitagliano, L., Berisio, R., Mazzarella, L., and Zagari, A. (2001). Structural Bases of Collagen Stabilization Induced by Proline Hydroxylation. Biopolymers *58*, 459–464. https://doi.org/10.1002/1097-0282(20010415)58:5<459::AID-BIP1021>3.0.CO;2-V.
8. DeRider, M.L., Wilkens, S.J., Waddell, M.J., Bretscher, L.E., Weinhold, F., Raines, R.T., and Markley, J.L. (2002). Collagen stability: Insights from NMR spectroscopic and hybrid density functional computational investigations of the effect of electronegative substituents on prolyl ring conformations. J. Am. Chem. Soc. *124*, 2497–2505. https://doi.org/10.1021/ja0166904.
9. Improta, R., Benzi, C., and Barone, V. (2001). Understanding the role of stereoelectronic effects in determining collagen stability. 1. A quantum mechanical study of proline, hydroxyproline, and fluoroproline dipeptide analogues in aqueous solution. J. Am. Chem. Soc. *123*, 12568–12577. https://doi.org/10.1021/ja010599i.
10. Holmgren, S.K., Taylor, K.M., Bretscher, L.E., and Raines, R.T. (1998). Code for collagen's stability deciphered. Nature *392*, 666–667. https://doi.org/10.1038/33573.
11. Okuyama, K., Hongo, C., Wu, G., Mizuno, K., Noguchi, K., Ebisuzaki, S., Tanaka, Y., Nishino, N., and Bächinger, H.P. (2009). High-resolution structures of collagen-like peptides [(Pro-Pro-Gly)4-Xaa-Yaa-Gly-(Pro-Pro-Gly)4]: Implications for triple-helix hydration and Hyp(X) puckering. Biopolymers *91*, 361–372. https://doi.org/10.1002/bip.21138.
12. Vitagliano, L., Berisio, R., Mastrangelo, A., Mazzarella, L., and Zagari, A. (2001). Preferred proline puckerings in cis and trans peptide groups: Implications for collagen stability. Protein Sci. *10*, 2627–2632. https://doi.org/10.1110/ps.ps.26601a.
13. Berisio, R., Vitagliano, L., Mazzarella, L., and Zagari, A. (2002). Crystal structure of the collagen triple helix model [(Pro-Pro-Gly) 10] 3. Protein Sci. *11*, 262–270. https://doi.org/10.1110/ps.32602.
14. Ramachandran, G.N., and Kartha, G. (1955). Structure of collagen. Nature *176*, 593–595. https://doi.org/10.1038/176593a0.
15. Bürgi, H.B., Dunitz, J.D., and Shefter, E. (1973). Geometrical Reaction Coordinates. II. Nucleophilic Addition to a Carbonyl Group. J. Am. Chem. Soc. *95*, 5065–5067. https://doi.org/10.1021/ja00796a058.
16. Bürgi, H.B., Dunitz, J.D., Lehn, J.M., and Wipff, G. (1974). Stereochemistry of reaction paths at carbonyl centres. Tetrahedron *30*, 1563–1572. https://doi.org/10.1016/S0040-4020(01)90678-7.
17. Newberry, R.W., and Raines, R.T. (2017). The n→π* Interaction. Acc. Chem. Res. *50*, 1838–1846. https://doi.org/10.1021/acs.accounts.7b00121.
18. Choudhary, A., Gandla, D., Krow, G.R., and Raines, R.T. (2009). Nature of amide carbonyl-carbonyl interactions in proteins. J. Am. Chem. Soc. *131*, 7244–7246. https://doi.org/10.1021/ja901188y.
19. Bretscher, L.E., Jenkins, C.L., Taylor, K.M., Derider, M.L., and Raines, R.T. (2001). Conformational Stability of Collagen Relies on a Stereoelectronic Effect. J. Am. Chem. Soc. *123*, 777–778. https://doi.org/10.1021/ja005542v.
20. Jenkins, C.L., Lin, G., Duo, J., Rapolu, D., Guzei, I.A., Raines, R.T., and Krow, G.R. (2004). Substituted 2-azabicyclo[2.1.1]hexanes as constrained proline analogues: Implications for collagen stability. J. Org. Chem. *69*, 8565–8573. https://doi.org/10.1021/jo049242y.
21. Hinderaker, M.P., and Raines, R.T. (2003). An electronic effect on protein structure. Protein Sci. *12*, 1188–1194. https://doi.org/10.1110/ps.0241903.
22. Hodges, J.A., and Raines, R.T. (2006). Energetics of an n → π* interaction that impacts protein structure. Org. Lett. *8*, 4695–4697. https://doi.org/10.1021/ol061569t.
23. Shoulders, M.D., Hodges, J.A., and Raines, R.T. (2006). Reciprocity of steric and stereoelectronic effects in the collagen triple helix. J. Am. Chem. Soc. *128*, 8112–8113. https://doi.org/10.1021/ja061793d.
24. Bogojeski, M., Vogt-Maranto, L., Tuckerman, M.E., Müller, K.R., and Burke, K. (2020). Quantum chemical accuracy from density functional

approximations via machine learning. Nat. Commun. *11*, 5223. https://doi.org/10.1038/s41467-020-19093-1.

25. Riplinger, C., and Neese, F. (2013). An efficient and near linear scaling pair natural orbital based local coupled cluster method. J. Chem. Phys. *138*, 034106. https://doi.org/10.1063/1.4773581.
26. Liakos, D.G., Sparta, M., Kesharwani, M.K., Martin, J.M.L., and Neese, F. (2015). Exploring the accuracy limits of local pair natural orbital coupled-cluster theory. J. Chem. Theory Comput. *11*, 1525–1539. https://doi.org/10.1021/ct501129s.
27. Bartlett, G.J., Choudhary, A., Raines, R.T., and Woolfson, D.N. (2010). $n \rightarrow \pi^*$ interactions in proteins. Nat. Chem. Biol. *6*, 615–620. https://doi.org/10.1038/nchembio.406.
28. Rahim, A., Saha, P., Jha, K.K., Sukumar, N., and Sarma, B.K. (2017). Reciprocal carbonyl-carbonyl interactions in small molecules and proteins. Nat. Commun. *8*, 78. https://doi.org/10.1038/s41467-017-00081-x.
29. Reed, A.E., Curtiss, L.A., and Weinhold, F. (1988). Intermolecular Interactions from a Natural Bond Orbital, Donor—Acceptor Viewpoint. Chem. Rev. *88*, 899–926. https://doi.org/10.1021/cr00088a005.
30. Landis, C.R., and Weinhold, F. (2014). The NBO View of Chemical Bonding. In The Chemical Bond: Fundamental Aspects of Chemical Bonding, G. Frenking and S. Shaik, eds. (Wiley-VCH Verlag GmbH & Co. KGaA), pp. 91–120. https://doi.org/10.1002/9783527664696.ch3.
31. Egli, J., Schnitzer, T., Dietschreit, J.C.B., Ochsenfeld, C., and Wennemers, H. (2020). Why Proline? Influence of Ring-Size on the Collagen Triple Helix. Org. Lett. *22*, 348–351. https://doi.org/10.1021/acs.orglett.9b03528.
32. Rondi, A., Rodriguez, Y., Feurer, T., and Cannizzo, A. (2015). Solvation-driven charge transfer and localization in metal complexes. Acc. Chem. Res. *48*, 1432–1440. https://doi.org/10.1021/ar5003939.
33. Salo, A.M., Rappu, P., Koski, M.K., Karjalainen, E., Izzi, V., Drushinin, K., Miinalainen, I., Käpylä, J., Heino, J., and Myllyharju, J. (2024). Collagen prolyl 4-hydroxylase isoenzymes I and II have sequence specificity towards different X-Pro-Gly triplets. Matrix Biol. *125*, 73–87. https://doi.org/10.1016/j.matbio.2023.12.001.
34. Ramachandran, G.N., and Kartha, G. (1954). Structure of collagen. Nature *174*, 269–270. https://doi.org/10.1038/174269c0.
35. Shah, N.K., Ramshaw, J.A., Kirkpatrick, A., Shah, C., and Brodsky, B. (1996). A host-guest set of triple-helical peptides: Stability of Gly-X-Y triplets containing common nonpolar residues. Biochemistry *35*, 10262–10268. https://doi.org/10.1021/bi960046y.
36. Chan, V.C., Ramshaw, J.A., Kirkpatrick, A., Beck, K., and Brodsky, B. (1997). Positional Preferences of Ionizable Residues in Gly-X-Y Triplets of the Collagen Triple-helix. J. Biol. Chem. *272*, 31441–31446. https://doi.org/10.1074/JBC.272.50.31441.
37. Holmgren, S.K., Bretscher, L.E., Taylor, K.M., and Raines, R.T. (1999). A hyperstable collagen mimic. Chem. Biol. *6*, 63–70. https://doi.org/10.1016/S1074-5521(99)80003-9.
38. Beck, K., Chan, V.C., Shenoy, N., Kirkpatrick, A., Ramshaw, J.A., and Brodsky, B. (2000). Destabilization of osteogenesis imperfecta collagen-like model peptides correlates with the identity of the residue replacing glycine. Proc. Natl. Acad. Sci. USA *97*, 4273–4278. https://doi.org/10.1073/pnas.070050097.
39. Persikov, A.V., Ramshaw, J.A., Kirkpatrick, A., and Brodsky, B. (2000). Amino acid propensities for the collagen triple-helix. Biochemistry *39*, 14960–14967. https://doi.org/10.1021/bi001560d.
40. Inouye, K., Kobayashi, Y., Kyogoku, Y., Kishida, Y., Sakakibara, S., and Prockop, D.J. (1982). Synthesis and physical properties of (hydroxyproline-proline-glycine)10: Hydroxyproline in the X-position decreases the melting temperature of the collagen triple helix. Arch. Biochem. Biophys. *219*, 198–203. https://doi.org/10.1016/0003-9861(82)90149-7.
41. Lu, K.P., Finn, G., Lee, T.H., and Nicholson, L.K. (2007). Prolyl cis-trans isomerization as a molecular timer. Nat. Chem. Biol. *3*, 619–629. https://doi.org/10.1038/nchembio.2007.35.
42. Ishikawa, Y., and Bächinger, H.P. (2013). A molecular ensemble in the rER for procollagen maturation. Biochim. Biophys. Acta - Mol. Cell Res. *1833*, 2479–2491. https://doi.org/10.1016/j.bbamcr.2013.04.008.
43. Ishikawa, Y., Boudko, S., and Bächinger, H.P. (2015). Ziploc-ing the structure: Triple helix formation is coordinated by rough endoplasmic reticulum resident PPIases. Biochim. Biophys. Acta Gen. Subj. *1850*, 1983–1993. https://doi.org/10.1016/j.bbagen.2014.12.024.
44. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Petersson, G.A., Nakatsuji, H., Li, X., et al. (2019). Gaussian 16 (Wallingford, CT: Gaussian, Inc.).
45. Neese, F. (2012). The ORCA program system. WIREs Comput. Mol. Sci. *2*, 73–78. https://doi.org/10.1002/wcms.81.
46. NBO 7.0., Glendening, E.D., Badenhoop, J.K., Reed, A.E., Carpenter, J.E., Bohmann, J.A., Morales, C.M., Karafiloglou, P., Landis, C.R., and Weinhold, F. (2018). Theoretical Chemistry Institute (University of Wisconsin).
47. Head-Gordon, M., Pople, J.A., and Frisch, M.J. (1988). MP2 energy evaluation by direct methods. Chem. Phys. Lett. *153*, 503–506. https://doi.org/10.1016/0009-2614(88)85250-3.
48. Frisch, M.J., Head-Gordon, M., and Pople, J.A. (1990). Semi-direct algorithms for the MP2 energy and gradient. Chem. Phys. Lett. *166*, 281–289. https://doi.org/10.1016/0009-2614(90)80030-H.
49. Frisch, M.J., Head-Gordon, M., and Pople, J.A. (1990). A direct MP2 gradient method. Chem. Phys. Lett. *166*, 275–280. https://doi.org/10.1016/0009-2614(90)80029-D.
50. Riplinger, C., Sandhoefer, B., Hansen, A., and Neese, F. (2013). Natural triple excitations in local coupled cluster calculations with pair natural orbitals. J. Chem. Phys. *139*, 134101. https://doi.org/10.1063/1.4821834.
51. Grimme, S. (2006). Semiempirical GGA-type density functional constructed with a long-range dispersion correction. J. Comput. Chem. *27*, 1787–1799. https://doi.org/10.1002/jcc.20495.
52. Grimme, S., Antony, J., Ehrlich, S., and Krieg, H. (2010). 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. https://doi.org/10.1063/1.3382344/926936.
53. Grimme, S., Ehrlich, S., and Goerigk, L. (2011). Effect of the damping function in dispersion corrected density functional theory. J. Comput. Chem. *32*, 1456–1465. https://doi.org/10.1002/JCC.21759.
54. McLean, A.D., and Chandler, G.S. (1980). Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, Z=11-18. J. Chem. Phys. *72*, 5639–5648. https://doi.org/10.1063/1.438980.
55. Krishnan, R., Binkley, J.S., Seeger, R., and Pople, J.A. (1980). Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. J. Chem. Phys. *72*, 650–654. https://doi.org/10.1063/1.438955.
56. Kendall, R.A., Dunning, T.H., and Harrison, R.J. (1992). Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. J. Chem. Phys. *96*, 6796–6806. https://doi.org/10.1063/1.462569.
57. Woon, D.E., and Dunning, T.H. (1993). Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. J. Chem. Phys. *98*, 1358–1371. https://doi.org/10.1063/1.464303.
58. Dunning, T.H., Jr., Peterson, K.A., and Wilson, A.K. (2001). Gaussian basis sets for use in correlated molecular calculations. X. The atoms aluminum through argon revisited. J. Chem. Phys. *114*, 9244–9253. https://doi.org/10.1063/1.1367373.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Data S1. Optimized Cartesian coordinates of all the single amino acid and large models used in the study.	Mendeley data repository	https://doi.org/10.17632/ychk3jfdfk.1
Software and algorithms		
Gaussian16	Frish et al.[44]	https://gaussian.com
ORCA 4.2.1	Neese, F.[45]	https://orcaforum.kofo.mpg.de/app.php/portal
NBO 7.0	Glendening et al.[46]	https://nbo6.chem.wisc.edu/
NBO7Pro@Jmol	Glendening et al.[46]	https://nbo6.chem.wisc.edu/
ChemDraw Professional 22.0	Revvity Signals	https://revvitysignals.com/products/research/chemdraw
Chimera 1.16	UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.	https://www.cgl.ucsf.edu/chimera/
Origin 2021	OriginLab Corporation	https://www.originlab.com/

METHOD DETAILS

A thorough benchmarking of the density functional theory (DFT) method used in this work was performed against correlated *ab initio* quantum chemical methods, *viz*. 2nd-order Møller-Plesset perturbation theory (MP2)[47–49] and domain-based local pair natural orbital coupled cluster method with single-, double, and triple-excitations (DLPNO-CCSD(T)).[50] Considering the vast data set, we restricted our benchmarking to the single amino acid model, like 4-HyP or other proline derivatives. The relative energy between the two pyrrolidine ring conformers, $\Delta E_{endo-exo}$ = *E*[4(*R*)-*endo*] – *E*[4(*R*)-*exo*], was taken as the key parameter for the benchmarking study. To select an appropriate DFT method, we began with a total of twenty-four density functionals, involving four pure generalized gradient approximation (GGA), seven hybrid GGA, three pure meta-GGA, seven hybrid meta-GGA, and three double-hybrid GGA functionals (Table S9). The non-covalent interaction due to the dispersion effect was considered during the DFT calculations using Grimme's empirical dispersion corrections, D2,[51] D30,[52] and D3 with Becke-Johnson damping (D3BJ).[53] Three different dispersion corrections were needed due to the lack of parameters for a single dispersion correction in the Gaussian 16[44] program for different functionals. On top of this, DFT functionals with varying percentages of Hartree-Fock exchange (%HF-X), starting from 0% to 55% were selected, to investigate the role of %HF-X in relative energies.

To obtain the most accurate DFT functional to be employed for the geometry and electronic structure, first, we sought to calibrate the DFT functionals against the gold-standard ab initio methods, i.e., DLPNO-CCSD(T). Importantly, the DLPNO-CCSD(T) method is capable of estimating relative energies (ΔE) within the chemical accuracy limit of ∼1 kcal/mol. However, as this method is available with the ORCA[45] suite of quantum chemical programs, we first calibrated the ab initio MP2 method available in the ORCA program (Figure S1). While optimized equilibrium geometries were obtained for all DFT and MP2 methods, the MP2-level geometries optimized in Gaussian 16 were used for the calibration of MP2 against DLPNO-CCSD(T) in ORCA by single point energy (SPE) calculations (Table S10). All equilibrium geometries were verified to be local minima through the harmonic vibrational frequency calculation. Pople's triple-ζ quality basis set with single diffuse and double polarization functions, 6-311+G(d,p)[54,55] was used for the DFT-level, and Dunning's correlation consistent augmented triple-ζ basis set with diffuse functions, aug-cc-pVTZ[56–58] was used for the MP2-level geometry optimization. The DLPNO-CCSD(T) energy calculations were performed in conjunction with Dunning's aug-cc-pVTZ basis set. The geometry optimization and energy calculations at the MP2 level of theory were accelerated with the resolution of the identity (RI) for the Coulomb part and a chain of spheres algorithm for the Hartree-Fock exchange part (RIJCOSX). The RIJCOSX approximation was also used to accelerate the energy calculations at the DLPNO-CCSD(T) level. The electrostatic effect due to solvation was taken into account during the geometry optimizations through the implicit solvent model based on density (SMD) and water as solvent.

Once the ab initio MP2 method is calibrated against the DLPNO-CCSD(T) energies, we selected the MP2 method as the standard for our benchmarking of DFT functionals. To access an extended set of DFT functionals, we re-optimized the 4-HyP conformers with twenty-four DFT functionals (*vide supra*) along with the ab initio MP2 level using the Gaussian 16 program. The basis set, dispersion correction, and effect due to solvation were used consistently as described above.

To investigate the electronic structure, chemical bonding, and stabilization due to $n\rightarrow\pi^*$ and $\sigma\rightarrow\sigma^*$ charge-transfer interactions, natural bond orbital (NBO) analysis was performed on the DFT-optimized geometries. The NBO 7.0[46] program suite was used for this purpose. The stabilization provided by $n\rightarrow\pi^*$ charge-transfer interactions was quantified using energy obtained from the second-order perturbation theory ($E^{(2)}$) implemented in the NBO 7.0 program. The natural bond orbitals were viewed with the "View" module of NBOPro7@Jmol visualization software.

QUANTIFICATION AND STATISTICAL ANALYSIS

Gaussian 16[44] and ORCA[45] programs were used for the geometry optimization. NBO 7.0[46] program was used to quantify the charge-transfer interactions.