**A Hidden Aggregation-Prone Structure in the Heart of Hypoxia Inducible Factor Prolyl Hydroxylase**

**Short Title:** aggregation-prone state of PHD2 unfolding

*Hamid Hadi-Alijanvand [1,2], Elizabeth A. Proctor [3], Feng Ding [4,5], Nikolay V. Dokholyan [4,6,7], Ali A. Moosavi-Movahedi [2,8*]*

[1] Department of Biological Sciences, Institute for Advanced Studies in Basic sciences (IASBS), Zanjan, Iran

[2] Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran

[3] Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[4] Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, School of Medicine, Chapel Hill, NC 27599, USA

[5] Department of Physics and Astronomy, Clemson University, Clemson, SC 29634, USA

[6] Curriculum in Bioinformatics and Computational Biology, University of North Carolina at Chapel Hill, School of Medicine, Chapel Hill, NC 27599, USA

[7] Program in Molecular and Cellular Biophysics, University of North Carolina at Chapel Hill, School of Medicine, Chapel Hill, NC 27599, USA

[8] Center of Excellence in Biothermodynamics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran

* Corresponding Author;

E-mail: moosavi@ut.ac.ir

Tel: +9821-66403957

Fax: +9821-66404680

Address: Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran

**Supporting Information**

We compute the accessible surface area (ASA) of tryptophan residues for the unfolding intermediates (Fig. S1). The ASA changes of tryptophan residues 148 and 203, located at the edge of and within the active site lumen, respectively, indicate that the unfolding of this region begins at the end of the lumen corruption. This exposure occurs in two steps, which correlate with the structural species existing at temperatures 280 K and 318 K (Fig. 6, Intermediate states 1 and 5).

The tryptophan residue 203 remains buried until 318 K (Intermediate state 5). The exposure of tryptophan residue 181 increases in three steps, at 280, 301, and 348 K (Intermediate states 1, 3, and 6, respectively). A comparison of Intermediate state temperatures in the 2D-PMF-w with the temperatures of important structural changes in the 2D-PMF-g reveals that the first two steps of tryptophan 181 exposures occur simultaneously with the unfolding of helices, with the third step occurring in parallel with the unfolding of the active site lumen. W72 is located near the active site lumen entrance and just at the end of the first strand of a beta-sheet that is conserved until 331 K (based on 2D-PMF-g data). The ASA of this residue does not change until 348 K (Intermediate state 5). However, W72 is largely exposed at higher temperatures, which corresponds to the unfolding of the active site lumen and the most stable beta-sheet structure. We observe that tryptophan residues have varying accessibilities among the species of the last 2D-PMF-w intermediate state (well number 6). This variation aids in interpretation of the fluorescence thermal unfolding profile; in the first transition, tryptophan residues 148 and 181 become exposed, and during the last fluorescence transition tryptophan residues 72 and 203 become exposed.

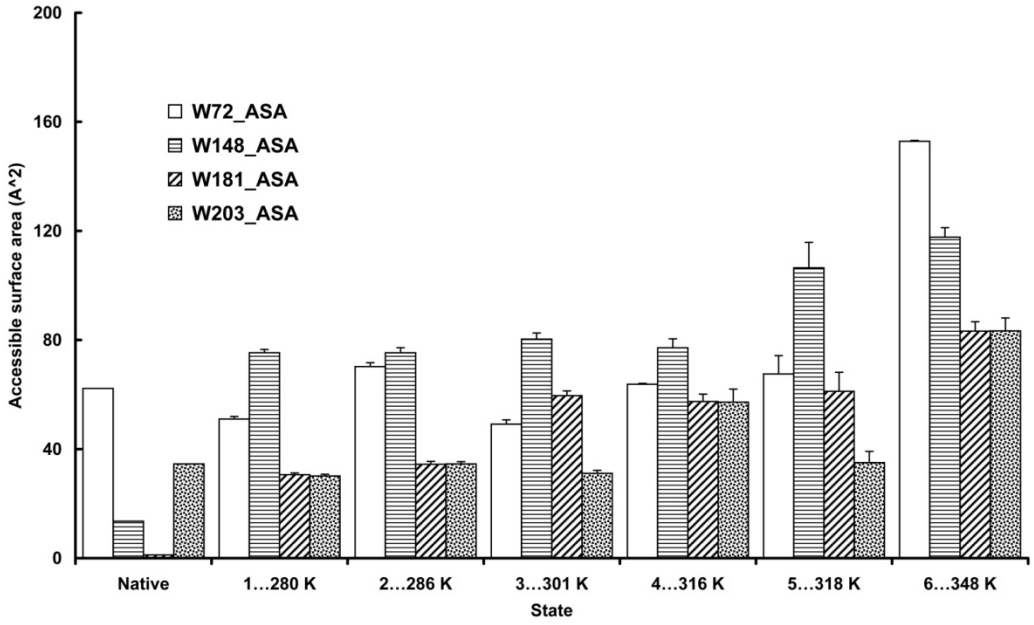


**Figure S1.** Different tryptophan residues of PHD2 become accessible to solvent step by step during the progression of unfolding process. Intermediate state number and temperature is indicated on x-axis (Intermediate state number correspond to Figure 6 in main text) , with native state information derived from the crystal structure. Error bars represent the standard error of mean.

**Figure S2.** Animated GIF file shows the order of PHD2 structure's destruction defined using results of DMD.

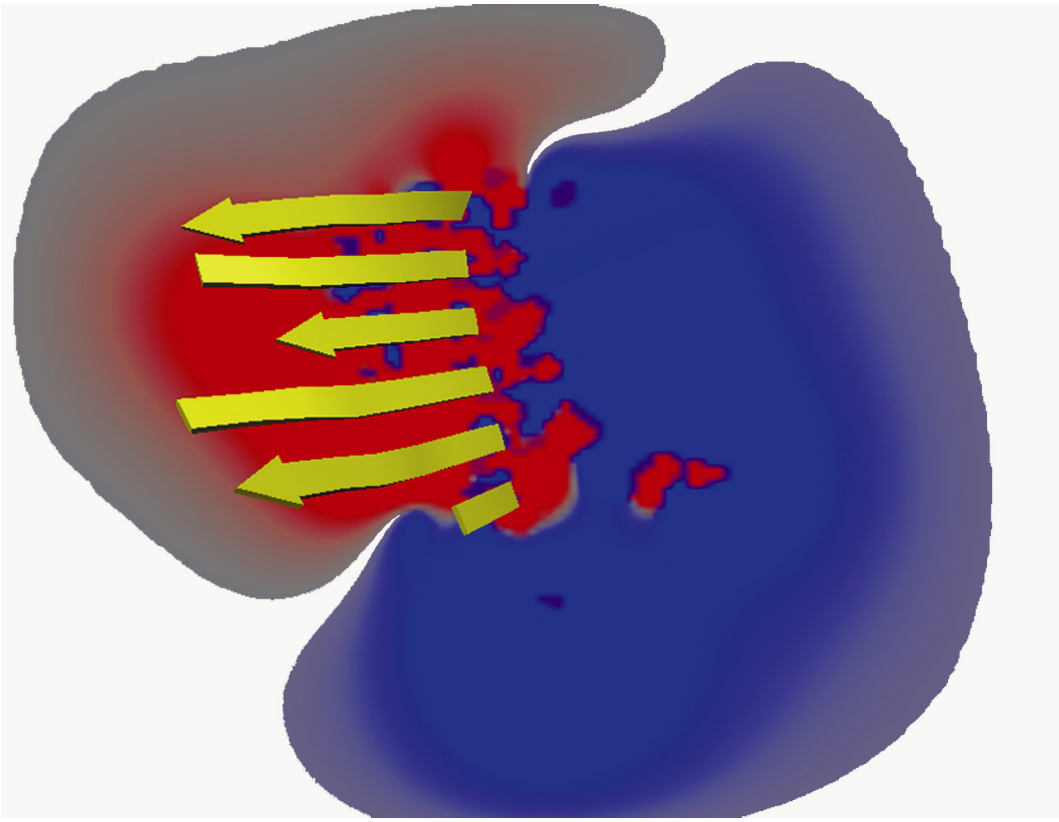

**Figure S3.** The electrostatic potential map of aggregation-prone intermediate is computed to define possible role of electrostatic interaction in assembly of aggregation-prone intermediate of PHD2 unfolding. The structure of intermediate state is represented in cartoon. Positive and negative electrostatic potential regions defined by blue and red regions, respectively.