

# タンパク質ベースの材料における微視的な熱輸送特性

Tingting WANG

名古屋大学理学研究科

## 1. Introduction

The thermal transport properties of proteins are intricately linked to their structures and dynamic behavior. Advances in computational simulation techniques, such as the normal mode approach and molecular dynamics (MD)<sup>1-4</sup> based on the Irving-Kirkwood type formulation, have enabled us to explore the thermal transport properties of polymers, including proteins, at the atomic level. It's important to note that folded native proteins do not merely function as thermal conduits; they also operate as sophisticated molecular nanomachines within cells. While intrinsically disordered proteins behave differently, the molecular functions of globular proteins emerge primarily through the process of protein folding. Consequently, we are particularly interested in understanding the types of "communication" that occur through native contacts within folded proteins. We also anticipate that these processes may be related to specific patterns of intramolecular thermal transport, akin to vibrational energy transport, both of which are governed by the protein's structure and dynamics. However, the mechanisms by which vibrational energy and heat flow within protein molecules remain a topic of active debate.<sup>5-7</sup> This is partly due to the inherent heterogeneity and anisotropy of proteins. Nonuniform energy flow occurs along the polypeptide chain and through non-bonded native contacts, resulting in varying transport efficiencies across different regions within the protein interior, where interactions such as van der Waals forces, electrostatic forces, and hydrogen bonding come into play. Therefore, it is essential to first grasp the fundamental properties of protein heat currents before delving deeper into the study of molecular functions in the context of thermal transport or vibrational energy transfer.

To quantitatively evaluate the competition between different thermal transport pathways—specifically, covalent bonds and non-bonded contacts—and various types of interactions within proteins, we introduced the concept of inter-residue thermal conductivity. This concept is based on the autocorrelation function of the inter-residue heat current between pairs of residues in non-bonded native contact. Utilizing equilibrium molecular dynamics (EMD) simulations and the Green-Kubo formula, we examined the inherent thermal transport properties through residue contacts in HP36.

## 2. Methods

### 2.1 Inter-residue heat current analysis

The site-selective heat/energy current analysis is based on the atomistic expression

of the heat/energy current for molecular system. The inter-atomic heat current between atoms  $i$  and  $j$ , denoted as  $\mathbf{h}_{ij}$ , can be expressed as:

$$\mathbf{h}_{ij} = (\mathbf{r}_i - \mathbf{r}_j) \left\{ \frac{1}{2} \mathbf{F}_{ij} \cdot (\mathbf{v}_i + \mathbf{v}_j) \right\}, \quad \mathbf{h} = \sum_{i < j} \mathbf{h}_{ij} \quad (1)$$

, where  $\mathbf{h}$  gives the macroscopic heat current of the entire molecule,  $\mathbf{F}_{ij}$  is the force acting on atom  $i$  from atom  $j$ ,  $\mathbf{r}_i$  ( $\mathbf{r}_j$ ) and  $\mathbf{v}_i$  ( $\mathbf{v}_j$ ) are position vector and velocity of atom  $i$  ( $j$ ), respectively. It should be noted that the derivation of mathematical expression of the pairwise interatomic forces for macromolecules is not straightforward because their potential energy functions contain multibody terms. To explore possible relationships between thermal energy transport and molecular mechanism of protein functions, it would be helpful to introduce coarse grained expression than Eq. (1). Thus, we define the heat current between a pair of residues  $\alpha$  and  $\beta$ , ( $1, \dots, \alpha(\beta), \dots, N_r$ ), expressed as:

$$\mathbf{h}_{\alpha,\beta} = \sum_{i \in \alpha} \sum_{j \in \beta} \mathbf{h}_{ij}, \quad (2)$$

where  $N_r$  is the total number of residues. By using the following quantity:

$$\Lambda_{\alpha,\beta} = \int_0^\infty \langle \mathbf{h}_{\alpha,\beta}(t) \cdot \mathbf{h}_{\alpha,\beta}(0) \rangle dt, \quad (3)$$

the inter-residue thermal conductivity,  $\lambda_{\alpha,\beta}$ , can be defined as:

$$\lambda_{\alpha,\beta} = \frac{1}{3(V_\alpha + V_\beta)k_B T^2} \int_0^\infty \langle \mathbf{h}_{\alpha,\beta}(t) \cdot \mathbf{h}_{\alpha,\beta}(0) \rangle dt = \frac{\Lambda_{\alpha,\beta}}{3(V_\alpha + V_\beta)k_B T^2} \quad (4)$$

, where  $V_\alpha$  ( $V_\beta$ ) is the volume of residue  $\alpha$  ( $\beta$ ),  $k_B$  is the Boltzmann constant,  $T$  is the absolute temperature. The angle brackets represent ensemble average. We used the VLDP (Voronoi Laguerre Delaunay Protein) method to calculate the volume of each atom, and the volume of each residue was calculated as a summation of the volume of the constituent atoms.

## 2.2 Equilibrium MD simulations

### 2.2.1 Protein model

For the analysis of thermal energy transport properties, we build a protein system of villin headpiece (HP36) based on the nuclear magnetic resonance (NMR) structure (PDB code: 1VII). The protonation states of all ionizable residues were kept in their standard states at pH = 7. After the protein molecule was solvated by a truncated octahedral box of TIP3P water molecules with 2 sodium and 4 chloride ions, the total number of atoms for the simulation box became 7589. Amber ff19SB force field was used for protein molecule. For calculations of long-range electrostatic interactions, the particle mesh Ewald (PME) method was used under the periodic boundary condition. A 9 Å distance cutoff was used to calculate the non-bonded particle-particle interactions. All MD simulations were performed by using the AMBER 19 program.

### 2.2.2 MD Simulations

Energy minimization of the system was performed through three-stage optimization for the positions of (1) hydrogen, (2) sidechain, and (3) mainchain atoms. After the energy minimization, we generated five different Maxwell-Boltzmann atomic velocity

distributions at  $T = 0.1$  K. Each of them was used as an initial condition for the subsequent MD simulation during which the temperature was gradually increased from  $T = 0.1$  to 300 K for 50 ps at constant volume with positional restraints imposed on the mainchain atoms. The equations of motion were integrated with a time step of 2.0 fs with the SHAKE constraints applied for the bonds involving hydrogen atoms.

For each branch of the five simulation runs, a *NVT* simulation followed at  $T = 300$  K for 50 ps with keeping the same positional restraints, then another *NVT* simulation was conducted for 200 ps at  $T = 300$  K with no positional restraints. Finally, isothermal-isobaric (*NPT*) simulation was performed for 700 ps at  $T = 300$  K,  $P = 0.987$  atm with Nose-Hoover thermostat and barostat.

Starting from each end point of the previous simulation runs, we continued an *NPT* MD simulation for 56 ns consisting of 50-ns, 1-ns, and 5-ns runs with time steps of 2.0, 0.5, and 0.5 fs, respectively. After the first segment of 50-ns, the SHAKE constraints were switched off for the protein. From the last 5-ns trajectory, 10 snapshots of atomic coordinates and velocities were extracted every 0.5 ns. Thus, we obtained 50 ( $= 5 \times 10$ ) initial conditions, from each of which we performed constant volume, constant energy (*NVE*) MD simulations for 1 ns with a time step of 0.5 fs, while atomic coordinates and velocities were saved every 1.0 fs with the time points of velocity snapshots adjusted to those of atomic coordinates.

## 2.3 Classification of nonbonded residue-residue interactions

### 2.3.1 Interacting residue pairs

In this study, we constructed three different sets,  $L$ ,  $M$ , and  $S$ , of non-bonded residue pairs: Dataset  $L$  was defined as residue pairs,  $(\alpha, \beta)$ , with the closest interatomic distance,  $r_{\min}(\alpha, \beta)$ ,  $\leq 6$  Å in at least one frame in the fifty *NVE* MD trajectories, while Dataset  $M$  ( $S$ ) as those with  $\langle r_{\min}(\alpha, \beta) \rangle \leq 6(4)$  Å, where the angle brackets represent the average of all frames appeared in the all trajectories. Accordingly, the number of pairs became 268, 157, and 104 for Dataset  $L$ ,  $M$ , and  $S$ , respectively. Dataset  $S$  mainly contains residue pairs in non-bonded contact, while datasets  $M$  and  $L$  include weakly interacting residue pairs in addition to those in direct contact.

### 2.3.2 Interaction types

The residue-residue interaction types are classified into four types based on the protein structure that appeared in MD simulation trajectories: H-bonding,  $\pi$  stacking, electrostatic interactions between charged or polar residue, and hydrophobic interactions. H-bonded pairs,  $(\alpha, \beta)$ s, were identified by the *cpptraj* module of AmberTools 19, with a geometric criterion, *i.e.*, the distance between atoms  $X \in \alpha$  and  $Y \in \beta$  is less than 3.0 Å and the angle  $X-H \cdots Y$  or  $Y-H \cdots X$  falls within the range of  $145^\circ - 180^\circ$ , where each of  $X$  and  $Y$  are either oxygen or nitrogen atoms. The  $\pi$  stacking interactions were identified in the NMR structure of HP36 using the RING3.0 server<sup>51</sup>. A relaxed distance threshold was applied, where the centroid-to-centroid distance between two aromatic rings was required to be within 7.0 Å. Any residue pairs not

labeled as “H-bonding”, “ $\pi$  stacking”, or “electrostatic” were classified as “hydrophobic interactions”.

### 3. Results

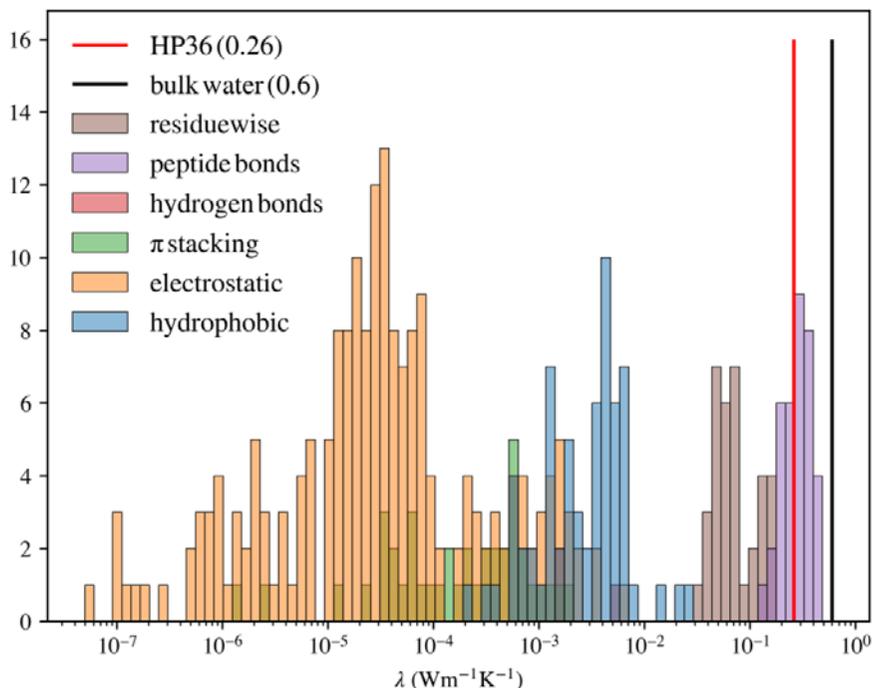


Figure 1: Histogram of occurrence frequency of local thermal conductivities, *i. e.*,  $\lambda_{\alpha,\beta}$ , for different types of residue pairs ( $\alpha$ ,  $\beta$ )s in dataset L. The residue-wise thermal conductivities,  $\lambda_{\alpha,\alpha}(1, \dots, \alpha, \dots, N_r)$ , and those between adjacent residue pairs  $\lambda_{\alpha,\alpha+1}(1, \dots, \alpha, \dots, N_r - 1)$  were taken from the previous study for comparison.<sup>2</sup> Also, the value of the thermal conductivity of the entire HP36 molecule is indicated by the red dashed line, as well as that of bulk water by the black dashed line. It should be noted that we used  $V_\alpha$ , instead of  $(V_\alpha + V_\beta)$ , in Equation 4 to evaluate  $\lambda_{\alpha,\alpha}$ .

Figure 1 shows the occurrence frequency of the calculated local thermal conductivities. The histogram reveals that the H-bonding pairs have the largest average value of among all types of the nonbonded interactions, with a peak at around  $2.5 \times 10^{-2} \text{ W m}^{-1} \text{ K}^{-1}$ , followed by those with electrostatic interactions with values ranging from  $10^{-5}$  to  $10^{-3} \text{ W m}^{-1} \text{ K}^{-1}$  with having a broader and less sharply peaked distributions than that of H-bonding pairs. In the hydrophobic core of HP36, there are three  $\pi$  stacking contacts: Phe7-Phe11 ( $\lambda_{7,11} = 4.9 \times 10^{-3}$ ), Phe7-Phe18 ( $\lambda_{7,11} = 1.5 \times 10^{-3}$ ), and Phe11-Phe18 ( $\lambda_{11,18} = 1.5 \times 10^{-3}$ ), with  $\lambda_{\alpha,\beta}$  values comparable with those of H-bonding contacts. Hydrophobic interactions constitute the majority of non-covalent residue pairs, showing a broad distribution spanning from  $10^{-7}$  to  $10^{-2} \text{ W m}^{-1} \text{ K}^{-1}$ . Their  $\lambda_{\alpha,\beta}$  values are relatively smaller, with a peak at around  $5.9 \times 10^{-4} \text{ W m}^{-1} \text{ K}^{-1}$ , than the other types of

interactions. In summary, the occurrence frequency of local thermal conductivities exhibits a strikingly broad distribution, where different types of non-bonded interactions have significantly different contributions to the thermal transport in the protein, with H-bonds playing a dominant role.

The molecular mechanism of thermal energy transport in proteins remains a subject of considerable debate.<sup>7-9</sup> To address this issue, we conducted a site-selective heat current analysis (Figure 1), which highlights significant heterogeneity in the distribution of local thermal transport coefficients at the microscopic level. Specifically, the thermal conductivities for individual residues, denoted as  $\lambda_{(\alpha, \alpha)}$ , range from 0.08 to 0.3 W m<sup>-1</sup> K<sup>-1</sup>, while those for peptide-bonded residue pairs,  $\lambda_{(\alpha, \alpha+1)}$ , range from 0.01 to 0.08 W m<sup>-1</sup> K<sup>-1</sup>. Both of these ranges are higher than those observed for residue pairs connected solely by non-bonded interactions, suggesting that the polypeptide chain serves as the primary conduit for thermal energy transport. However, it is important to note three exceptions:  $\lambda_{3,6}=2.4\times 10^{-2}$ ,  $\lambda_{4,15}=1.3\times 10^{-2}$ , and  $\lambda_{5,8}=2.1\times 10^{-2}$ . The magnitudes of these values are comparable to the local thermal conductivities along the polypeptide chain. Interestingly, the residue pairs Ser3-Asp6 and Asp4-Arg15 were also identified as having relatively high local energy diffusivity in a theoretical study using non-equilibrium MD simulations and a master equation model.<sup>9</sup> This suggests their potential roles as critical “shortcuts” in the thermal transport network of the HP36 protein.

#### 4. Reference

- (1) Yu, X.; Leitner, D. M. Heat Flow in Proteins: Computation of Thermal Transport Coefficients. *J. Chem. Phys.* **2005**, *122* (5), 54902.
- (2) Wang, T.; Yamato, T.; Sugiura, W. Site-Selective Heat Current Analysis of  $\alpha$ -Helical Protein with Linear-Homopolymer-like Model. *J. Chem. Phys.* **2023**, *158* (21), 214105.
- (3) Hayashi, Y.; Shiomi, J.; Morikawa, J.; Yoshida, R. RadonPy: Automated Physical Property Calculation Using All-Atom Classical Molecular Dynamics Simulations for Polymer Informatics. *Npj Comput. Mater.* **2022**, *8* (1), 1-15.
- (4) Ohara, T.; Yuan, T. C.; Torii, D.; Kikugawa, G.; Kosugi, N. Heat Conduction in Chain Polymer Liquids: Molecular Dynamics Study on the Contributions of Inter- and Intramolecular Energy Transfer. *J. Chem. Phys.* **2011**, *135* (3), 034507. <https://doi.org/10.1063/1.3613648>.
- (5) Kondoh, M.; Mizuno, M.; Mizutani, Y. Importance of Atomic Contacts in Vibrational Energy Flow in Proteins. *J. Phys. Chem. Lett.* **2016**, *7* (11), 1950-1954.
- (6) Reid, K. M.; Yamato, T.; Leitner, D. M. Scaling of Rates of Vibrational Energy Transfer in Proteins with Equilibrium Dynamics and Entropy. *J. Phys. Chem. B* **2018**, *122* (40), 9331-9339.
- (7) Deniz, E.; Valiño-Borau, L.; Löffler, J. G.; Eberl, K. B.; Gulzar, A.; Wolf, S.; Durkin, P. M.; Kaml, R.; Budisa, N.; Stock, G.; et al. Through Bonds or Contacts? Mapping

Protein Vibrational Energy Transfer Using Non-Canonical Amino Acids. *Nat. Commun.* **2021**, *12* (1), 3284.

(8) Mizutani, Y.; Mizuno, M. Time-Resolved Spectroscopic Mapping of Vibrational Energy Flow in Proteins: Understanding Thermal Diffusion at the Nanoscale. *J. Chem. Phys.* **2022**, *157* (24), 240901.

(9) Buchenberg, S.; Leitner, D. M.; Stock, G. Scaling Rules for Vibrational Energy Transport in Globular Proteins. *J. Phys. Chem. Lett.* **2016**, *7* (1), 25-30.