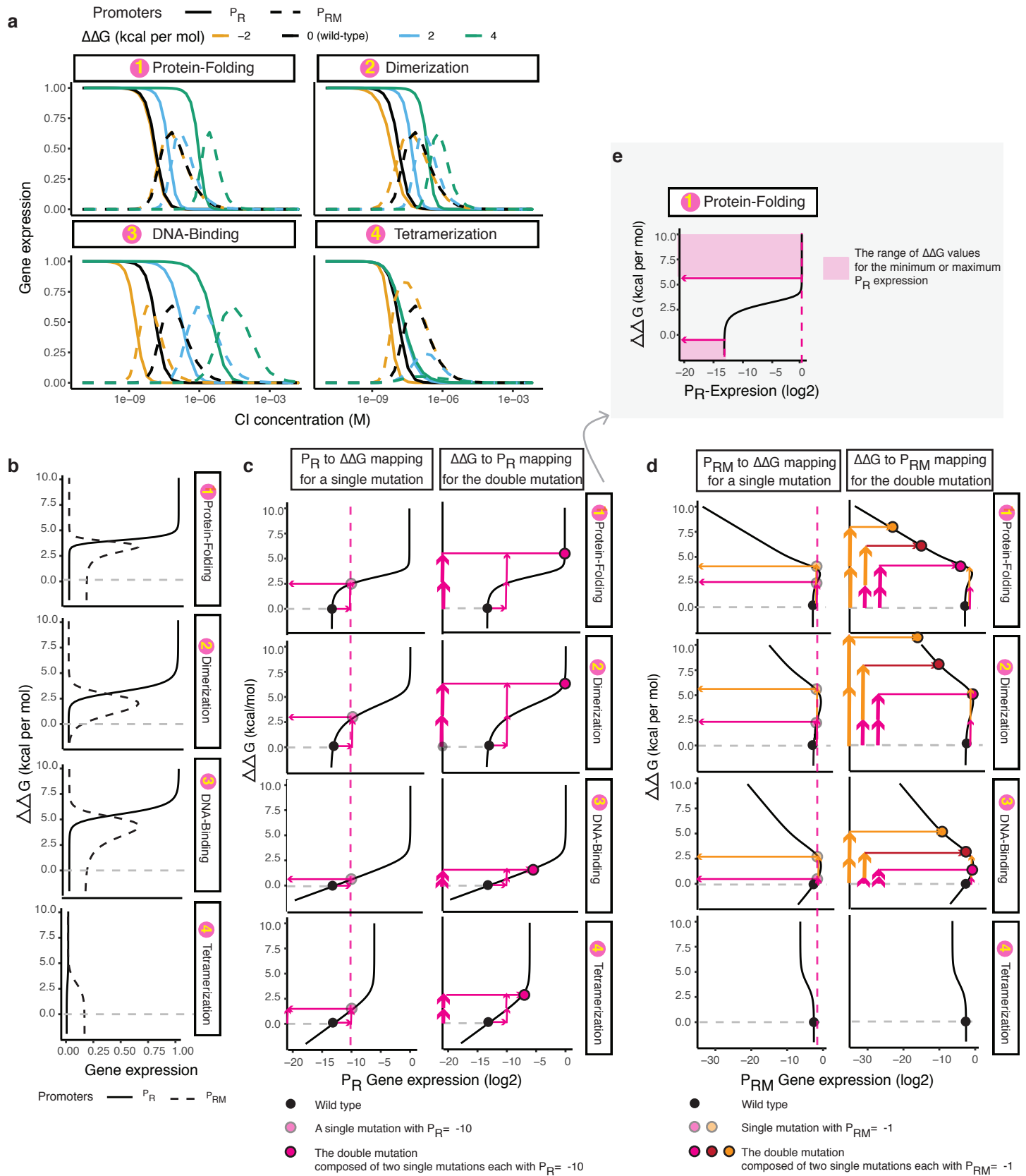


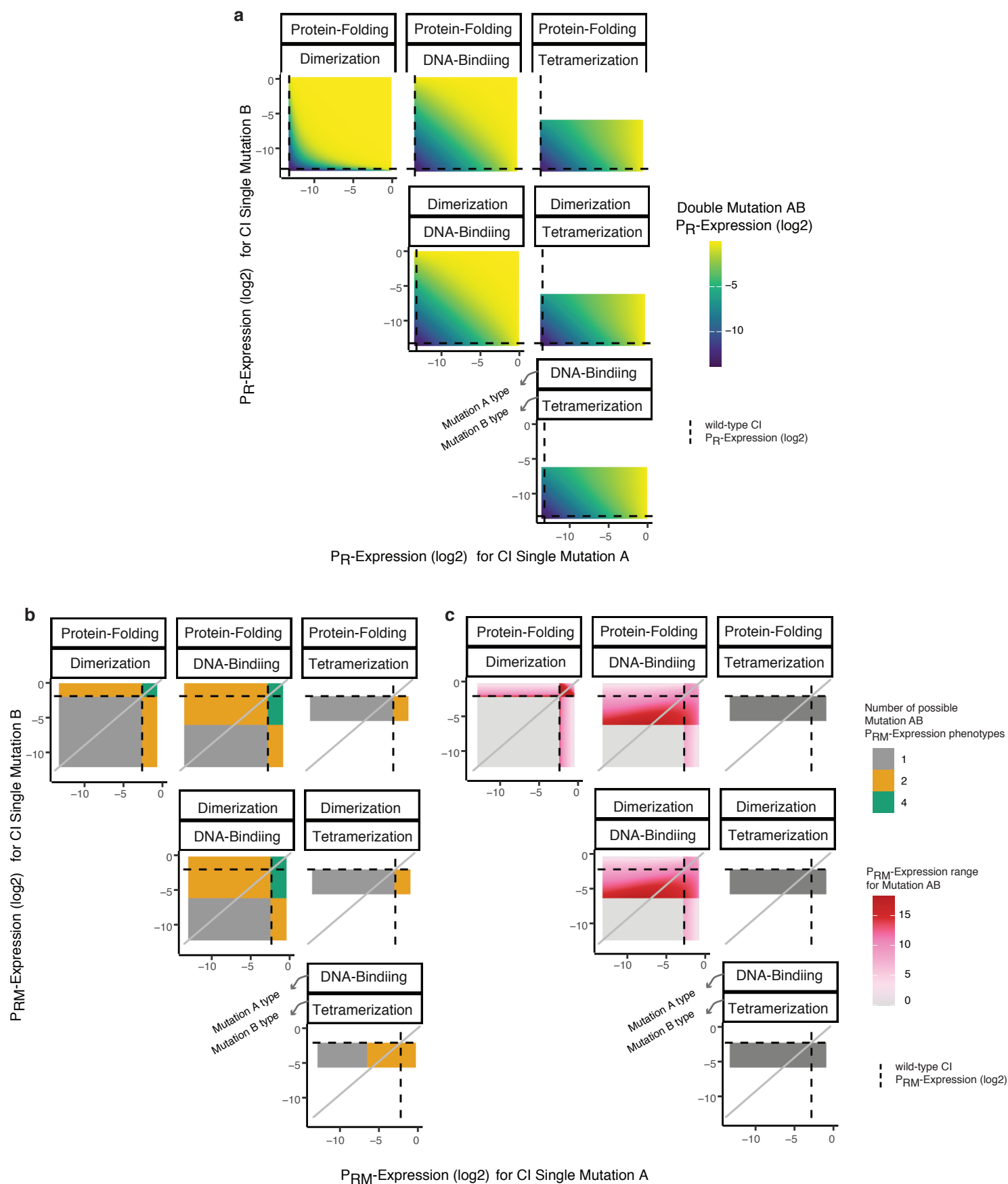
**Supplementary Information**

**Biophysical ambiguities prevent accurate genetic prediction**

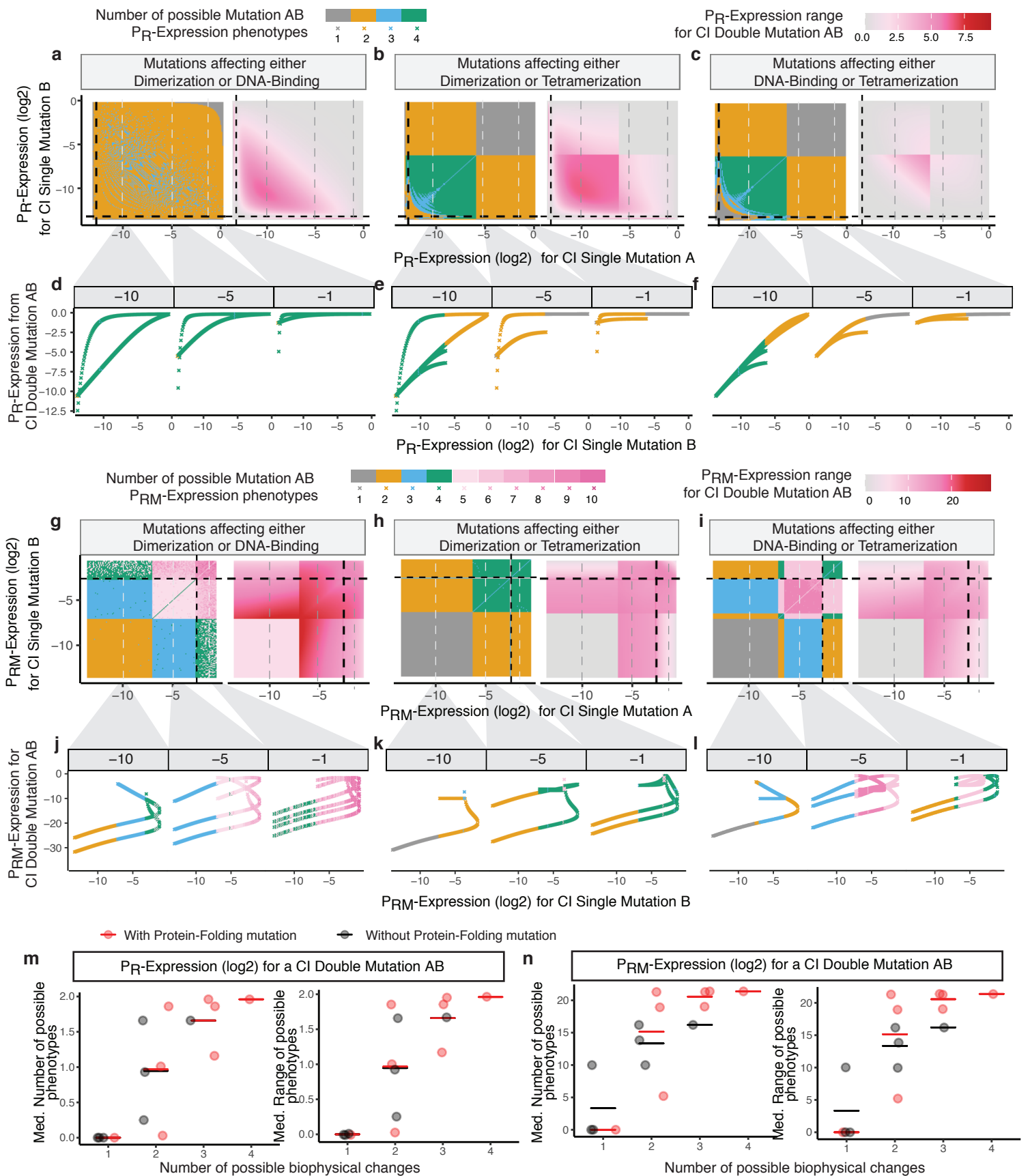
**Li et al.**



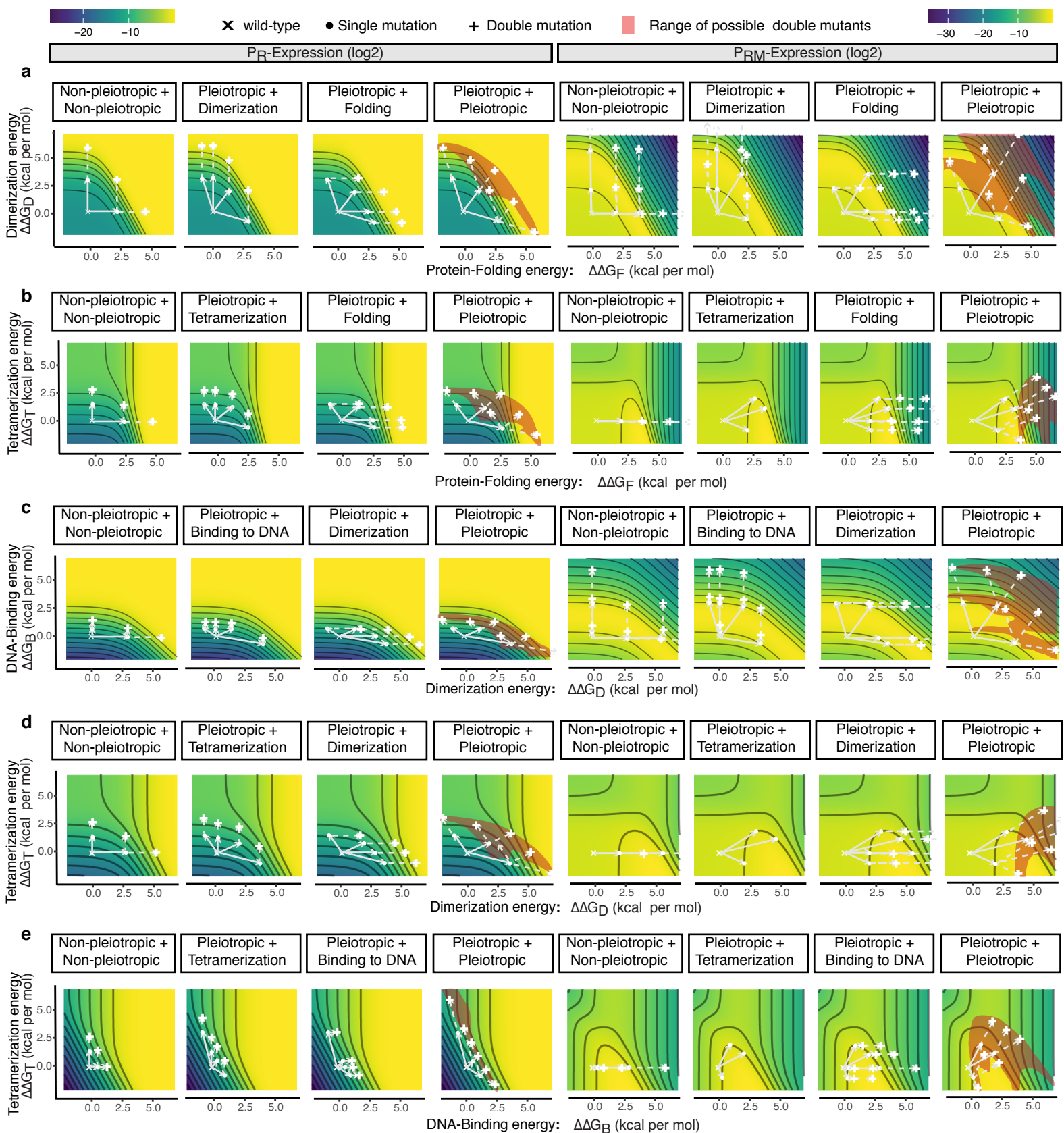
**Supplementary Figure 1. Relationships between biophysical parameters and phenotypes. (a)** Dose-response curves of the  $P_R$  and  $P_{RM}$  promoters with altered biophysical parameters. **(b)**  $P_R$  and  $P_{RM}$  phenotype - Free energy functions for mutations affecting the free energy of protein-folding, dimerization, DNA-binding, and tetramerization with the phenotypes. **(c, d)** Scheme for mapping  $P_R$  **(c)** and  $P_{RM}$  **(d)** phenotypes to biophysical parameters for single mutations (left panels) and from biophysical parameters to phenotypes for double mutations (right panels). Phenotypes are shown in  $\log_2$  scale. **(e)** Ambiguous mapping of phenotype to free energy changes due to measurement uncertainty when mutant phenotypes are close to minimum or maximum possible values. Source data are provided as a Source Data file.

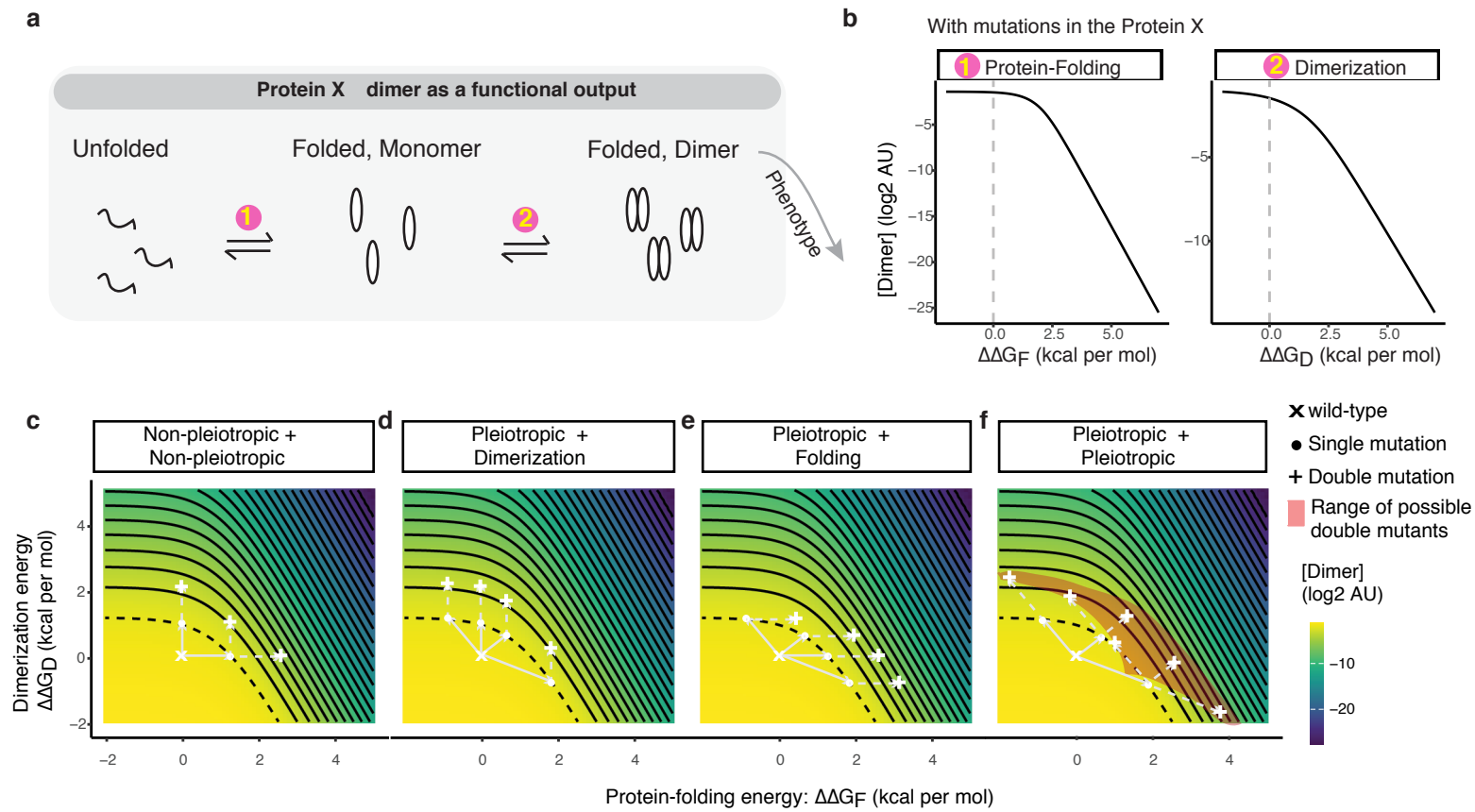


**Supplementary Figure 2. Combining mutations affecting two different biophysical parameters. (a)** Double mutant  $P_R$  phenotypes. **(b, c)** Number **(b)** and range **(c)** of possible double mutant  $P_{RM}$  phenotypes. Source data are provided as a Source Data file.



**Supplementary Figure 3. Biophysical ambiguity confounds phenotype prediction.** (a – c) Double mutant P<sub>R</sub> phenotypes when two single mutants can affect either of the two indicated biophysical parameters (but not both). Number of possible double mutant P<sub>R</sub> phenotypes (left panels) and range of possible double mutant P<sub>R</sub> phenotypes (right panels) are shown for each combination of single mutant phenotypes. (d – f) Examples showing how a mutation with a known phenotype combines with other mutations leading to 1 to 4 possible double mutant P<sub>R</sub> phenotypes. (g – i) Double mutant P<sub>RM</sub> expression when two single mutants are allowed to affect either of the two indicated biophysical parameters. Number (left panels) and range of possible double mutant P<sub>RM</sub> phenotypes (right panels) are shown for each combination of single mutant P<sub>RM</sub> phenotypes. (j – l) Examples showing how a mutation with a known P<sub>RM</sub> phenotype combines with other mutations, leading to many different possible double mutant P<sub>RM</sub> phenotypes. (m, n) The uncertainty of possible double mutant P<sub>R</sub> (m) and P<sub>RM</sub> (n) phenotypes increases with increasing biophysical ambiguity. Each data point represents the median number (left panels) and the median range (right panels) of double mutant phenotypes for a given combination of biophysical types. Horizontal lines denote the mean of the data points. n = 4, 6, 4 and 1 respectively for the groups with number of possible biophysical parameters equal to 1, 2, 3 and 4. Source data are provided as a Source Data file.





**Supplementary Figure 5. Biophysical ambiguity in a protein dimer system.** (a) Statistical thermodynamic model of protein dimerization. Protein X exists as 3 states: unfolded, folded monomer and folded dimer. The partitioning of these molecules depends on Gibbs free energy differences between states. (b) Mutations result in additive changes in the free energy of protein folding and dimerization, altering the protein dimer concentration. (c – f) Free energy-phenotype landscapes for mutations that affect the free energy of folding (x-axis) and/or dimerization energy (y-axis). Phenotypic isochores are drawn with an interval of 1 in log(2) scale. A continuous range of free energy changes can underlie an observed phenotype (dashed isochore). Combining two mutations with the same effect can result in a range of double mutant phenotypes (red shaded areas in f). Example double mutant outcomes are shown when neither (c), one (d, e) or both (f) mutations are pleiotropic. Source data are provided as a Source Data file.

**Supplementary Table 1.** Statistical thermodynamics of CI regulatory system

Cis-states	CI states	# CI <sub>2</sub> bound	P <sub>R</sub>	P <sub>RM</sub>	Statistical thermodynamic transition states	Energy-terms between the statistical thermodynamic transition states	Relative distribution of cis-regulatory element states based on the statistical thermodynamics	CI Relative molecule numbers, based on the statistical thermodynamics
	Unfolded <b>CI<sub>(U)</sub></b>				$CI_{(U)}/CI_{(M)}$	<b>-ΔG<sub>F</sub></b>		
	Monomer <b>CI<sub>(M)</sub></b>				$CI_{(M)}^2/CI_2$	<b>-ΔG<sub>D</sub></b>		
<b>c1</b>	Free Dimer <b>CI<sub>2</sub></b>	<b>0</b>	<b>ON</b>	<b>OFF</b>	Reference	<b>0</b>	$f_1 = \exp\left(\frac{0}{RT}\right) CI_2^0$	<b>2CI<sub>2</sub></b>
<b>c2</b>	Bound to OR3	<b>1</b>	<b>ON</b>	<b>OFF</b>	$c^2/(CI_2 \cdot OR)$	<b>ΔG<sub>2</sub></b> = ΔG <sub>OR3</sub>	$f_2 = \exp\left(\frac{-\Delta G_2}{RT}\right) CI_2^1$	<b>1 · 2OR · exp</b> $\left(\frac{-\Delta G_2}{RT}\right) CI_2^1$
<b>c3</b>	Bound to OR2	<b>1</b>	<b>OFF</b>	<b>ON</b>	$c^3/(CI_2 \cdot OR)$	<b>ΔG<sub>3</sub></b> = ΔG <sub>OR2</sub>	$f_3 = \exp\left(\frac{-\Delta G_3}{RT}\right) CI_2^1$	<b>1 · 2OR · exp</b> $\left(\frac{-\Delta G_3}{RT}\right) CI_2^1$
<b>c4</b>	Bound to OR1	<b>1</b>	<b>OFF</b>	<b>OFF</b>	$c^4/(CI_2 \cdot OR)$	<b>ΔG<sub>4</sub></b> = ΔG <sub>OR1</sub>	$f_4 = \exp\left(\frac{-\Delta G_4}{RT}\right) CI_2^1$	<b>1 · 2OR · exp</b> $\left(\frac{-\Delta G_4}{RT}\right) CI_2^1$
<b>c5</b>	Bound to OR1, OR3	<b>2</b>	<b>OFF</b>	<b>OFF</b>	$c^5/(CI_2^2 \cdot OR)$	<b>ΔG<sub>5</sub></b> = ΔG <sub>OR1</sub> + ΔG <sub>OR3</sub>	$f_5 = \exp\left(\frac{-\Delta G_5}{RT}\right) CI_2^2$	<b>2 · 2OR · exp</b> $\left(\frac{-\Delta G_5}{RT}\right) CI_2^2$
<b>c6</b>	Bound to OR2, OR3	<b>2</b>	<b>OFF</b>	<b>OFF</b>	$c^6/(CI_2^2 \cdot OR)$	<b>ΔG<sub>6</sub></b> = ΔG <sub>OR2</sub> + ΔG <sub>OR3</sub> + ΔG <sub>T</sub>	$f_6 = \exp\left(\frac{-\Delta G_6}{RT}\right) CI_2^2$	<b>2 · 2OR · exp</b> $\left(\frac{-\Delta G_6}{RT}\right) CI_2^2$
<b>c7</b>	Bound to OR1, OR2	<b>2</b>	<b>OFF</b>	<b>ON</b>	$c^7/(CI_2^2 \cdot OR)$	<b>ΔG<sub>7</sub></b> = ΔG <sub>OR1</sub> + ΔG <sub>OR2</sub> + ΔG <sub>T</sub>	$f_7 = \exp\left(\frac{-\Delta G_7}{RT}\right) CI_2^2$	<b>2 · 2OR · exp</b> $\left(\frac{-\Delta G_7}{RT}\right) CI_2^2$
<b>c8</b>	Bound to OR1, OR2, OR3	<b>3</b>	<b>OFF</b>	<b>OFF</b>	$c^8/(CI_2^3 \cdot OR)$	<b>ΔG<sub>8</sub></b> = ΔG <sub>OR1</sub> + ΔG <sub>OR2</sub> + ΔG <sub>OR3</sub> + ΔG <sub>T</sub>	$f_8 = \exp\left(\frac{-\Delta G_8}{RT}\right) CI_2^3$	<b>3 · 2OR · exp</b> $\left(\frac{-\Delta G_8}{RT}\right) CI_2^3$
<b>Sum</b>							$\sum_{i\{1:8\}} f_i$	<b>CI (Total)</b>

Red colour highlights values derived from the information that each CI dimer is composed of 2 CI monomers; light blue highlights number of CI dimer molecules bound to OR1-3; green highlights promoter states being ON.

**Supplementary Table 2. Parameters**

	Value	Unit	Reference
$\Delta G_F$	-1	kcal per mol	1, 2
$\Delta G_D$	-10.886	kcal per mol	2, 3
$\Delta G_{OR1}$	-11.7	kcal per mol	2, 3
$\Delta G_{OR2}$	-10.1	kcal per mol	2, 3
$\Delta G_{OR3}$	-10.1	kcal per mol	2, 3
$\Delta G_T$	-2	kcal per mol	2, 3
OR	1e-9	M	2, 3
$CI_{(Total)}$	8.4e-7	M	2

**Supplementary References**

1. Huang, G. S. & Oas, T. G. Structure and Stability of Monomeric  $\lambda$  repressor: NMR Evidence for Two-State Folding. *Biochemistry* **34**, 3884–3892 (1995).
2. Li, X., Lalic, J., Baeza-Centurion, P., Dhar, R. & Lehner, B. Changes in gene expression predictably shift and switch genetic interactions. *Nat. Commun.* **10**, 3886 (2019).
3. Ackers, G. K., Johnson, A. D. & Shea, M. A. Quantitative model for gene regulation by  $\lambda$  phage repressor. *Proc. Natl. Acad. Sci. U. S. A.* **79**, 1129–33 (1982).