

Biophysical Journal, Volume 122

Supplemental information

**Phonon-assisted electron-proton transfer in [FeFe] hydrogenases:
Topological role of clusters**

Yann Chalopin, Stephen P. Cramer, and Simon Arragain

Supplementary Informations : Phonon-assisted electron-proton transfer in [FeFe] hydrogenases: Topological role of clusters

Yann Chalopin, Stephen P. Cramer, and Simon Arragain
(Dated: February 27, 2023)

INTRODUCTION

MATERIALS AND METHOD

Construction of the dynamical matrix

To estimate harmonic interactions between residues, we implemented and compared two very distinct force fields, namely an elastic network model (ENM), as well as an all-atom force field through molecular dynamics simulations. The elastic model employed here has no cut-off length. For the MD simulations, we employed the Amber force field to take into account interactions between all the atoms of the enzyme. We have also simulated the interactions with the FeS cluster and H-cluster according to the force field proposed by Kim. The solvent is accounted explicitly for the water atoms with the TIP3P potential, and by centering the system in a box such that the distance between an atom of the complex and the edge of the box is 10Å. The equilibrium trajectories of the system were obtained after minimizing the cohesion energy and an equilibration time of 0.1 ns in order to avoid too large conformational changes.

details

The constants α_{ij} of the system are obtained by tracking the equilibrium displacement fluctuations of the C_α . By introducing the correlation matrix \mathbf{C} of displacements

$$\mathbf{C} : (\mathbf{C})_{ij} = \beta \langle X_i X_j \rangle, \quad (1)$$

with $\beta = 1/K_b T$. The force constant matrix A is obtained by inversion

$$\mathbf{A} : \alpha_{ij} = (\mathbf{C}^{-1})_{ij} \quad (2)$$

$$\mathbf{A} : \alpha_{ij} = \begin{cases} (\mathbf{C}^{-1})_{ij} & i \neq j \\ -\sum_{i=j} \alpha_{ij} & i=j \end{cases} \quad (3)$$

However, the correlation matrix \mathbf{C} contains noise which produces inaccurate matrix inversion by conventional methods. Instead it is decomposed in singular values as

$$\mathbf{C} = U S V^T, \quad (4)$$

which allows to obtain the inverse matrix from

$$\mathbf{C}^{-1} = V S^{-1} U^T. \quad (5)$$

The correlation matrix was obtained and refined by ensemble averaging 20 distinct equilibrated trajectories at room temperature ($T = 300K$). Figure 3-A(top) compares the location hotspots obtained from MD and ENM calculations. The correlation between the two quantities reaches 74%.

It can be seen that for a total of 574 residues, about 15 residues (< 3%) are thermal spots. As derived in a previous study, these domains populated with a high phonon density of states correspond to distinct, uncoupled clusters of residues (Figure 3-B). To further illustrate this assumption, we have projected this quantity onto the three-dimensional structure of the hydrogenase (Figure 3-C and Figure 3-D). The localization landscape reveals that the rigid, central region of the enzyme undergoes complex dynamic patterns, comprising small domains that form an irregular spatial structure. As discussed later, this distribution of thermal spots is crucial in the working of H₂ases.

RESULTS

The topological role of the clusters

We provide an additional figure for this section

TUTORIAL : HOW TO COMPUTE THE LOCALIZATION LANDSCAPE

In this section, we will look at step by step how to calculate the localization landscape in detail. To do this, we will consider the simplest folded protein consisting of three residues. This simplicity allows us to present all the calculation steps that can be done by hand or analytically by the reader.

We start by expressing the dynamic matrix $L = -\Delta$ corresponding to the discrete Laplacian associated with the topology of the figure 5-B.

$$L = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}$$

This matrix can be seen as the sum of a degree matrix minus the adjacency matrix $L = D - A$. Now we express the matrix L_h obtained from Equation 1 describing the dynamics of the system in the harmonic regime defined by the wave vector $k = \pi/a = \pi$. So:

$$Lh = c - \Delta_h \begin{bmatrix} c-1 & 1 & 0 & 0 \\ 1 & c-3 & 1 & -1 \\ 0 & 1 & c-2 & 1 \\ 0 & -1 & 1 & c-2 \end{bmatrix}$$

We choose c to be greater than the largest eigenvalue of the system (here we choose for example $c = 6$).

$$Lh = c - \Delta_h \begin{bmatrix} 5 & 1 & 0 & 0 \\ 1 & 3 & 1 & -1 \\ 0 & 1 & 4 & 1 \\ 0 & -1 & 1 & 4 \end{bmatrix}$$

The solution of the linear system defined by Equation (3)

$$L_h u_h = (1) \tag{6}$$

gives us $u_h = [0.125, 0.375, 0.075, 0.325]$. This localization landscape is plotted on Figure 5-C and represented in space on Figure 5-D with an identical color code to the one representing the results of the article. In the first case (C), the landscape is plotted as a function of the residue index encountered in the PDB in order to avoid discontinuity in the case where the sequence contains gaps.

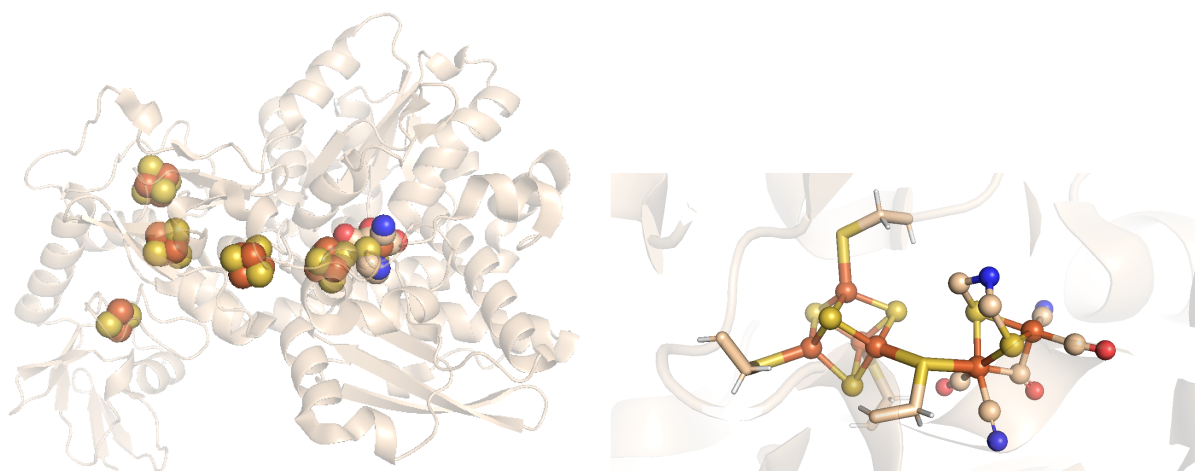


FIG. 1: **Structure** [FeFe] hydrogenase (CpI) with all the FeS clusters and the details of the H-Cluster with carbon in orange, oxygen in red, sulfur in yellow, nitrogen in blue et iron deep red (PDBid: 4xdc).

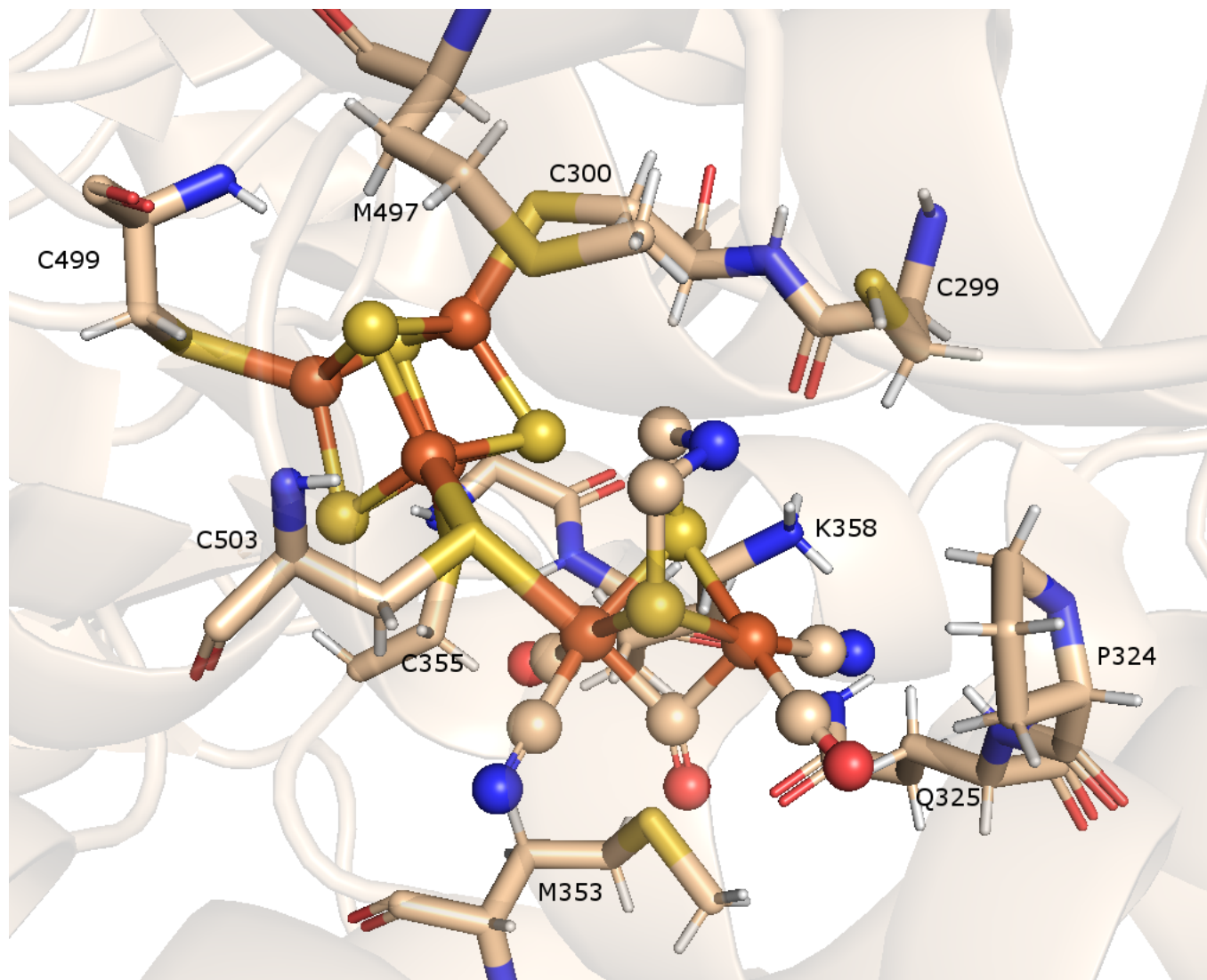


FIG. 2: **Structure** CpI amino acid around 4\AA of the subcluster of the H-Cluster (Top part). The amino acids inCpI M353, C299, C298, M497, K358,P324, Q325. CpI additional residues around the subcluster might also be important (S323, I268, A230, P231, S232, P354,F417,

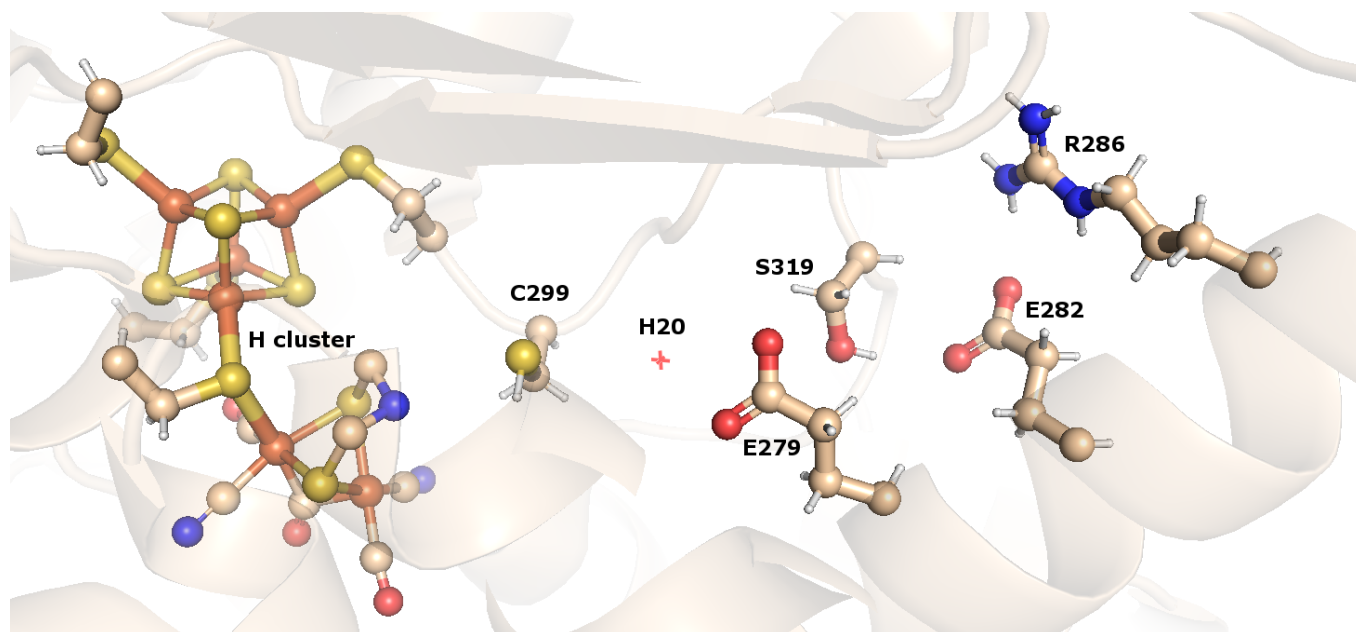


FIG. 3: **Structure** Proton transport chains in CpI with in the pathway 2, C299, S298, some H₂O molecules [Ginovska-Pangovska 2014].

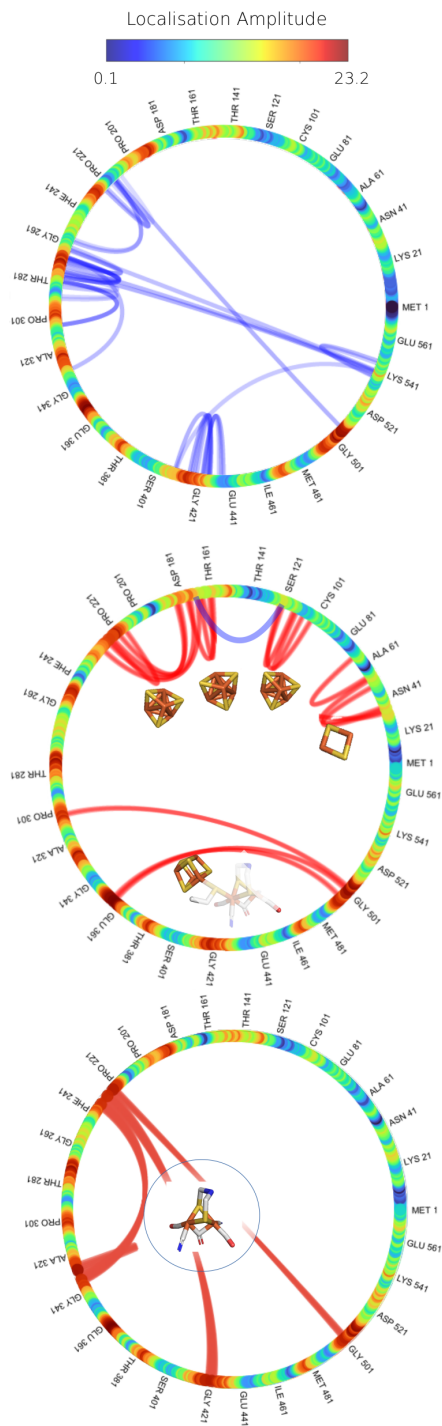


FIG. 4: **Thermal hot spots and distal contacts** top: Main interactions between non consecutive residues in the sequence. This indicates that specific hot-spots are dynamically coupled by distal contacts. middle: The FeS clusters allows to couple most of the thermal hot-spots of the structure. They play the role of cofactor-contacts. bottom: The H-cluster is at the center of a cavity made of hot-spot. The pocket is associated to network of connected residues that are the main hot-spots. Units are in ω^{-2}

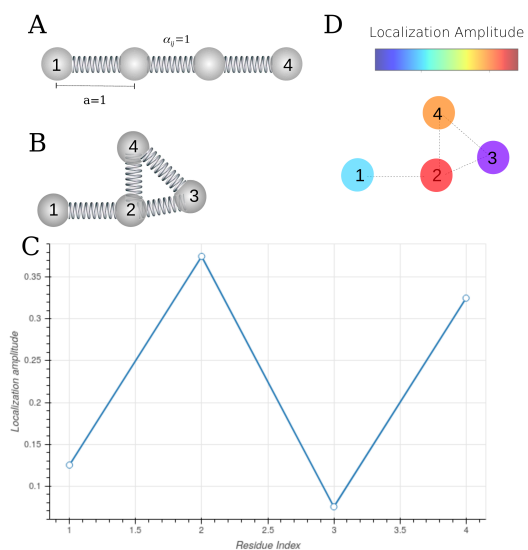


FIG. 5: **Computing the localization landscape top:**
 A - Simple homogeneous non folded model of protein.
 B. Simple folding of a 4 residues protein. C.
 Corresponding localization lanscape. D. Color coded
 spatial representation of the localization landscape.