Nanoscale

Hydrogen-Bond Configuration Modulates Energy Transfer Efficiency in Helical Protein Nanotubes

Journal:	*Nanoscale*
Manuscript ID	NR-ART-08-2020-006031.R1
Article Type:	Paper
Date Submitted by the Author:	05-Nov-2020
Complete List of Authors:	He, Jinlong; Utah State University Zhang, Lin; University of Pittsburgh, Mechanical Engineering and Materials Science; Shandong University, Department of Engineering Mechanics Liu, Ling; Temple University

Hydrogen-Bond Configuration Modulates Energy Transfer Efficiency in Helical Protein Nanotubes

Jinlong He,[ab] *Lin Zhang,*[cd] *and Ling Liu*[*a]

[a]*Department of Mechanical Engineering, Temple University, Philadelphia, PA 19122, United States*

[b]*Department of Mechanical and Aerospace Engineering, Utah State University, Logan, Utah 84322, United States*

[c]*Department of Engineering Mechanics, School of Civil Engineering, Shandong University, Jinan, 250061, P.R. China*

[d]*Department of Mechanical Engineering and Materials Science, University of Pittsburgh, Pittsburgh, PA 15261, United States*

*E-mail: ling.liu@temple.edu

Abstract. Energy transport in proteins is critical to a variety of physical, chemical, and biological processes in living organisms. While strenuous efforts have been made to study vibrational energy transport in proteins, thermal transport processes across the most fundamental building blocks of proteins, i.e. helices, are not well understood. This work studies energy transport in a group of "isomer" helices. The π-helix is shown to have the highest thermal conductivity, 110% higher than that of the α-helix and 207% higher than that of the 3_{10}-helix. The H-bond connectivity is found to govern thermal transport mechanisms including phonon spectral energy density, dispersion, mode-specific transport, group velocity, and relaxation time. The energy transport is strongly correlated with H-bond strength which is also modulated by the H-bond connectivity. These fundamental insights provide a novel perspective for understanding energy transfer in proteins and guiding a rational molecule-level design of novel materials with configurable H-bonds.

Introduction

Energy transport and atomic vibrations in biomacromolecules such as proteins are critical to a variety of physical, chemical and biological processes taking place in living organisms, from structural construction, catalysis, to molecular recognition.[1-5] In principle, the vibrational energy flow is strongly associated with structural characteristics of proteins including the conformation, amino acid sequence, and intermolecular affinities.[6-9] Due to their polymeric nature, proteins generate intricate macromolecular structures as they attain their biological states in the process of folding. Secondary structures are formed as an information bridge that links the primary sequence with substructures to shape the protein systems. Representative protein secondary structures include the β-sheet, 3_{10}-helix, α-helix and π-helix. They are stable and energetically favorable, primarily due to many interlocking hydrogen bonds (H-bonds) that exist between carbonyl (C=O) and amide (N−H) groups (**Fig. 1**g). Establishing the relationship between protein structures and energy transport pathways is an important area of fundamental interest to better understand protein reactivity and functionality. Further, the molecular insights gained from the study of nanoscale thermal transport in proteins may also inspire the design of new materials with novel structures and unprecedented properties.

Strenuous efforts have been made to probe energy transport in proteins and protein secondary structures such as the heme cooling cofactor,[10] peptide helices,[11] and the β-sheet.[12, 13] On one hand, evidences show that heat diffusion is integrated with allosteric communication pathways[14-16] which are closely related with the cellular and

physiological functions of proteins.[17-19] On the other hand, studies demonstrate that heat spreads through multiple pathways including the covalent backbone as well as the relatively weak inter/intra-molecular interactions in protein secondary structures. As such, the energy transport processes may be modulated by tuning the inter/intra-molecular interactions,[20] where the H-bond is arguably one of the most important factors. Although being a secondary bond, the H-bond is critical in proteins as it largely controls their structural stability,[21, 22] catalytic properties,[23-25] chemical reactivity,[26, 27] and energy transport.[8] Recent studies have reported that the H-bond significantly enhances thermal transport in the β-sheet of spider silk proteins[12] and a variety of polymer blends.[28]

Despite the progress, vibrational energy transport in helical protein structures are underexplored. As biomolecular "nanowires", protein helices differ from any other nanowires in that they uniquely feature tortuous covalent backbones interlocked by H-bonds. Different H-bond connectivities lead to different types of helices in the protein database, which include the 3_{10}-helix, α-helix, and the π-helix. While the 3_{10}-helix widely exists in many biological channels and membrane proteins[29] (e.g. ABC-ATPase SufC as shown in **Fig. 1**a), the α-helix is commonly found in globular proteins[30] (e.g. T4 Lysozyme as shown in **Fig. 1**b) and the π-helix is usually identified near functional sites of proteins[31] (e.g. PGRP-SA as shown in **Fig. 1**c). All composed of single amino acid chains, these helices are differentiated by the different atomic sites of forming H-bonds. For example, **Fig. 1**d-f illustrates ribbon schematics of the 3_{10}-helix, α-helix and π-helix, respectively. In the 3_{10}-helix, H-bonds are established between the CO group

of the *i*-th residue and the NH group of the (*i*+3)th residue. By comparison, H-bonds of the α-helix are formed between the *i*-th and the (*i*+4)th residues, while those of the π-helix are formed between the *i*-th and the (*i*+5)th residues. The different H-bond network leads to significantly different structural characteristics of the three helices including the number of amino acids per turn and the helical pitch. Understanding energy transport in such helical structures may fundamentally reveal how the H-bond, together with the helical backbone, dictates vibrational energy flow in proteins. A previous study has shown that the polymer chain conformation significantly influences thermal transport as a short straight chain is elongated to be a torturous long chain.[32]

This work investigates vibrational energy transfer and thermal transport mechanisms in the three key building blocks of proteins including the 3_{10}-, α-, and π-helices. To focus on the important effects of H-bond connectivity, three model helices are assumed which have the same amino acid sequence of poly-Glycine. The molecular structures of these helix "isomers" are shown in **Fig. 1**h-i. These "isomer" helices feature different H-bond connectivities as illustrated in **Fig. 1**d-f for the 3_{10}-, α-, and π-helices, respectively. Using the molecular dynamic (MD) simulation and vibrational energy transfer analysis, we demonstrate that the π-helix has the largest thermal conductivity, followed by the α-helix and then the 3_{10}-helix. Energy transfer mechanisms are explored via the analysis of phonon dispersion, spectral thermal conductivity, relaxation time, and group velocity. Results show that the different H-bond networks that form the "isomer" helices fundamentally change the behavior of both low-frequency acoustic and semi-optical phonons. The energy transport properties

of the three types of helices are strongly correlated with the different strengths of intramolecular H-bonds as revealed by quantum-chemical analyses based on the density functional theory (DFT) calculations. The results enrich our fundamental understanding of the energy transfer in protein structures and the critical role of interlocking H-bonds. The understanding may open new avenues to design novel biomacromolecules, e.g. synthetic spider silk with designed "blueprints" of amino acid sequences,[33] with extraordinary energy transfer properties for a wide range of applications.

Results and Discussion

Thermal Conductivity. Vibrational energy transport in "isomer" protein helices is first studied by quantifying and comparing how fast energy is conducted in these structures. **Fig. 1**k plots the thermal conductivity calculated by MD simulation for the 3_{10}-, α- and π-helices with various lengths including 2 nm, 3 nm, 5 nm, 10 nm, 20 nm, and 40 nm. Since helical segments in proteins are usually short with only a few nanometers in length, vibrational energy carriers, i.e. phonons, are not relaxed and their transport is constrained by and sensitive to the geometry. As shown in **Fig. 1**k, the thermal conductivity (k_x) of the helices increases nonlinearly with the length (L), conforming to the ballistic-diffusive empirical formula, $k_x = k_{x,\infty}\left(1 - e^{-L/L_c}\right)$, where $k_{x,\infty}$ is the thermal conductivity in the diffusive limit and L_c denotes a characteristic length of the ballistic-to-diffusive transition. Another ballistic-diffusive characteristic is demonstrated in **Error! Reference source not found.**, which shows that $1/k_x$ and $1/L$ are almost linearly correlated except when the helix length is extremely short. When the characteristic length is very short, thermal transport is strongly influenced by

multiple scattering mechanisms such as phonon-phonon scattering, multi-phonons scattering, and phonon-boundary scattering. These multi-scattering mechanisms are not mutually independent, leading to nonlinearity at large $1/L$ or short helix lengths. By fitting computational results of k_x with the empirical formula, thermal conductivity in the diffusive limit ($k_{x,\infty}$) is predicted to be 6.36 W m^{-1} K^{-1} for the 3_{10}-helix, 9.31 W m^{-1} K^{-1} for the α-helix, and 19.54 W m^{-1} K^{-1} for the π-helix. Among the three helical "isomers", the π-helix is the most thermally conductive, followed by the α-helix (52.35% lower $k_{x,\infty}$) and 3_{10}-helix (67.45% lower $k_{x,\infty}$). The phonon mean free path that corresponds to fully diffusive phonon transport is approximated by $L_{0.99}$, namely, the length at which thermal conductivity $k_x = 0.99k_{x,\infty}$. Based on the empirical formula with predicted $k_{x,\infty}$, the phonon mean free path is approximated as 139.39 nm, 148.10 nm and 160.02 nm for the 3_{10}-, α- and π-helices, respectively. The π-helix, which has the higher thermal conductivity, is also found to have longer phonon mean free path. The empirical equation is important because nearly all helix segments in proteins are short in length so they are in the ballistic-to-diffusive transition regime.

Phonon Vibrational Spectra. Effects of H-bond connectivity on phonon vibrational spectra can be found in the phonon dispersion curves plotted for the three isomer helices (**Fig. 2**a). The calculations are thermodynamically stable as no imaginary frequency is found in the curves. Only the regime with frequency < 30 THz is displayed. Phonons with higher frequencies are associated with localized interactions so they have low group velocities and short relaxation times.[34-37] Each helix shows three lowest-lying phonon branches including two transverse acoustic branches (TA1 and TA2) and one

longitudinal acoustic branch (LA). The two TA branches show almost linear dispersion, whereas the LA branch is nearly quadratic. Among the three acoustic branches, the LA branch has the largest slope, followed by the TA1 branch and then the TA2 branch. Although most optical phonon branches are relatively flat indicating negligible contributions to energy transport, all three helices show some optical branches that have significant slopes especially in the frequency range of 5-10 THz.

The group velocity for three acoustic branches and two selected optical branches are plotted in **Fig. 2**b. The two chosen optical branches (indicated by arrows in **Fig. 2**a) contribute significantly to energy transport as discussed in the following section using the mode-based thermal conductivity analysis. All branches except LA show bell-shaped group velocity curves that peak approximately in the middle. By comparison, the LA branch uniquely features an unsymmetrical curve with curve branching at the high frequency end. Comparing the curves, it is readily seen that the LA branch has the highest group velocity followed by the TA2 and TA1 branches, echoing the findings relating to the slope of dispersion curves. The two optical branches show similar group velocities as the TA1 and TA2 branches.

Importantly, the three isomer helices show critical differences as a result of their different H-bond connectivities. On the one hand, the π-helix exhibits the highest group velocity in all shown branches, followed by the α- and 3_{10}-helices. These results agree well with the ranking of the isomers in terms of their overall thermal conductivities. On the other hand, an interesting "redshift" is identified among the isomers. For example, the frequency at which the LA branch's group velocity peaks is found to redshift from

1.50 THz for the π-helix to 0.97 THz for the α-helix, and then further redshift to 0.71 THz for the 3_{10}-helix. The same trend persists in the other acoustic and optical phonon modes of importance to the energy transport processes.

Mode-Specific Contribution to Energy Transport. Based on the PSED analysis (**Error! Reference source not found.**), **Fig. 2**a uses the line color to plot the contribution to energy transport made by phonons of particular wave vector and frequency, i.e. $\tilde{k}_x(\boldsymbol{\kappa},\nu)$. It is apparent that the LA branch, especially when ξ = 0.05 ~ 0.4, makes the most prominent contribution to energy transport. The TA1 and TA2 branches in similar frequency ranges also contribute significantly. Summing up $\tilde{k}_x(\boldsymbol{\kappa},\nu)$ for the three branches, the LA branch is found to contribute about 38-40% thermal conductivity for all of the three helices, followed by the TA2 and TA1 branches which contribute about 13% and about 5-7%, respectively. The three acoustic branches together account for 57.70%, 58.21% and 59.99% of the overall thermal conductivity for the 3_{10}-, α- and π-helices, respectively. Therefore, heat is mainly carried by acoustic phonons in biomolecular helices.

In addition to acoustic phonons, some optical branches especially the two indicated by arrows in **Fig. 2**a also make important contributions as determined by their values of $\tilde{k}_x(\boldsymbol{\kappa},\nu)$. According to the mode-based thermal conductivity results shown in **Fig. 2**c, semi-optical branches in the range of 5-10 THz contribute 31.32%, 30.35% and 31.24% to the overall thermal conductivity for the 3_{10}-, α- and π-helices, respectively. **Fig. 2**d further shows that, for all helix isomers, the cumulative thermal conductivity increases significantly only in the frequency ranges of acoustic phonons and low-

frequency semi-optical phonons. The cumulative thermal conductivity curve almost plateaus after 30 THz. Therefore, acoustic phonons and semi-optical phonons with low frequencies make the most important contributions to energy transfer in biomolecular helices. They together account for about 90% of the overall thermal conductivity of the helix isomers under investigation.

The mode-based analysis also provides a means to verify the thermal conductivities calculated by MD simulation shown in **Fig. 1**k. Using the mode-based thermal conductivity results, the overall thermal conductivity can be calculated via $k_x = \sum_{\boldsymbol{\kappa}}\sum_{\nu}\tilde{k}_x(\boldsymbol{\kappa},\nu)$ which sums up the contributions made by all phonon branches. By using this approach, the overall thermal conductivity is found to be 5.84 W m^{-1} K^{-1} for the 3_{10}-helix, 8.47 W m^{-1} K^{-1} for the α-helix, and 18.43 W m^{-1} K^{-1} for the π-helix. These results are in good agreement with the conductivities calculated using MD directly. Both show profound effects of the H-bond configuration on energy transport.

Phonon Relaxation Time. The different H-bond connectivity in helices also influences phonon-phonon scattering. **Fig. 3**a plots full-mode relaxation time (τ) within the first Brillouin zone as a function of phonon frequency (f) for the 3_{10}-, α- and π-helices. On the one hand, relaxation time is in general shortened with increasing f in the low-frequency range where phonons carry most of the heat. The largest τ in this regime is about 10.93 ps for the 3_{10}-helix, 13.98 ps for the α-helix, and 20.20 ps for the π-helix. On the other hand, at higher frequencies, τ is well below a threshold that is frequency-independent. The threshold is about 4.45 ps, 7.04 ps and 11.06 ps for the 3_{10}-, α-, π-helices, respectively. In both frequency regimes, the isomer helices show different

upper bounds of τ as a result of their different H-bond networks. The π-helix is shown to have the longest τ, indicating the weakest phonon scattering. By contrast, the 3_{10}-helix is found to have the shortest τ and the strongest phonon scattering, mainly due to its slender helical structure with more frequent H-bonds along the backbone.

Fig. 3b further plots τ for selected phonon branches that contribute the most to energy transport, including three acoustic branches and two semi-optical branches indicated by arrows in **Fig. 2**a. All of these branches are in the low-frequency regime, and an approximately linear relationship is found between $\log(\tau)$ and $\log(f)$. Among the five branches, the acoustic branches are found to have the most significant frequency dependence of τ. As f increases, τ in these three branches descends quickly and almost monotonically indicating increasing phonon scattering. For all three isomer helices, the rate of descendance is the highest in LA with $\tau \propto f^{-1.77}$, very close to the well-known Klemens scaling relationship.[38] The descendance is slower in the other two acoustic branches, with $\tau \propto f^{-0.66}$ for TA1 and $\tau \propto f^{-0.73}$ for TA2. Compared with the acoustic branches, the two semi-optical branches show much slower descendance with approximately $\tau \propto f^{-0.19}$.

Interestingly, in the vicinity of Brillouin zone center, the $\tau - f$ data points for the LA branch show a tendency to diverge from the relationship of $\tau \propto f^{-1.77}$. When phonons are populated at small wave vectors, it has been previously discovered for a single polymer chain that Umklapp scattering is "frozen out" and Normal scattering dominates phonon interactions in the LA branch.[39] Hence, near the Brillouin zone center, LA phonon scattering is fundamentally different from that in the other regions

governed by a mixture of Umklapp and Normal scattering. Since the predominant Normal scattering is known to usually have high relaxation rates, τ of the LA branch drops at low $\boldsymbol{\kappa}$ and low f, leading to the downwards divergence as shown in **Fig. 3**b.

H-bonding Strength. Given the structure-governed thermal transport mechanisms discovered above, the H-bonding as a structural linker of helices is important to the energy transport processes. For quantitative understanding, the H-bonding is analyzed using two DFT-based methods, namely, the core-valence bifurcation index (CVBI) and the H-bond energy. Provided by equilibrium MD simulations, the initial structures of helices were intercepted into minimum sizes that each contains a H-bond (with four residues for the 3_{10}-helix, five for the α-helix, and six for the π-helix), as shown in **Error! Reference source not found.**. The structures were geometrically optimized by DFT, feeding wave functions into subsequent H-bond analyses.

Fig. 4a plots normalized electron density ranging from 0 to 1 for the three helices in the vicinity of four atoms forming the H-bond which include N (donor), H, O (acceptor), and C. Due to the different H-bond connectivities, geometry of the H-bond varies in the isomer helices despite the same type of H-bond in nature. Two important measures of a H-bond, i.e. the H-bond length (H···O) and the H-bond angle (N ─ H··· O) are found to be 2.12Å and 21.36°, 1.96 Å and 17.33°, and 1.83 Å and 15.69° for the 3_{10}-, α-, π-helices, respectively. **Fig. 4**b further plots the ELF distributions along the N ─ H covalent bond and the H···O H-bond, respectively, for the three helices. Their (3,−1) critical points are identified as the local minimum. Interestingly, while both ELF(C−V, D) and ELF(C−V, A) are close among the helices (variations < 4.32% and

< 3.54%, respectively), ELF(D−H, A) as one of the most important quantities of H-bonding is found to vary significantly, i.e. 0.057 for the 3_{10}-helix, 0.0681 for the α-helix and 0.0839 for the π-helix. Based on the ELF values, CVBI is calculated to be 0.0378, 0.0226 and 0.0071 for the 3_{10}-, α-, π-helices, respectively. On the one hand, the positive CVBIs indicate that the H-bonds in all helices under investigation are primarily of the electrostatic nature. On the other hand, the π-helix is found to have the lowest CVBI, suggesting the strongest H-bond among the three isomers.

Further, the H-bond energy is evaluated based on the QTAIM theory using the electron density at the H-bond critical point (ρ_{BCP}). ρ_{BCP} is found to increase from 0.0176 to 0.0235 and 0.0341 for the 3_{10}-, α- and π-helices, respectively. The H-bond energy is the strongest in the π-helix with -6.86 Kcal mol^{-1}, followed by the α-helix with -4.50 Kcal mol^{-1} and the 3_{10}-helix with -3.18 Kcal mol^{-1}. The results of H-bond energy agree well with the CVBI results. Importantly, the isomer helices demonstrate different H-bond strengths despite the same type of H-bond in nature, which contribute directly to the differences in energy transport mechanisms and the overall thermal conductivities as revealed in previous sections.

Conclusion

To summarize, this work provides a novel perspective on vibrational energy transport in proteins by studying three isomer helices. The isomers are specially designed to have the same simplest all-Glycine amino acid sequence interlocked by different H-bond networks representing three important types of helices that widely exist in proteins as building blocks. This design allows for an in-depth analysis of the role of H-bond

connectivity on the thermal conductivity and associated phonon transport mechanisms, while neglecting other factors including the amino acid sequence and side chains. The MD simulation predicts that the π-helix has the highest thermal conductivity of 19.54 W m^{-1} K^{-1} in the diffusive limit, which is 109.88% higher than that of the α-helix and 207.23% higher than that of the 3_{10}-helix. The thermal conductivities are highly length-dependent in the ballistic-to-diffusive transition regime, with a phonon mean free path of about 140-160 nm. The important contributions made by the H-bond to energy transport is underpinned by transport mechanisms including phonon spectral energy density, dispersion, mode-specific transport, group velocity, and relaxation time. The three acoustic branches together account for about 57.7-60% of the overall thermal conductivity, in which the LA branch contributes about two thirds. Additionally, semi-optical branches in the 5-10 THz range also contribute about 30.3-31.3%. The different H-bond connectivities cause an increase in group velocities and a blueshift in key frequencies defining the group velocity from the 3_{10}-helix to the π-helix. The unique H-bond connectivity in the π-helix also raises the phonon relaxation time indicating lower phonon scattering. Particularly, the LA branch diverges from the relationship of $\tau \propto f^{-1.77}$ at small wave vectors indicating predominant Normal scattering. Finally, calculations based on DFT and QTAIM reveal that the π-helix has the strongest H-bonds among the isomers, despite the fact that all these H-bonds are of the same type. The stronger H-bonds in the π-helix, together with the unique helical structure enabled by its H-bond connectivity, reduce scattering and increase group velocities of low-frequency acoustic and semi-optical phonons, drastically enhancing energy transport.

The insights provided by this study not only enrich the fundamental understanding of energy transport in proteins, but may also stimulate a rational design of helical molecular structures with configurable H-bond networks to enable new energy transport mechanisms and desired thermal properties.

Method Section

Models: Three "isomer" helices were created with the same amino acid sequence of poly-Glycine, where Glycine has a composition of –HN–$HC_{\alpha}H$–C=O–. The three helices include a 3_{10}-helix, a α-helix and a π-helix. They feature different atomic sites to form H-bonds (**Fig. 1**d-f): the 3_{10}-helix has H-bonds formed between the CO group at the *i*-th residue and the NH group at the (*i*+3)th residue; the α-helix has H-bonds between the *i*-th and (*i*+4)th residues; and the π-helix has H-bonds between the *i*-th and (*i*+5)th residues. The different sites of H-bonding vary the structural properties of helices. The 3_{10}-helix has 3 amino acids per turn, a helical pitch of 0.60 nm per turn, and a rise per helical residue of 0.20 nm. The α-helix has 3.6 residues per turn, a helical pitch of 0.54 nm per turn, and a rise per helical residue of 0.15 nm. The π-helix has 4.4 residues per turn, a helical pitch of 0.48 nm per turn, and a rise per helical residue of 0.12 nm.

Molecular dynamics (MD): MD simulation was conducted with LAMMPS.[40] Interatomic interactions were described by the CHARMM22 force field[41] that is widely used for peptide and protein simulations. Several studies have investigated thermal transport properties and vibrational modes of proteins using MD with the CHARMM force field; and their results are in excellent agreement with experimental measurements

and theoretical models.[42-46] Although every force field is different, the choice of the force field should not fundamentally change key conclusions drawn from this study (e.g. effects of hydrogen bond on thermal conductivities, group velocities, relaxation times, etc.) The particle-particle particle-mesh method (PPPM) was used to describe long range Coulomb interactions with a root mean square accuracy of 10^{-6}. Time step was set to be 0.25 fs, small enough to capture most vibrational modes. Periodic boundary conditions were imposed to the helices along all directions. Large spaces were left along transverse directions to prevent the model from interacting with its periodic images. Angular momentum was zeroed out at every step to eliminate potential rigid-body rotations. Initial molecular structures were optimized at 300 K to ensure the formation of H-bonds at selected sites leading to the 3_{10}-helix, α-helix and π-helix. The structures were then equilibrated at 300 K for 20 ns under the NVT ensemble. Identification of the H-bond was based on a geometric rule: (1) distance between the donor (D) atom and the acceptor (A) atom is < 3.5 Å; and (2) the angle of D-H-A is < 30°.

Thermal conductivity calculation: Nonequilibrium molecular dynamics (NEMD) simulations were performed to calculate thermal conductivities. The NEMD simulation setup is shown in **Error! Reference source not found.**a-c for the 3_{10}-, α- and π-helices, respectively. Each simulation system was divided into *N* slabs along the length direction. Atoms in the two end slabs were fixed to avoid an entropic state of helical chains. Thermal gradient was established by imposing Langevin thermostats of 325 K and 285 K in two slabs next to the fixed slabs, respectively, leading to a system temperature difference of ΔT = 40°C. The system was equilibrated for 20 ns under the NVE

ensemble, leading to a smooth thermal gradient between the heat source and heat sink slabs. Production runs were then performed for a duration of 10 ns. At every time step, the energy needed to keep ΔT a constant was evaluated and denoted by ΔE. Heat flow in the system, J, was calculated by $J = \Delta E/(A\Delta t)$, where Δt is the time step, and A is the effective cross sectional area of the helices as shown in **Error! Reference source not found.**d-f. Outside and inside radii of the helical tubes were determined by spatially and temporally averaging atomic positions during 2000 steps of MD calculations at equilibrium. The cross-sectional area, A, is calculated to be 20.36 Å^2, 22.82 Å^2 and 26.26 Å^2 for the 3_{10}-, α-, and π-helices, respectively. The thermal conductivity, k, was calculated by $k = J/(dT/dX)$. Here, dT/dX is the thermal gradient which is obtained by fitting a linear region of the temperature profile as shown in **Fig. 1**j.

Phonon spectral energy density (PSED): The PSED analysis was performed to obtain phonon properties including the dispersion curve, group velocity, relaxation time, and the mode-based thermal conductivity. The approach has been verified and applied for various materials systems.[34, 36, 47-50] The PSED analysis uses atomic velocities from the MD simulation which naturally includes anharmonic effects, phonon-phonon scattering, and Normal and Umklapp scattering. To calculate PSED, the smallest atomic group that represents a helical structure is defined as the unit cell, and each helix model has multiple unit cells along the length direction. The unit cell has the length of 13.37 Å, 12.08 Å and 10.38 Å, respectively, for the 3_{10}-, α- and π-helix. Based on lattice dynamic and Parseval's theorem,[51, 52] atomic trajectories in the real space are mapped to normal vibration modes in the frequency space, $q\left(\boldsymbol{\kappa}_{\nu};t\right)$. Here, $\boldsymbol{\kappa}$ is the wave vector,

ν labels the polarization branch, and *t* is time. The PSED is a weighted average of the contributions made by all atoms to the amplitude of normal modes, i.e. $\Phi(\boldsymbol{\kappa},\omega) = \frac{1}{2\pi\tau_0 N_T}\sum_\alpha\sum_j^{N_A} m_j\left|\int_0^{\tau_0}\sum_l^{N_T}\dot{u}_\alpha\binom{l}{j};t\right) \times \exp\left[i\boldsymbol{\kappa}\cdot\boldsymbol{r}_0\binom{l}{0} - i\omega t\right]dt\right|^2$. Here, $\Phi(\boldsymbol{\kappa},\omega)$ is the PSED as a function of $\boldsymbol{\kappa}$ and the angular frequency ω, τ_0 is the simulation time, N_T is the total number of unit cells, N_A is the total number of atoms in an unit cell, m_j is mass of the *j*-th atom, $\dot{u}_\alpha\binom{l}{j};t\right)$ denotes the velocity of the *j*-th atom in the *l*-th unit cell along the α-th direction at simulation time *t*, $\boldsymbol{r}_0\binom{l}{0}$ is the equilibrium position of the *l*-th unit cell. For one-dimensional helices, the wave vector is limited to $\kappa_x = 2\pi n_x/a_x N_T$, where a_x is the unit cell length along the *x*-direction, and n_x is an integer ranging from 0 to $N_T - 1$. Then, κ_x is mapped into $N_x/2$ points ranging from 0 to $2\pi/a_x$, which correspond to the first Brillouin zone in the real space from Γ to K point. Atomic velocities were collected from MD simulations of τ_0 = 1 ns under the NVE ensemble at 298 K.

Phonon dispersion: Based on the PSED profile, phono dispersion can be obtained by identifying peaks of the degenerate branches and connecting the peaks into continuous curves along the wave vector direction. The dispersion curves correlate phonon frequency ($f = \omega/2\pi$, in the unit of THz) with the reduced wave vector ($\xi = \kappa_x /(2\pi/a_x)$, dimensionless). Dispersion curves in this work were obtained at a resolution of 0.001 THz × 0.01.

Mode-based thermal conductivity: By solving the Boltzmann transport equation[53] with phonon relaxation time approximation,[54] the overall thermal conductivity is $k_x = \sum_{\boldsymbol{\kappa}}\sum_\nu \tilde{k}_x(\boldsymbol{\kappa},\nu)$, where $\tilde{k}_x(\boldsymbol{\kappa},\nu)$ is the thermal conductivity contributed by each phonon

mode that belongs to a phonon branch, ν, with a wave number of κ_x. The mode-based thermal conductivity has an expression of $\tilde{k_x}(\boldsymbol{\kappa},\nu) = c_{ph}v_{g,x}^2(\boldsymbol{\kappa},\nu)\tau(\boldsymbol{\kappa},\nu)$, where c_{ph} is the volumetric phonon specific heat, $v_{g,x}(\boldsymbol{\kappa},\nu)$ is the group velocity, and $\tau(\boldsymbol{\kappa},\nu)$ is the phonon relaxation time. In classical systems, $c_{ph} = k_B/V$, where k_B is Boltzmann constant and V is the system volume. The group velocity $v_{g,x}(\boldsymbol{\kappa},\nu) = \partial\omega/\partial\kappa_x$, which can be calculated based on the PSED result by using numerical differentiation.

Relaxation time: PSED can be reconstructed into a different form which is in terms of the phonon relaxation time, i.e. $\Phi(\boldsymbol{\kappa},f) = \sum_{\nu}^{3n} C(\boldsymbol{\kappa},\nu)/([4\pi\tau(\boldsymbol{\kappa},\nu)(f - f_0(\boldsymbol{\kappa},\nu))]^2 + 1)$, where f is the phonon frequency, $C(\boldsymbol{\kappa},\nu)$ is a mode-dependent constant, and $f_0(\boldsymbol{\kappa},\nu)$ is the frequency at which PSED reaches a localized peak. At each $\boldsymbol{\kappa}$, the PSED data is converted into a single-variable data series in terms of f. The data series has $3n$ local peaks where n is the number of atoms and $3n$ is the number of phonon branches in the system. The data series is fitted using multiple Lorentzian functions, $\Phi = {I}/{\left(1 + \left[(f - f_c)/\gamma\right]^2\right)}$, near local peaks. Here, I is the peak magnitude, f_c is the frequency at the peak, and γ is the half-width at the half-peak. Fitting was done by using the Fityk software with the Levenberg-Marquarde algorithm.[55] The reconstructed PSED function can be described as a superposition of the Lorentzian functions identified for all phonon branches at a given wave vector. Finally, relaxation time associated with all phonon modes, $\tau(\boldsymbol{\kappa},\nu)$, and the corresponding frequency, $f_0(\boldsymbol{\kappa},\nu)$, can be determined based on the parameters of the Lorentzian functions.

DFT calculation: Quantum-chemical calculations were conducted to

quantitatively study the H-bonds using Gaussian.[56] The calculations were based on the hybrid B3LYP exchange-correlation functional with Grimme's DFT-GD3(BJ) empirical dispersion correction. The B3LYP-GD3(BJ) theory has been widely used for studying H-bonds.[57] The ma-TZVPP basis set was employed. It is the "minimally augmented" version of the def2-TZVPP basis set,[58, 59] for which *s* and *p* type diffuse basis functions are added to non-H atoms. Using this theory, equilibrium geometries of the 3_{10}-, α- and π-helices were obtained by geometric optimization. Frequency calculations were performed to confirm structural stabilization. Wave functions were output for subsequent calculations of the H-bond chemistry.

Quantum-chemical H-bond calculation*:* Based on the wave functions calculated by DFT, strength of the H-bond was quantified using two numeric indices: (1) the core-valence bifurcation index (CVBI), and (2) the H-bond energy evaluated at the bond critical point. On the one hand, the CVBI is a widely used approach based on the electron localization function (ELF) theory.[60, 61] The ELF takes the form of $\eta(r) = 1/\left\{1 + \left[D(r)/D_0(r)\right]^2\right\}$, where $D(r)$ is the Pauli kinetic energy density and $D_0(r)$ is the Thomas-Fermi kinetic energy density. For a H-bond taking the general form of D−H⋯A, bifurcation points of the ELF are first identified through a topological analysis, based on which the (3,−1) critical points are located. The CVBI is calculated by CVBI = ELF(C−V, D) − ELF(DH⋯A),[62, 63] where ELF(C−V, D) is the core-valence bifurcation value at donor (D), and ELF(DH⋯A) stands for the valence-valence bifurcation value at the H-bond. The CVBI is a numeric indicator of the strength of the H-bond; the lower CVBI, the stronger H-bond. A negative CVBI suggests that the H-bond has a covalence

characteristic, whereas a positive CVBI means that the H-bond has an electrostatic nature. On the other hand, the H-bond strength can be evaluated directly based on the electron density at the bond critical point, i.e. the (3,−1) critical point associated with DH···A, on the basis of the quantum theory of atoms in molecules (QTAIM). The equation takes the form of $E_{\mathrm{HB}} = A\rho_{\mathrm{BCP}} + B$, where E_{HB} is the H-bond energy, ρ_{BCP} is the electron density at the bond critical point, and A and B are parameters depending on the H-bond type and the level of quantum chemical calculation.[64-67] In this work, A = –223.08 and B = 0.7423 were adopted,[68] which were optimized and proven to effectively characterize the H-bond strength in biologically important substances. The calculation was performed by using the MULTIWFN program at the same level of theory.[69]

Acknowledgements

This work is financially supported by CAREER Award No. CBET-1751610 from the National Science Foundation. This research made use Idaho National Laboratory computing resources which are supported by the Office of Nuclear Energy of the U.S. Department of Energy and the Nuclear Science User Facilities under Contract No. DE-AC07-05ID14517

Author Contributions

L.L conceived the idea, and proposed the research methodology and guidance. J.H. and

L.Z. performed the simulation and calculations. All authors contributed to the analysis and discussion of the data and the writing of the manuscript.

Competing Interests

The authors declare no conflict of interest.

Author Information

Corresponding Author: Ling Liu (E-mail: ling.liu@temple.edu)

References

1. P. Kukura, D. W. McCamant, S. Yoon, D. B. Wandschneider and R. A. Mathies, *Science*, 2005, **310**, 1006.
2. P. M. Champion, *Science*, 2005, **310**, 980.
3. G. S. Engel, T. R. Calhoun, E. L. Read, T.-K. Ahn, T. Mančal, Y.-C. Cheng, R. E. Blankenship and G. R. Fleming, *Nature*, 2007, **446**, 782-786.
4. R. J. D. Miller, *Annu. Rev. Phys. Chem.*, 1991, **42**, 581-614.
5. R. Mülhaupt, *Macromol. Chem. Phys.*, 2013, **214**, 159-174.
6. D. M. Leitner, *Annu. Rev. Phys. Chem.*, 2008, **59**, 233-259.
7. M. C. Gather and S. H. Yun, *Nat. Commun.*, 2014, **5**, 5722.
8. A. S. Davydov, *J. Theor. Biol.*, 1977, **66**, 379-387.
9. G. Li, D. Magana and R. B. Dyer, *Nat. Commun.*, 2014, **5**, 3100.
10. D. E. Sagnella and J. E. Straub, *J. Phys. Chem. B*, 2001, **105**, 7057-7063.
11. E. H. G. Backus, P. H. Nguyen, V. Botan, A. Moretto, M. Crisma, C. Toniolo, O. Zerbe, G. Stock and P. Hamm, *J. Phys. Chem. B*, 2008, **112**, 15487-15492.
12. L. Zhang, T. Chen, H. Ban and L. Liu, *Nanoscale*, 2014, **6**, 7786-7791.
13. L. Zhang, Z. Bai, H. Ban and L. Liu, *PCCP*, 2015, **17**, 29007-29013.
14. H. S. Lee and Y. Zhang, *Proteins Struct. Funct. Bioinf.*, 2012, **80**, 93-110.
15. N. Ota and D. A. Agard, *J. Mol. Biol.*, 2005, **351**, 345-354.
16. J. Liu, G. J. Tawa and A. Wallqvist, *PLoS ONE*, 2013, **8**, e81980.
17. K. Seldeen, B. Deegan, V. Bhat, D. Mikles, C. McDonald and A. Farooq, *The FEBS journal*, 2011, **278**, 2090-2104.
18. E. Laine, C. Auclair and L. Tchertanov, *PLoS Comput. Biol.*, 2012, **8**, e1002661.
19. H. Kong, K. Xia, N. Ren, Y. Cui, R. Liu, Q. Li, M. Lv, J. Shi, Q. Yan, Z. Cui, C. Fan, Y. Zhu and L. Wang, *Nanoscale*, 2018, **10**, 18055-18063.
20. H. Ma, E. O'Donnel and Z. Tian, *Nanoscale*, 2018, **10**, 13924-13929.
21. C. Pace, J. M. Scholtz and G. Grimsley, *FEBS Lett.*, 2014, **588**.
22. P. Li, H. Chen, J. A. Schott, B. Li, Y. Zheng, S. M. Mahurin, D.-e. Jiang, G. Cui, X. Hu, Y. Wang, L. Li and S. Dai, *Nanoscale*, 2019, **11**, 1515-1519.
23. E. L. Ash, J. L. Sudmeier, E. C. De Fabo and W. W. Bachovchin, *Science*, 1997, **278**, 1128.
24. G. Miño and R. Contreras, *Chem. Phys. Lett.*, 2010, **486**, 119-122.
25. W. Foster, K. Miyazawa, T. Fukuma, H. Kusumaatmaja and K. Voïtchovsky, *Nanoscale*, 2020, **12**, 5452-5463.
26. S. Seo, K. Lee, M. Min, Y. Cho, M. Kim and H. Lee, *Nanoscale*, 2017, **9**, 3969-3979.
27. J. Han, Y. Shen and W. Feng, *Nanoscale*, 2016, **8**, 14139-14145.
28. H. Huang, Y. Hu, J. Zhang, H. Sato, H. Zhang, I. Noda and Y. Ozaki, *J. Phys. Chem. B*, 2005, **109**, 19175-19183.
29. R. Vieira-Pires and J. Morais-Cabral, *J Gen Physiol*, 2010, **136**, 585-592.
30. M. Levitt and C. Chothia, *Nature*, 1976, **261**, 552-558.
31. R. Cooley, D. Arp and P. Karplus, *J. Mol. Biol.*, 2010, **404**, 232-246.
32. X. Wei and T. Luo, *PCCP*, 2019, **21**, 15523-15530.

33. F. Teulé, A. R. Cooper, W. A. Furin, D. Bittencourt, E. L. Rech, A. Brooks and R. V. Lewis, *Nat. Protoc.*, 2009, **4**, 341-355.
34. J. Goicochea, M. Madrid and C. Amon, *J. Heat Transfer*, 2010, **132**, 012401.
35. Y. Zhou, X. Zhang and M. Hu, *Phys. Rev. B*, 2015, **92**, 195204.
36. J. A. Thomas, J. E. Turney, R. M. Iutzi, C. H. Amon and A. J. H. McGaughey, *Phys. Rev. B*, 2010, **81**, 081411.
37. A. Aiyiti, S. Hu, C. Wang, Q. Xi, Z. Cheng, M. Xia, Y. Ma, J. Wu, J. Guo, Q. Wang, J. Zhou, J. Chen, X. Xu and B. Li, *Nanoscale*, 2018, **10**, 2727-2734.
38. P. G. Klemens and F. E. Simon, *Proc R Soc Lond A, Math Phys Sci*, 1951, **208**, 108-133.
39. X. Wang, M. Kaviany and B. Huang, *Nanoscale*, 2017, **9**, 18022-18031.
40. S. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1-19.
41. A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586-3616.
42. D. M. Leitner and T. Yamato.
43. M. Kondoh, M. Mizuno and Y. Mizutani, *The Journal of Physical Chemistry Letters*, 2016, **7**, 1950-1954.
44. J. E. Elenewski, K. A. Velizhanin and M. Zwolak, *Nat. Commun.*, 2019, **10**, 4662.
45. X. Yu and D. M. Leitner, *J. Phys. Chem. B*, 2003, **107**, 1698-1707.
46. X. Yu and D. M. Leitner, *J. Chem. Phys*, 2005, **122**, 054902.
47. N. de Koker, *Phys. Rev. Lett.*, 2009, **103**, 125902.
48. J. Shiomi and S. Maruyama, *Phys. Rev. B*, 2006, **73**, 205420.
49. A. J. H. McGaughey and M. Kaviany, *Phys. Rev. B*, 2004, **69**, 094303.
50. D. Donadio and G. Galli, *Phys. Rev. Lett.*, 2009, **102**, 195901.
51. M. T. Dove, *Introduction to Lattice Dynamics*, Cambridge University Press, Cambridge, 1993.
52. W. Rudin, *Real and complex analysis, 3rd ed*, McGraw-Hill, Inc., 1987.
53. J. E. Turney, E. S. Landry, A. J. H. McGaughey and C. H. Amon, *Phys. Rev. B*, 2009, **79**, 064301.
54. G. Srivastava, *The Physics of Phonons*, 2019.
55. M. Wojdyr, *J. Appl. Crystallogr.*, 2010, **43**, 1126-1128.
56. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N.

Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Journal*, 2016.
57. L. Goerigk and S. Grimme, *J. Chem. Theory Comput.*, 2011, **7**, 291-309.
58. F. Weigend and R. Ahlrichs, *PCCP*, 2005, **7**, 3297-3305.
59. J. Zheng, X. Xu and D. Truhlar, *Theor. Chem. Acc.*, 2011, **128**, 295-305.
60. A. D. Becke and K. E. Edgecombe, *J. Chem. Phys*, 1990, **92**, 5397-5403.
61. T. Lu and F. Chen, *Acta Phys. Chim. Sin.*, 2011, **27**, 2786.
62. F. Fuster and B. Silvi, *Theor. Chem. Acc.*, 2000, **104**, 13-21.
63. M. E. Alikhani, F. Fuster and B. Silvi, *Struct. Chem.*, 2005, **16**, 203-210.
64. T. Y. Nikolaienko, L. A. Bulavin and D. M. Hovorun, *PCCP*, 2012, **14**, 7441-7447.
65. O. h. Brovarets, Y. Yurenko and D. Hovorun, *J. Biomol. Struct. Dyn.*, 2014, **32**, 1-29.
66. Ahmed T. Ayoub, Travis J. A. Craddock, M. Klobukowski and J. Tuszynski, *Biophys. J.*, 2014, **107**, 740-750.
67. S. Emamian, T. Lu, H. Kruse and H. Emamian, *J. Comput. Chem.*, 2019, **40**, 2868-2881.
68. A. V. Afonin, A. V. Vashchenko and M. V. Sigalov, *Org. Biomol. Chem.*, 2016, **14**, 11199-11211.
69. T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580-592.
70. H. Berman, K. Henrick and H. Nakamura, *Nat. Struct. Mol. Biol.*, 2003, **10**, 980-980.
71. H. Berman, K. Henrick, H. Nakamura and J. L. Markley, *Nucleic Acids Res.*, 2006, **35**, D301-D303.
72. P. D. B. c. ww, *Nucleic Acids Res.*, 2018, **47**, D520-D528.
73. S. Watanabe, A. Kita and K. Miki, *J. Mol. Biol.*, 2005, **353**, 1043-1054.
74. L. Liu, W. A. Baase and B. W. Matthews, *J. Mol. Biol.*, 2009, **385**, 595-605.
75. C.-I. Chang, S. Pili-Floury, M. Hervé, C. Parquet, Y. Chelliah, B. Lemaitre, D. Mengin-Lecreulx and J. Deisenhofer, *PLoS Biol.*, 2004, **2**, e277.

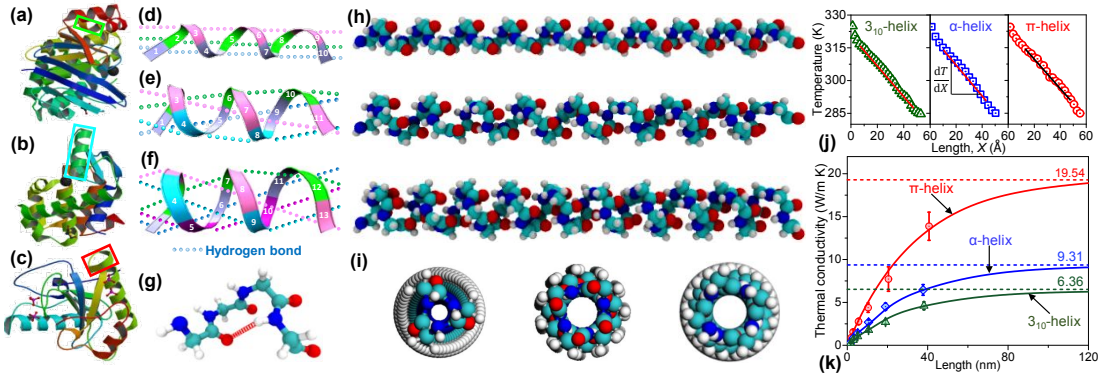


Fig. 1. Helix segments in example proteins:[70-72] a) ABC-ATPase SufC[73] with the 3_{10}-helix (in green box), b) T4 Lysozyme[74] with the α-helix (in cyan box), and c) PGRP-SA[75] with the π-helix (in red box). Different types of helices are interlocked by different H-bond networks: d) the 3_{10}-helix has H-bonds connecting residues *i* and *i* + 3; e) the α-helix has H-bonds connecting residues *i* and *i* + 4; and f) the π-helix has H-bonds connecting residues *i* and *i* + 5. g) Illustration of a H-bond between the carbonyl (C=O) and amide (N−H) groups. Three isomer helices with the same amino acid sequence of poly-Glycine but different H-bond connectivities to form three types of helices: h) side and i) cross-sectional views. j) Temperature profiles of the helices in NEMD simulation. k) Thermal conductivity of the 3_{10}-, α- and π-helices with varying lengths. Curves are fitted against an empirical equation of the length-dependent thermal conductivity. Dashed lines show thermal conductivities in the diffusive limit.

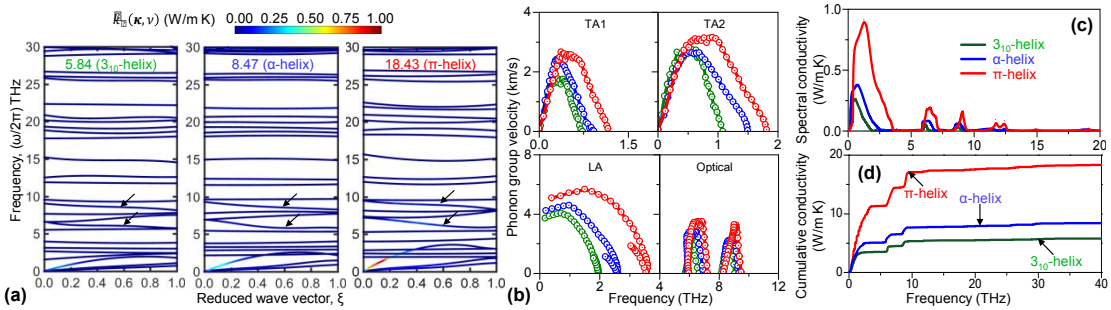


Fig. 2. a) Phonon dispersion curves of three isomer helices in the first Brillouin zone. Colors show mode-by-mode contributions to the thermal conductivity, i.e. $\tilde{k}_x(\boldsymbol{\kappa},\nu)$. Arrows indicate two semi-optical branches with significant $\tilde{k}_x(\boldsymbol{\kappa},\nu)$. b) Frequency-dependent group velocities for three acoustic branches and two selected optical branches of the 3_{10}-helix (green), α-helix (blue), and π-helix (red). c) Spectral thermal conductivity and d) cumulative thermal conductivity of the three isomer helices.

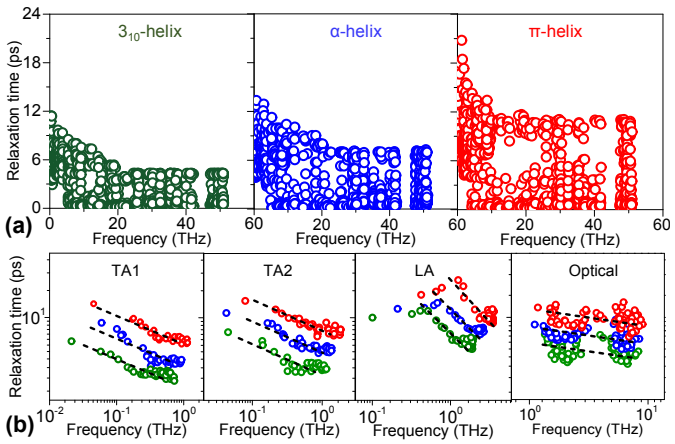


Fig. 3. a) Frequency-dependent relaxation time of three isomer helices. b) Relaxation time for three acoustic branches and two selected optical branches of the 3_{10}-helix (green), α-helix (blue), and π-helix (red). Back dashed lines are fitted against a relationship of $\tau \propto f^{-\alpha}$.

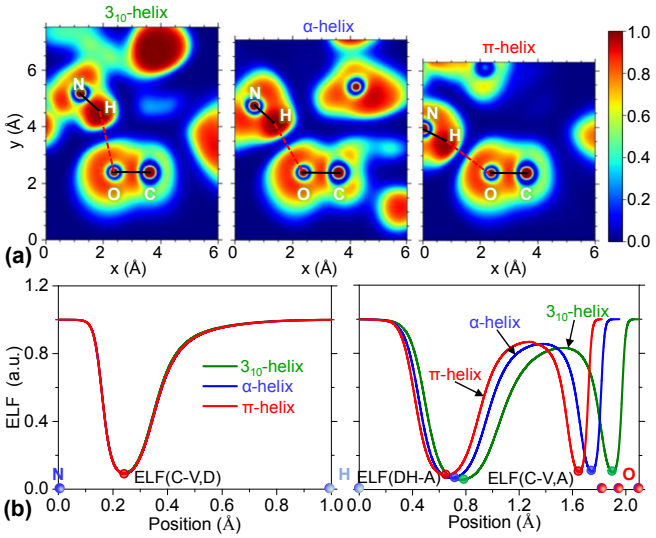
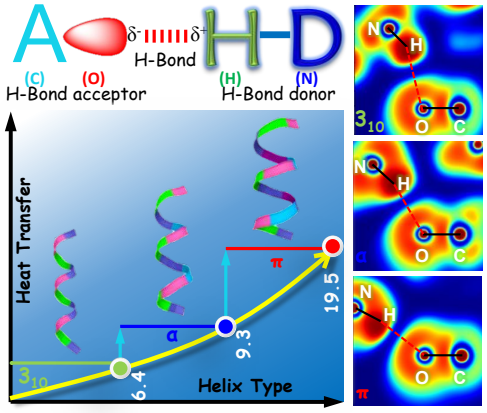


Fig. 4. a) Normalized electron density for the three isomer helices in the vicinity of four atoms forming the H-bond. b) The electron localization function (ELF) distributions along the N ─ H covalent bond and the H···O H-bond. Local minimum of these curves define ELF(C−V, D), ELF(DH···A), and ELF(C−V, A). Spheres on the horizontal axis indicate atomic locations.

TOC Graphics



As fundamental building blocks of proteins, helices take different forms including the 3_{10}-helix, α-helix, and π-helix that feature distinct H-bond networks. The H-bond connectivity is shown to dictate energy transfer in these helices as a part of energy transfer in proteins. Thermal conductivity is tuned in a large range by configuring H-bonds to modulate phonon spectral energy density, dispersion, mode-specific transport, group velocity, relaxation time, and H-bond strength