

Supplementary Material for

## **Conserved quantitative stability/flexibility relationships (QSFR) in an orthologous RNase H pair**

D.R. Livesay<sup>1</sup> and D.J. Jacobs<sup>2</sup>

<sup>1</sup>Department of Chemistry and Center for Macromolecular Modeling & Materials Design,  
California State Polytechnic University, Pomona

<sup>2</sup>Department of Physics and Astronomy, California State University, Northridge

### **Table of contents:**

Supp. Fig. 1: Comparison of best fit u,v pairs vs.  $\delta_{\text{nat}}$

Supp. Fig. 2: Model predictions vs. fitting error.

Supp. Fig. 3: Correlations within the three phenomenological parameters.

Supp. Fig. 4: Comparison of global flexibility at best fit u,v pairs along  $\delta_{\text{nat}}$ .

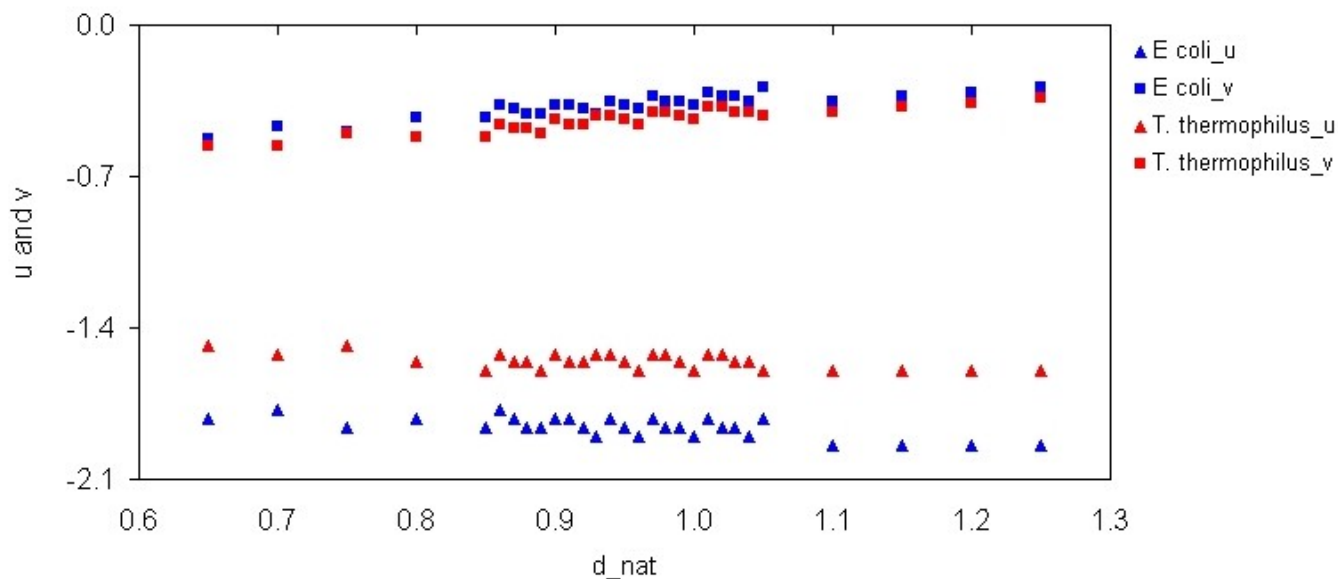
Supp. Fig. 5: Comparison of exemplar cooperativity correlation plots vis-à-vis parameterization

Supp. Fig. 6: Comparison of backbone flexibility predictions vis-à-vis parameterization.

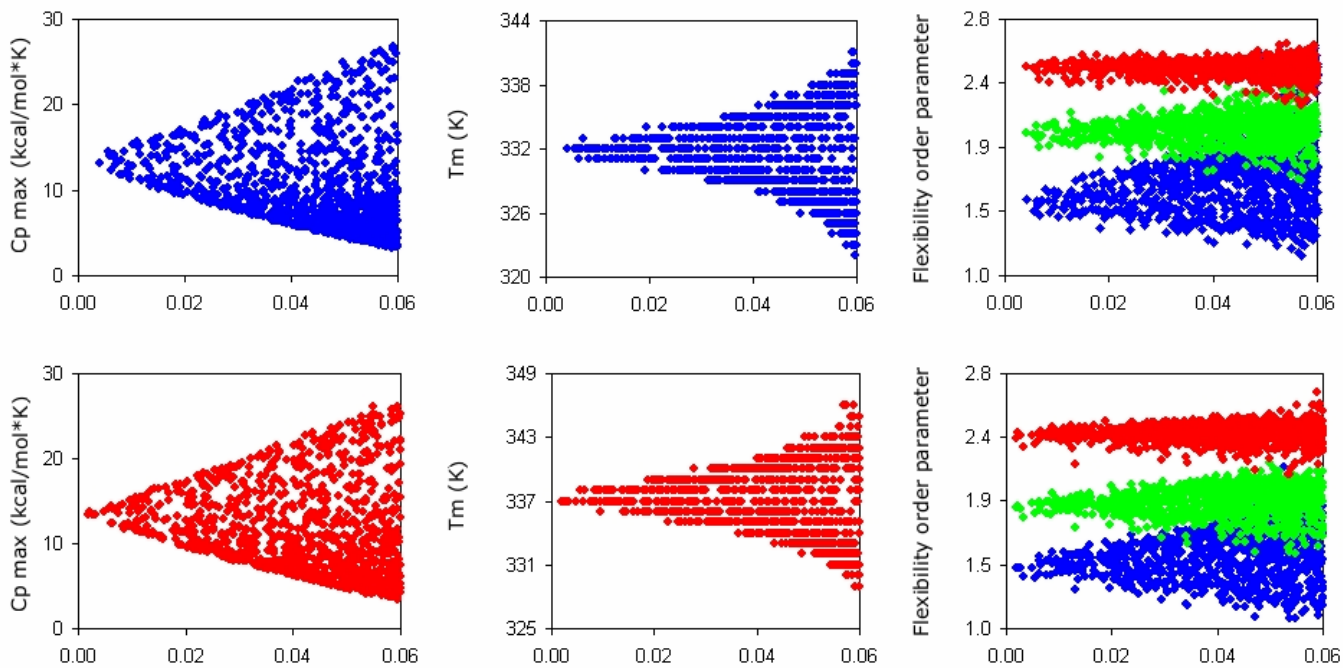
Supp. Fig. 7: Variation within cooperativity correlation plots.

Supp. Table 1: Exemplar parameter sets.

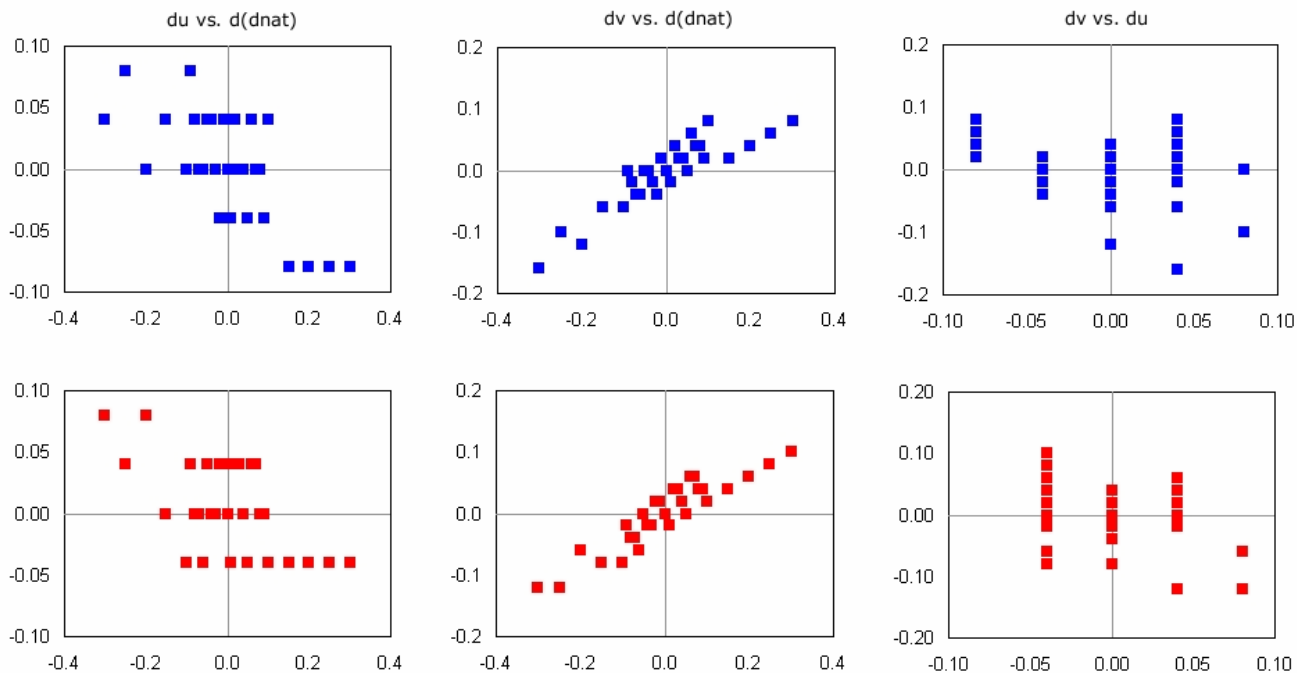
Supp. Table 2: Correlations of parameters and other model predictions.



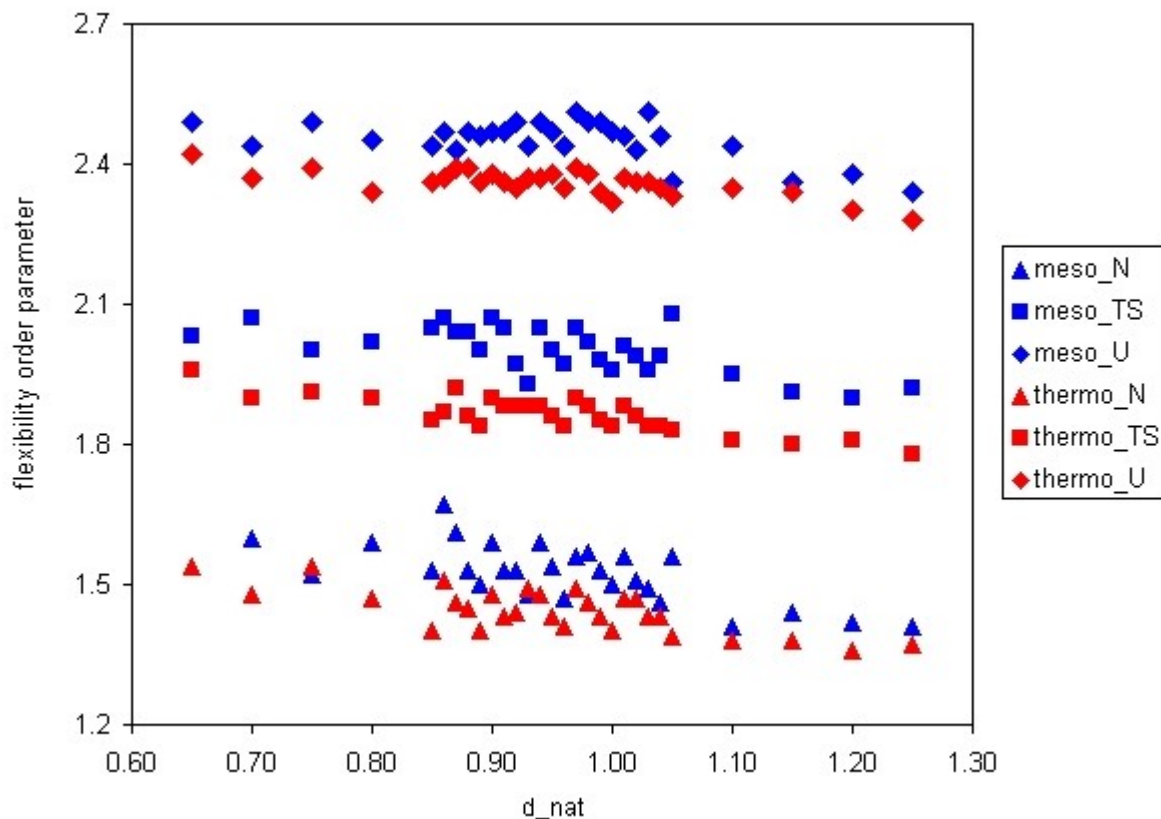
**Supp. Fig. 1.** Best fit  $u, v$  pairs are plotted against a coarse-grained  $\delta_{\text{nat}}$  range. In all instances, the best fit  $u, v$  pairs indicate that an additional cohesive force (interpreted as hydrophobic in nature) is present within the thermophilic ortholog. The robustness of the observed cohesive force strongly substantiates the conclusions made herein regarding the improved hydrophobic packing of the thermophilic ortholog.



**Supp. Fig. 2.**  $C_p$  maximum,  $T_m$ , and  $\theta$  at key points along the unfolding pathway are plotted against the global fitting error. The error indicates the least squares fitting error per residue normalized by peak height, which is the same as in Livesay et al., *FEBS Letters*, **576**, 2004. *E. coli* results are on the top, whereas *T. thermophilus* results are on the bottom. In the  $\theta$  plot, blue indicates  $\theta_{\text{nat}}$ , green indicates  $\theta_{\text{TS}}$ , and red indicates  $\theta_{\text{unf}}$ . The colors in the other four plots have no meaning other than blue is for *E. coli* and red is for *T. thermophilus*.

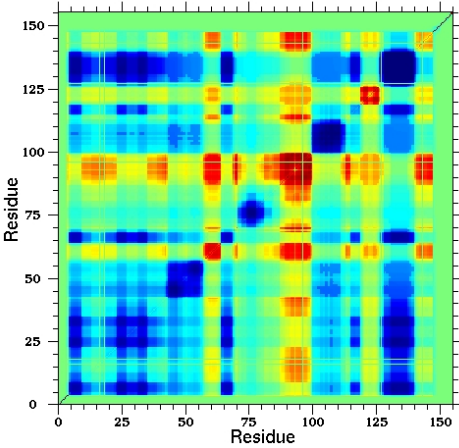
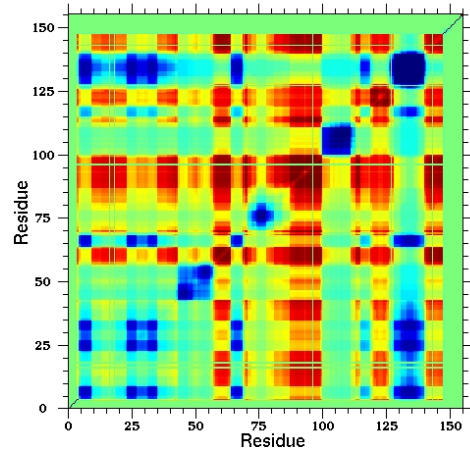
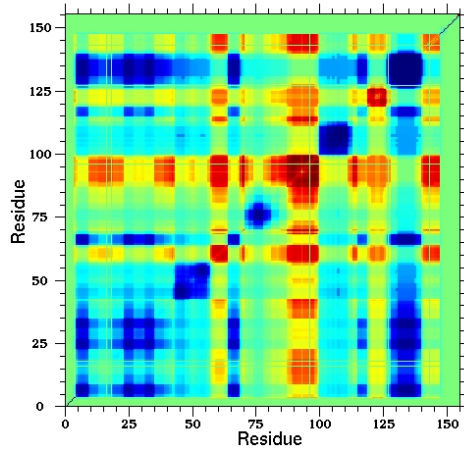
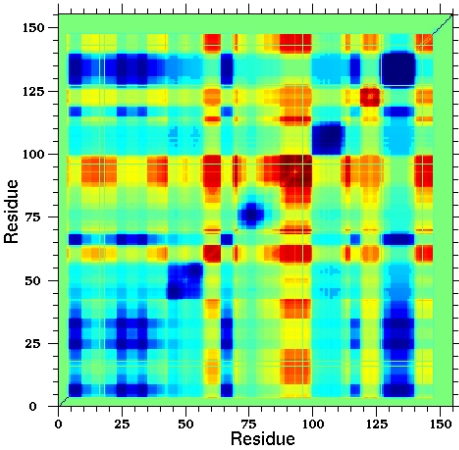
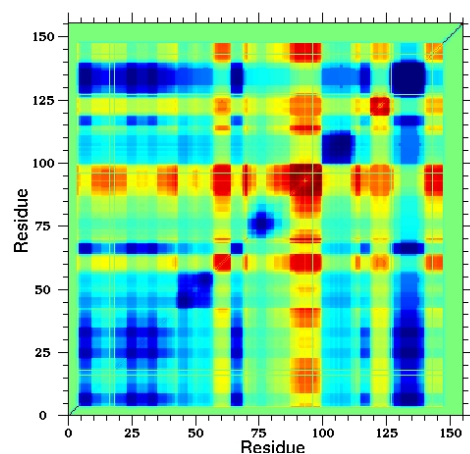
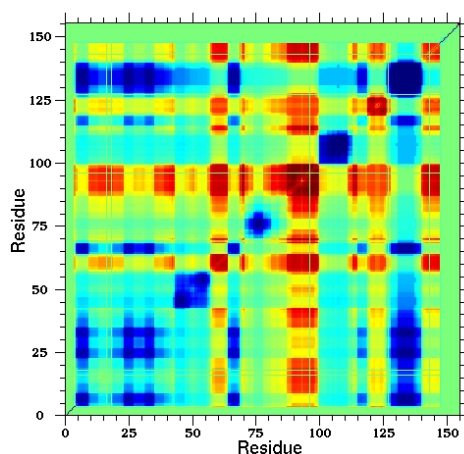
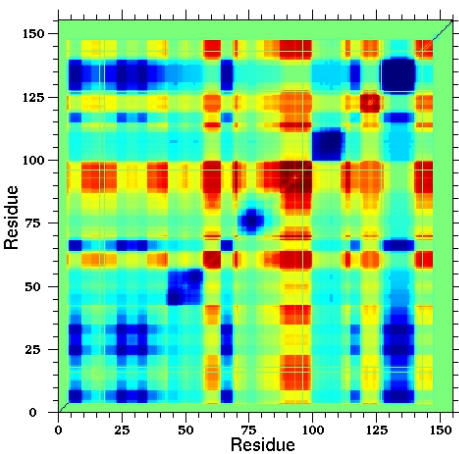
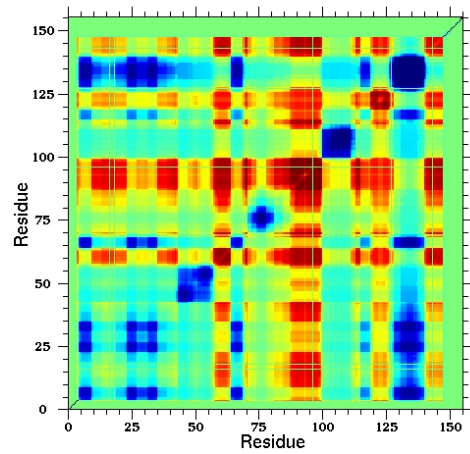
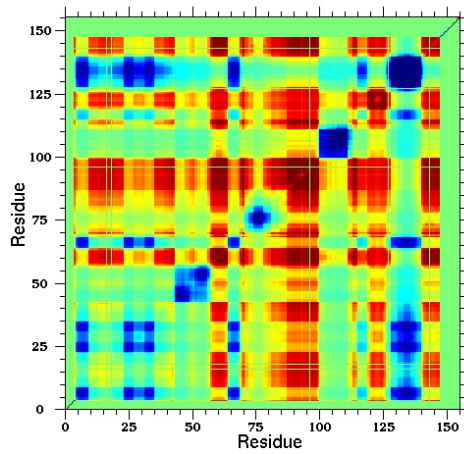
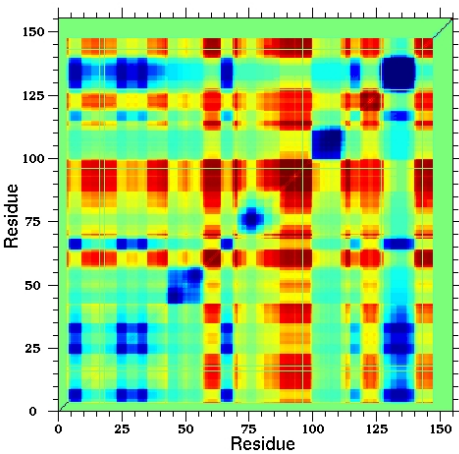


**Supp. Fig. 3.** Parameter deviations (from the values reported in Table 1) of the best  $u, v$  pairs for a given  $\delta_{\text{nat}}$  are plotted against each other to demonstrate how the three phenomenological parameters work in concert. *E. coli* results are plotted in blue, whereas *T. thermophilus* results are in red. In both cases, the correlation is strongest within the  $v, \delta_{\text{nat}}$  pair, demonstrating how the entropic and enthalpic costs of native torsion angles are coupled. The correlation is next strongest in the  $u, \delta_{\text{nat}}$  pair, and weakest in the  $u, v$  pair. Linear correlation coefficients are reported in Supp. Table 2. The  $u, v$  parameter pairs shown here are the same as in Supp. Fig. 1.



**Supp. Fig. 4.** Global flexibility order parameter,  $\theta$ , values at the Native and Unfolded minima and the straddling Transition State maximum. Values reported are for best  $u, v$  pair as a function of  $\delta_{\text{nat}}$  for the *E. coli* (blue) and *T. thermophilus* (red) structures, each at their respective  $T_m$ . This plot shows how global flexibility adjusts similarly within the mesophilic and thermophilic parameter sets. As a consequence, conservation within the global flexibility is maintained.

**Supp. Fig. 5 (next page).** Cooperativity correlation plots (at  $T_m$ ) of nine good exemplar *E. coli* parameter sets, the best fit parameter set, and an exemplar “bad” fit set. The good exemplar parameter sets are the same as those in Fig. 10. In all instances, the main flexibly correlated features are conserved. The best fit cooperativity correlation plot qualitatively appears to be the average of the nine good fits. This point is demonstrated quantitatively in Supp. Fig. 7. As discussed in the text, the main characteristics of the flexibility information is maintained even in the bad fit. However, the magnitude of the flexibly/rigidly correlated distinctions is not as great as the good fits. Similar results are observed for the *T. thermophilus* ortholog (unpublished results).

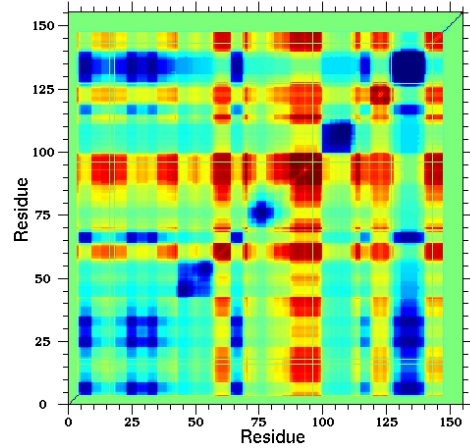


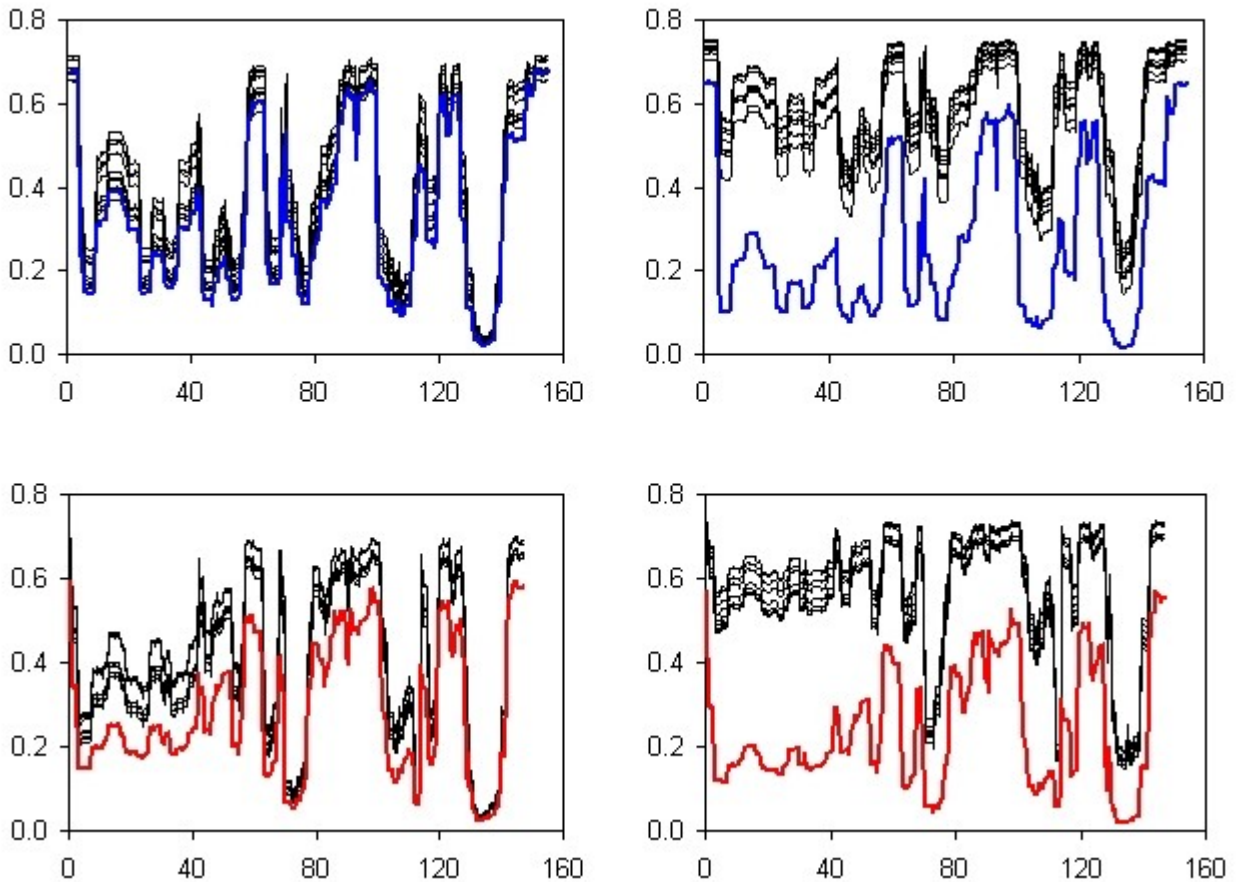
**Nine plots above** – randomly picked exemplar good fits over a coarse-grained  $d_{\text{nat}}$  range.

**Left** – an exemplar bad fit.

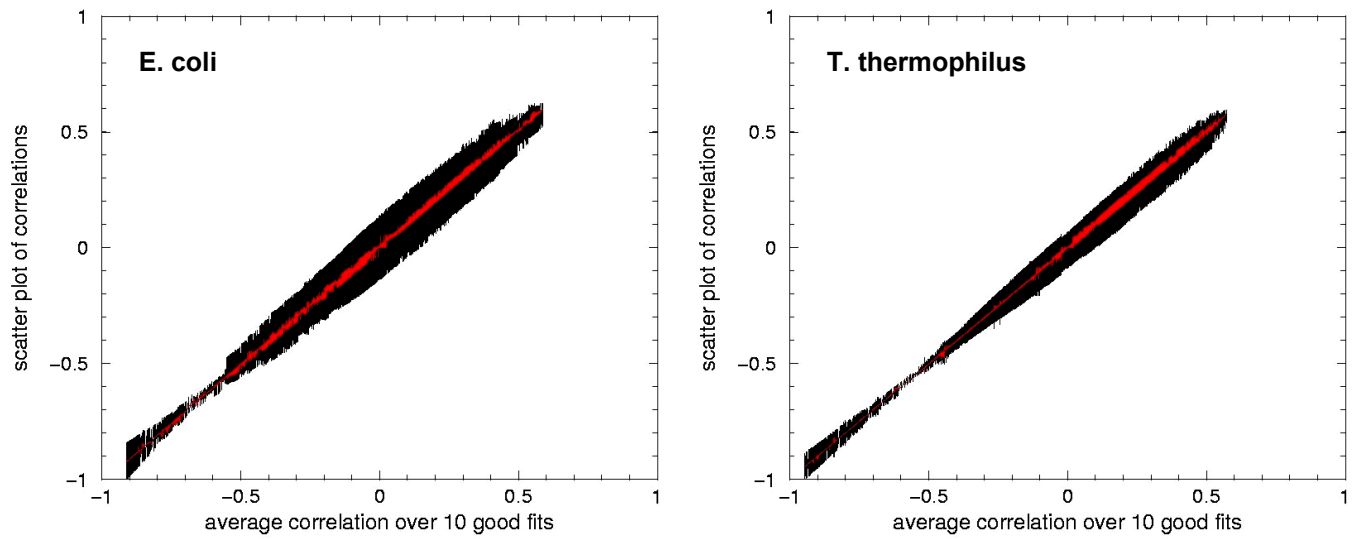
**Right** – The best-fit identified.

(Figure caption is located on preceding page.)





**Supp. Fig. 6.**  $P_R$  versus residue number of the nine mesophilic and nine thermophilic exemplar parameter sets shown in Fig. 10 (solid black lines). In all cases, the temperature is set to their respective  $T_m$  (which can deviate up to 3 K from the actual  $T_m$ ), and all have an error  $\leq 0.03$ . A single parameter set whose error is  $\sim 0.08$  (colored line) is also provided for comparison. *E. coli* results are shown on the top and *T. thermophilus* results are underneath. The native structure sub-ensemble is on the left, whereas the full thermodynamic ensemble is on the right. The differences within the sub-ensemble flexibility predictions are much less than those within the full thermodynamic ensemble. While  $C_p$ ,  $T_m$ , and  $\theta_{\text{nat,TS,unf}}$  values may vary significantly at such high errors, the flexibility information is qualitatively maintained, especially within the sub-ensemble averaging. The exact values of the parameter sets are provided in Supp. Table 1.



**Supp. Fig. 7.** Variability within cooperativity correlation plots (Supp. Fig. 5) the nine exemplar “good” fits (black) around the best-fit (red) is demonstrated. Within the cooperativity correlation plots, a single number describes the extent of flexibility/rigidity correlation within the ensemble. The number ranging from -1 (rigidly correlated) to +1 (flexibly correlated). Plotting the scatter of the ten values against their average reveals two important points. First, and most important, that the variation about the average is limited (never more than ~10%). Second, that the best-fit parameter set is roughly the average of the ten datasets.



E. coli						
	d_nat	u	v	Tm	Cp_max	Error
best	0.95	-1.86	-0.37	332	13.0	n/a
good	0.65	-1.82	-0.53	329	15.1	0.01
good	0.75	-1.82	-0.47	333	13.8	0.02
good	0.85	-1.82	-0.39	329	11.3	0.03
good	0.90	-1.90	-0.43	329	16.3	0.02
good	0.95	-1.82	-0.35	332	11.9	0.02
good	1.00	-1.94	-0.39	331	15.4	0.02
good	1.05	-1.82	-0.29	332	11.3	0.01
good	1.15	-1.82	-0.25	329	10.7	0.02
good	1.25	-2.02	-0.35	329	16.4	0.02
bad	0.65	-1.66	-0.55	351	10.6	0.08
T. thermophilus						
	d_nat	u	v	Tm	Cp_max	Error
best	0.95	-1.56	-0.44	337	15.3	n/a
good	0.65	-1.48	-0.56	337	10.2	0.03
good	0.75	-1.60	-0.58	337	16.5	0.01
good	0.85	-1.60	-0.52	337	15.4	0.01
good	0.90	-1.48	-0.42	339	11.5	0.02
good	0.95	-1.52	-0.40	335	11.4	0.03
good	1.00	-1.52	-0.38	337	11.6	0.02
good	1.05	-1.60	-0.42	337	14.4	0.01
good	1.15	-1.64	-0.40	337	15.0	0.02
good	1.25	-1.64	-0.38	337	15.3	0.03
bad	0.65	-1.68	-0.60	352	1.0	0.08

**Supp. Table 1.** Parameter values of the nine exemplar fits with error  $\leq 0.03$  and an additional parameter set with an error  $\sim 0.08$ . These parameter sets are used in Fig. 10 and Supp. Fig. 5-7.

		<b>E. coli</b>				
	<b>dnat</b>	<b>du</b>	<b>dv</b>	<b>error</b>	<b>Tm</b>	<b>Cp_max</b>
<b>dnat</b>						
<b>du</b>	-0.71					
<b>dv</b>	0.89	-0.32				
<b>error</b>	-0.21	0.25	-0.18			
<b>Tm</b>	-0.09	0.08	-0.09	-0.15		
<b>Cp_max</b>	0.04	-0.71	-0.38	-0.07	0.07	
<b>theta_n</b>	-0.65	0.93	-0.27	0.15	0.05	0.05
<b>theta_t</b>	-0.64	0.89	-0.28	0.19	-0.02	-0.02
<b>theta_u</b>	-0.55	0.46	-0.39	-0.09	0.12	0.12
		<b>T. thermophilus</b>				
	<b>dnat</b>	<b>du</b>	<b>dv</b>	<b>error</b>	<b>Tm</b>	<b>Cp_max</b>
<b>dnat</b>						
<b>du</b>	-0.60					
<b>dv</b>	0.92	-0.26				
<b>error</b>	-0.30	0.12	-0.29			
<b>Tm</b>	0.13	0.19	0.20	0.16		
<b>Cp_max</b>	0.20	-0.89	-0.17	0.05	-0.16	
<b>theta_n</b>	-0.75	0.95	-0.45	0.21	0.05	0.05
<b>theta_t</b>	-0.85	0.80	-0.64	0.26	-0.04	-0.04
<b>theta_u</b>	-0.74	0.72	-0.52	0.23	0.04	0.04
		<b>Absolute difference</b>				
	<b>dnat</b>	<b>du</b>	<b>dv</b>	<b>error</b>	<b>Tm</b>	<b>Cp_max</b>
<b>dnat</b>						
<b>du</b>	0.11					
<b>dv</b>	0.03	0.06				
<b>error</b>	0.09	0.13	0.11			
<b>Tm</b>	0.22	0.11	0.29	0.31		
<b>Cp_max</b>	0.16	0.17	0.21	0.13	0.23	
<b>theta_n</b>	0.10	0.02	0.17	0.06	0.00	0.00
<b>theta_t</b>	0.21	0.09	0.35	0.08	0.02	0.02
<b>theta_u</b>	0.18	0.26	0.13	0.32	0.08	0.08

**Supp. Table 2.** Linear correlation coefficients comparing how the three phenomenological parameters compensate for each other. Correlations with other model predictions are also provided. A significant correlation between the phenomenological parameters and model predictions is only observed for  $\theta$ 's of key points along the unfolding pathway. The absolute values of differences between the *E. coli* and *T. thermophilus* results are provided as well; no significant differences within the correlations are observed.