

# **SUPPLEMENTAL MATERIAL**

## **Conformational Flexibility of the C-Terminal Region Influences Distal Active Site Residues Across the Tautomerase Superfamily**

Christopher Argueta<sup>†, ,</sup>, Andrew Parkins<sup>†</sup>, and Georgios Pantouris<sup>\*</sup>

*Department of Chemistry, University of the Pacific, Stockton, CA 95211, USA*

<sup>†</sup>Equal contributing authors

**\*Corresponding author:** gpantouris@pacific.edu

**Table S1: Sequence identity between TSF members\***

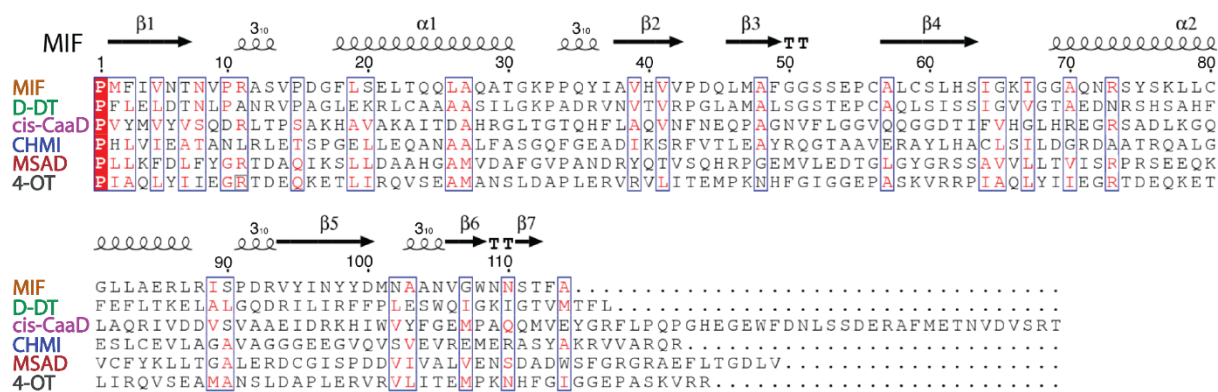
	<b>MIF</b>	<b>D-DT</b>	<b>cis-CaaD</b>	<b>CHMI</b>	<b>MSAD</b>	<b>4-OT</b>
<b>MIF</b>	-	34.2	16.5	15.7	19.1	14.6
<b>D-DT</b>	34.2	-	12.8	21.9	19.5	16.4
<b>cis-CaaD</b>	16.5	12.8	-	22.9	25.0	14.3
<b>CHMI</b>	15.7	21.9	22.9	-	24.7	16.0
<b>MSAD</b>	19.	19.5	25.0	24.7	-	25.8
<b>4-OT</b>	14.6	16.4	14.3	16.0	25.8	-

\*The numbers are expressed in percent (%) identity

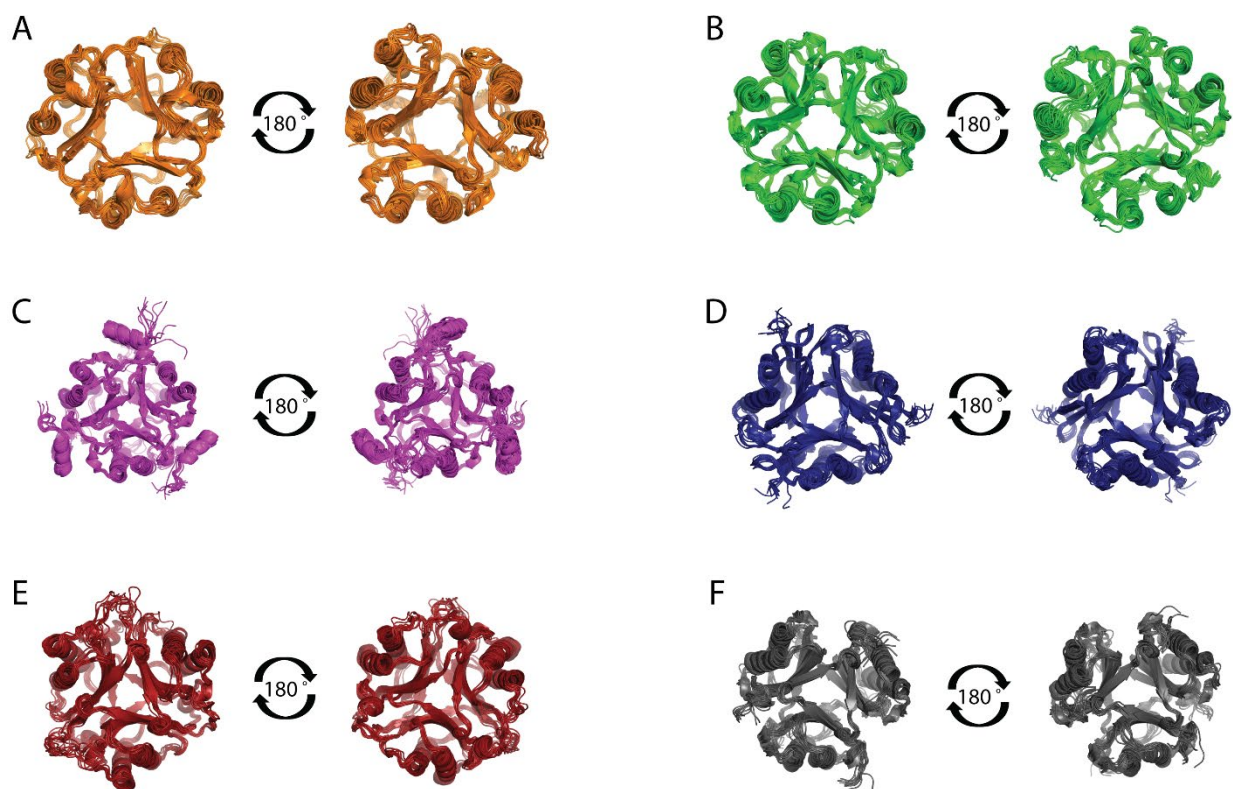
**Table S2: Average Root Mean Square fluctuation (RMSF) profiles of TSF proteins<sup>\*</sup>**

	Average RMSF (Å)
<b>MIF</b>	$0.71 \pm 0.2$
<b>D-DT</b>	$0.69 \pm 0.3$
<b>cis-CaaD</b>	$0.88 \pm 0.7$
<b>CHMI</b>	$0.67 \pm 0.5$
<b>MSAD</b>	$0.79 \pm 0.5$
<b>4-OT</b>	$0.87 \pm 0.7$

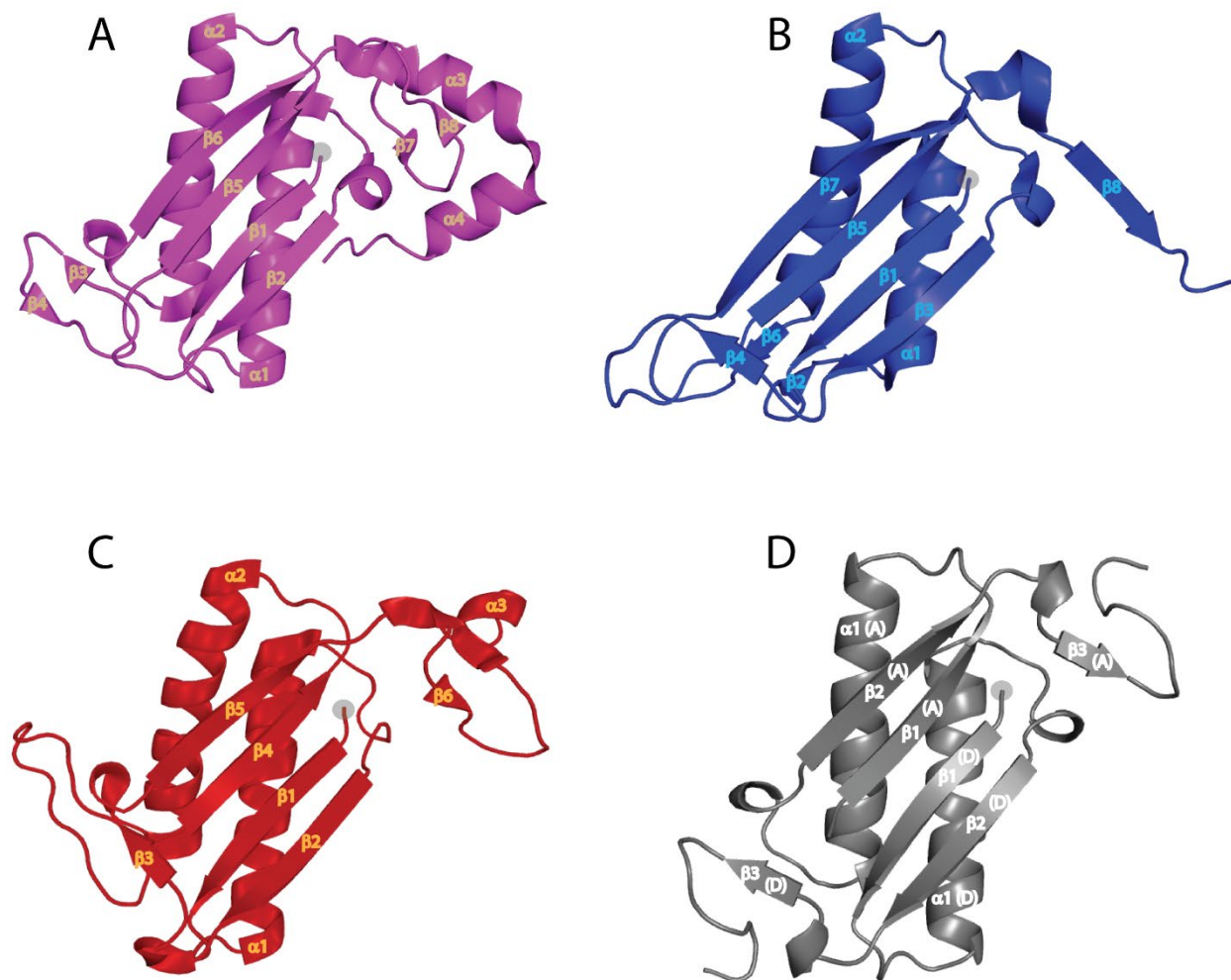
<sup>\*</sup>Data were obtained from three independent 1 $\mu$ s MD simulations



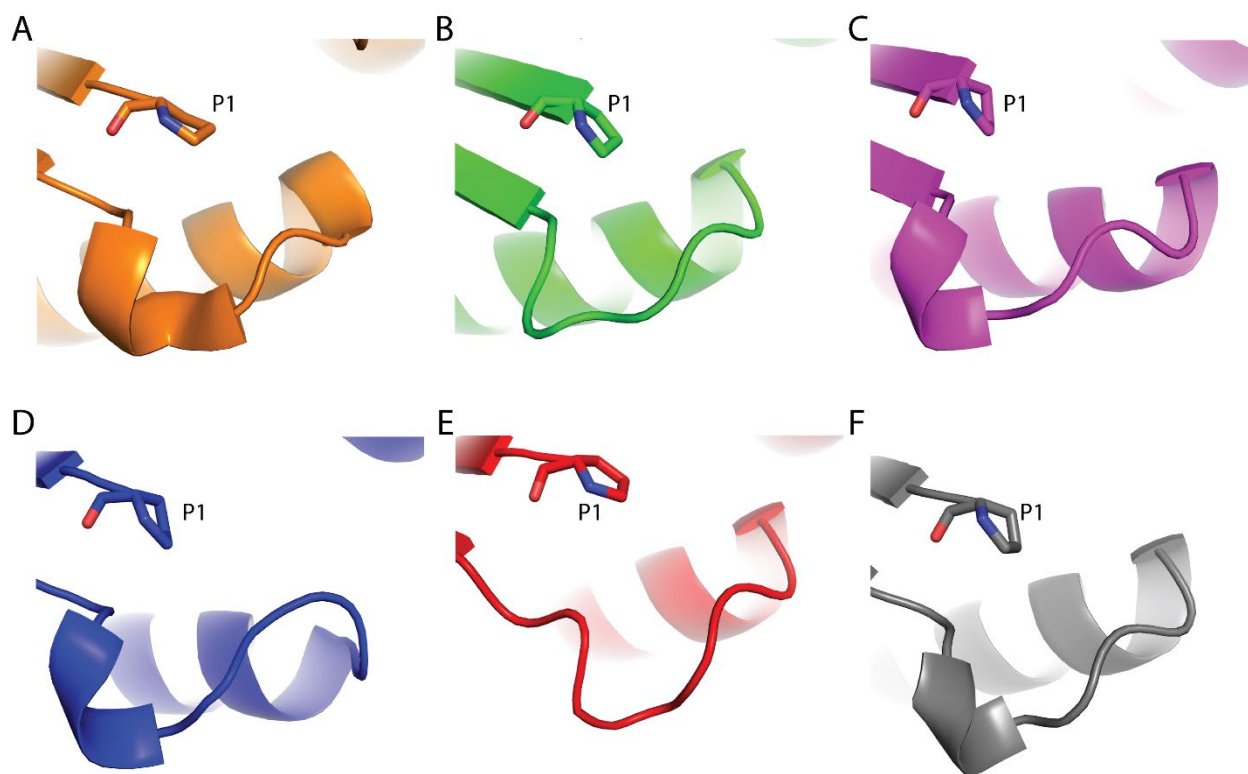
**Figure S1. Multiple sequence alignment of TSF representative members.** The amino acid sequences of human MIF (PDB:3DJH), human D-DT (PDB:1DPT), *Corynebacterium* cis-CaaD (PDB:2FLZ), *Pseudomonas aeruginosa* CHMI (PDB:3E6Q), *Pseudomonas pavonaceae* MSAD (PDB:2AAG), and *Pseudomonas* sp. CF600 4-OT (PDB:1OTF) were obtained from RCSB PDB (protein data bank) using the corresponding entries shown in each parenthesis. A red box and a white character (catalytic residue P1) designates strict identity. The red characters designate similarity in a group. The blue frame designates similarity across groups. For comparison, the secondary structure of MIF is provided on the top of the panel. Multiple sequence alignment was performed using Clustal Omega [1].



**Figure S2. Dynamic profiles of the six TSF proteins.** Snapshots of A) MIF, B) D-DT, C) cis-CaaD, D) CHMI, E) MSAD, and F) 4-OT structures were captured at 100 ns increments, having 0 ns and 1  $\mu$ s as the starting and ending points, respectively. For accuracy, the one microsecond trajectories were repeated in triplicate.



**Figure S3. The secondary structure features of the four bacterial proteins.** The **(A)** cis-CaaD (PDB: 2FLZ), **(B)** CHMI (PDB: 3E6Q), **(C)** MSAD (PDB: 2AAG) monomers and **(D)** 4-OT (PDB: 1OTF) dimer were used for this analysis. The position of P1 is highlighted with a transparent sphere.



**Figure S4. Illustration of the flexible loop, found in the active site environment of the six TSF proteins.** This  $\alpha 1/\beta 2$  loop (the corresponding  $\alpha 1/\beta 3$  loop for CHMI) is located next to the catalytic residue P1. The six illustrations were derived from the crystal structures of **A)** MIF (PDB:3DJH), **B)** D-DT (PDB:1DPT), **C)** cis-CaaD (PDB:2FLZ), **D)** CHMI (PDB:3E6Q), **E)** MSAD (PDB:2AAG), and **F)** 4-OT (PDB:1OTF). The catalytic residue P1 is shown as sticks.

## REFERENCES

1. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Soding, J.; Thompson, J. D.; Higgins, D. G., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **2011**, 7, 539.