nature structural & molecular biology

Implications of molecular heterogeneity for the cooperativity of biological macromolecules

Sergey V Solomatin[1,3], Max Greenfeld[1–3] & Daniel Herschlag[1]

Cooperativity, a universal property of biological macromolecules, is typically characterized by a Hill slope, which can provide fundamental information about binding sites and interactions. We demonstrate, through simulations and single-molecule FRET (smFRET) experiments, that molecular heterogeneity lowers bulk cooperativity from the intrinsic value for the individual molecules. As heterogeneity is common in smFRET experiments, appreciation of its influence on fundamental measures of cooperativity is critical for deriving accurate molecular models.

Cooperativity is a universal property of biological macromolecules that is exploited widely in biological regulation and function[1–4]. Cooperativity of ligand binding is defined as positive if the binding of one ligand molecule strengthens the binding of subsequent molecules, and as negative if the binding of one molecule weakens the binding of subsequent molecules. Quantitative measurements of cooperativity typically rely on fitting the Hill equation[5] to ligand-binding data:

$$f = \frac{([L]/L_{1/2})^n}{1+([L]/L_{1/2})^n} \quad (1)$$

where f is the fraction of the ligand bound, $[L]$ is the concentration of ligand and $L_{1/2}$ is the midpoint of binding. In the simple case of all-or-none binding, the Hill cooperativity coefficient n indicates the binding stoichiometry, and in situations in which the stoichiometry is known, the value of n provides fundamental information about intramolecular communication[6,7]. In bulk studies, individual macromolecules within an ensemble have traditionally been assumed to be identical. However, modern single-molecule assay methods are revealing that unexpected molecular heterogeneity exists in many instances (additional references are provided in the **Supplementary Note** because of space limitations)[8–12]. Here we address, through numerical simulations and single-molecule FRET experiments, how such heterogeneity distorts the cooperativity measured in bulk from the actual cooperativity of the individual molecules.

Consider an ensemble of macromolecules in which the affinity for a ligand varies between molecules. For simplicity, we assume that the affinity variation is characterized by a normal distribution (**Fig. 1a**). We refer to the s.d. of the distribution as the 'heterogeneity parameter' H. We further assume that all molecules have the same binding cooperativity—in this example, $n = 3$ (**Fig. 1b**, blue lines). When the individual binding curves are summed to give the overall curve for the ensemble, a shallower dependence on ligand concentration is obtained, with a bulk cooperativity parameter $n_{bulk} = 1.6$ for this example (**Fig. 1b**, red circles).

The origin of the decrease in apparent cooperativity can be readily seen from **Figure 1b**. The ensemble binding curve is the sum of individual binding curves, which are shifted relative to one another along the ligand concentration axis. The ensemble curve starts to rise before the highest-abundance binding curves rise because macromolecules with the highest affinities bind ligand at lower concentrations than the most highly represented molecules. At high ligand concentrations, the ensemble curve does not level off as sharply as the individual curves, because the macromolecules with the lowest affinities are not yet saturated at ligand concentrations that saturate other molecules. Thus, the curve for the ensemble of molecules has a shallower dependence on the ligand concentration than do the curves for the individual molecules.

The observed cooperativity derived from the ensemble curve becomes smaller as heterogeneity increases (**Fig. 1c**). With sufficient heterogeneity, cooperative binding can appear noncooperative

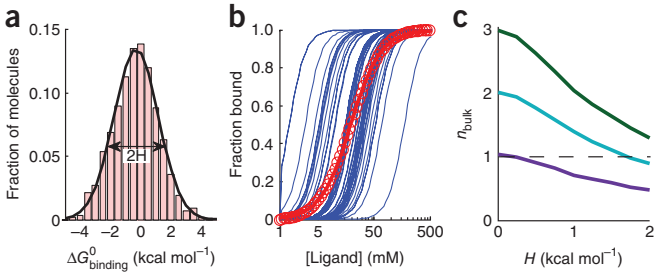

Figure 1 Simulations of ligand binding cooperativity with a heterogeneous population of molecules. (**a**) A simulated distribution of ligand-binding energies, $\Delta G^o_{binding}$, for 3,000 hypothetical molecules. The s.d. of the distribution, referred to as the heterogeneity parameter, H, is 1.5 kcal mol^{-1}. (**b**) Simulated ligand-binding curves for individual molecules with the cooperativity coefficient $n = 3$ (blue lines). Fifty randomly drawn curves are displayed. The bulk binding curve is the sum of curves of all of the molecules (red circles). The red line is the best Hill fit to the bulk binding curve, giving $n_{bulk} = 1.6$. (**c**) Reduction of the bulk cooperativity parameter n_{bulk} as a function of the heterogeneity parameter H for three different true values of n. The dashed line at $n_{bulk} = 1$ separates positive and negative observed cooperativity.

[1]Department of Biochemistry, Stanford University, Stanford, California, USA. [2]Department of Chemical Engineering, Stanford University, Stanford, California, USA. [3]These authors contributed equally to this work. Correspondence should be addressed to D.H. (herschla@stanford.edu).

Received 12 November 2010; accepted 28 February 2011; published online 15 May 2011; doi:10.1038/nsmb.2052

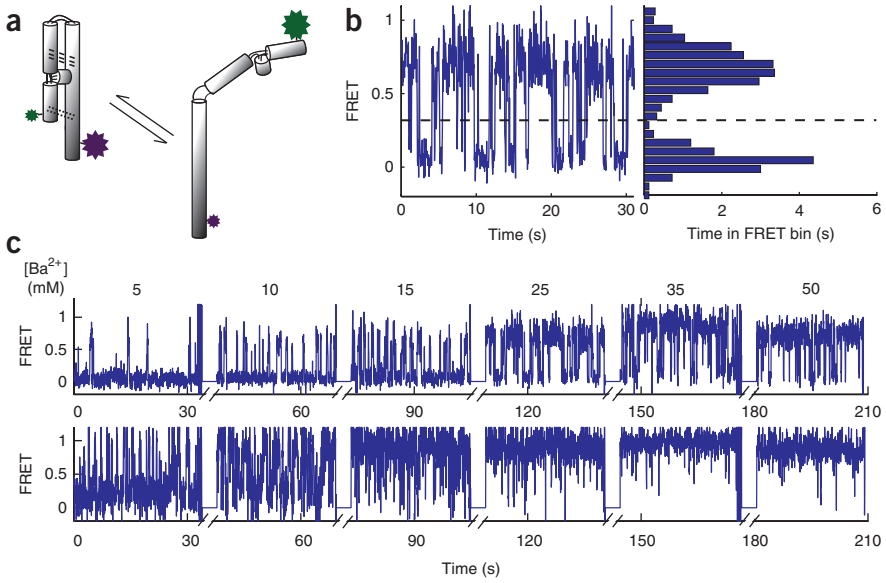

Figure 2 smFRET measurements of P4-P6 RNA folding. (**a**) Scheme of folding of fluorescent dye–labeled P4-P6. Green and purple stars indicate the positions of the donor (Cy3) and the acceptor (Cy5) dyes, respectively[13]. Dashed and dotted lines indicate tertiary interactions that stabilize the folded state. (The tertiary interaction indicated by the dashed lines involves bound Mg^{2+} ions and does not occur in the presence of Ba^{2+}; ref. 21.) (**b**) FRET trace of a single P4-P6 molecule showing fluctuations between a folded (high-FRET) and an unfolded (low-FRET) state. The dashed line is the threshold used to calculate $\Delta G^{o}_{folding}$ according to equation (2). (**c**) FRET traces of two individual P4-P6 molecules observed over a range of Ba^{2+} concentrations. Breaks in the *x* axis indicate points when data collection was suspended to change Ba^{2+} concentration.

(**Fig. 1c**; cyan curve gives $n = 1$ with a heterogeneity of $H = 1.7$ kcal mol^{-1}), and binding of a ligand that was cooperative or noncooperative can even appear anticooperative (**Fig. 1c**, cyan and magenta curves). Conversely, the more uniform the behavior of the molecules, the closer the value of n_{bulk} approaches the true value of n. Thus, only for a truly uniform ensemble can n_{bulk} be directly interpreted as reflecting the fundamental properties of individual molecules. To unambiguously interpret n_{bulk} measured for a potentially heterogeneous ensemble, the extent of heterogeneity must be known.

Single-molecule experiments have almost invariably revealed heterogeneity of biological macromolecules with respect to ligand binding, folding and activity[8–12]. Heterogeneity may be a property of macromolecules undetected in bulk experiments (from multiple long-lived conformations, or from multiple covalently distinct molecules) or an artifact of single-molecule experiments (such as interactions with heterogeneous surfaces used for immobilization of molecules). Some studies have provided evidence that heterogeneity is not an artifact of surface immobilization[10,12], so it is likely that heterogeneity exists in bulk experiments and distorts cooperativity measured in at least some experiments. Single-molecule experiments have unique abilities to directly measure the extent of heterogeneity and to obtain binding curves for individual molecules, and thus to directly determine n for each molecule.

Taking advantage of these abilities, we measured heterogeneity and cooperativity coefficients for ion-dependent folding of P4-P6 RNA using single-molecule FRET (smFRET). P4-P6 (**Fig. 2a**) is an independently folding domain from the *Tetrahymena thermophila* group I intron and is a convenient system for smFRET studies (**Fig. 2b,c**)[13]. The molecular basis for the cooperativity coefficients for ion-dependent folding is complex because most ions are not site-bound to RNA, but instead form a diffuse 'ion atmosphere'[14–17]. Nevertheless, the effect of heterogeneity on n arises from differences in the folding of individual molecules and does not depend on the underlying basis of n. It also does not depend on the molecular basis of heterogeneity, which for P4-P6 is mostly covalent[18]. In other systems heterogeneity is, at least in substantial part, conformational[12]. We used P4-P6 here as a powerful system to demonstrate the effect of heterogeneity on the cooperativity with actual experimental data. The overall interpretation holds regardless of whether the heterogeneity is covalent or conformational.

We measured heterogeneity of P4-P6 folding for an ensemble of 126 molecules in Ba^{2+} at a concentration close to the folding midpoint (**Fig. 2b**). These conditions allowed us to follow many folding transitions before dye photobleaching and thus to obtain the most accurate measurements of the folding free energy $\Delta G^{o}_{folding}$ (**Supplementary Methods**)[12,19]. The value of $\Delta G^{o}_{folding}$ for each molecule was calculated from smFRET traces as follows:

$$\Delta G^{o}_{folding} = -RT \times \ln\left(\frac{t_{folded}}{t_{unfolded}}\right) \quad (2)$$

where t_{folded} and $t_{unfolded}$ are the times each molecule spent folded and unfolded, respectively. Individual P4-P6 traces produced a range of $\Delta G^{o}_{folding}$ values from –1 kcal mol^{-1} to 1.5 kcal mol^{-1} (**Fig. 3a**). This range was much broader than expected from statistical variability arising from finite time sampling and corresponds to molecular heterogeneity (**Supplementary Methods** and **Supplementary Fig. 1**). Fitting a Gaussian to the histogram of $\Delta G^{o}_{folding}$ allowed us to calculate the heterogeneity parameter $H = 0.8 \pm 0.2$ kcal mol^{-1} . The same value of H was obtained for a much larger ensemble of 1,241 molecules (**Supplementary Fig. 2**), indicating that 126 molecules adequately cover P4-P6 heterogeneity.

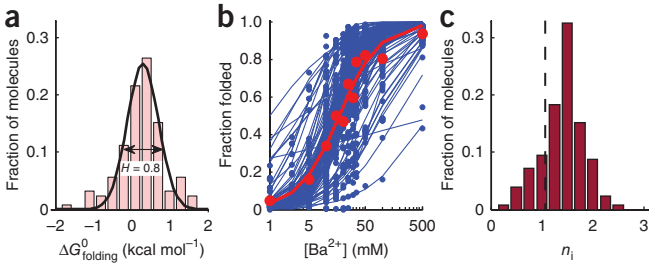

Figure 3 Cooperativity of Ba^{2+}-dependent folding for 126 P4-P6 molecules measured by smFRET. (**a**) Heterogeneity of $\Delta G^{o}_{folding}$ for 126 individual P4-P6 molecules at $[Ba^{2+}] = 10$ mM. Heterogeneity parameter $H = 0.8 \pm 0.2$ kcal mol^{-1} (the error is the s.e.m.). (**b**) Individual folding isotherms for 126 P4-P6 molecules and Hill fits to each of these isotherms (blue circles and blue lines, respectively). The bulk folding data (red circles) were calculated as the sum of the data for all of the molecules and fit with the Hill equation (red line). (**c**) The distribution of individual cooperativity parameters n_i obtained from the fits shown in **b**. The value of n_{bulk} is indicated by a dashed line.

At this level of heterogeneity, our simulations predict that n_{bulk} is substantially lower than the true value for individual molecules. To test this prediction, we designed single-molecule titration experiments and obtained folding isotherms for individual P4-P6 molecules (**Fig. 2c** and **Supplementary Methods**). The bulk folding isotherm was obtained from the total time all of the molecules spent in the folded state at each Ba^{2+} concentration. The best fit of the Hill equation to the bulk folding curve (**Fig. 3b**, red symbols) gave the cooperativity parameter $n_{bulk} = 1.1 \pm 0.14$. This value of n is indistinguishable from 1, which could have suggested an absence of folding cooperativity with respect to Ba^{2+}. However, when we fitted the Hill equation to the folding isotherms of the individual molecules (**Fig. 3b**) and directly determined the individual cooperativity coefficients, n_i, we observed $n_i = 1.5 \pm 0.3$ for most molecules (**Fig. 3c**). These measurements indicate that folding is, in fact, cooperative, and directly illustrate how molecular heterogeneity distorts ensemble-averaged cooperativity. Furthermore, individual molecules had different cooperativity coefficients, with n_i varying from as low as 0.5 ± 0.12 to as high as 2.7 ± 0.7 (**Fig. 3c**). The differences in cooperativity parameters between molecules were much larger than the errors associated with these parameters (**Supplementary Fig. 3**). These observations further underscore the wealth of mechanistic information that can be revealed in single-molecule experiments.

In conclusion, we have demonstrated, through simulation and the first reported experimental single-molecule titrations, how molecular heterogeneity distorts cooperativity observed in ensemble measurements. Knowledge of the cooperativity actually shown by individual molecules is critical for developing an atomic-level mechanistic understanding of macromolecular behavior. In addition, there is increasing awareness of molecular heterogeneity, and possible roles of heterogeneity in biological systems are being widely discussed (for example, ref. 20). We speculate that controlled heterogeneity, possibly established by variations in covalent modification of proteins, might be used to fine-tune cooperative behavior in biology.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank H. Mabuchi (Department of Applied Physics, Stanford University) for technical support and D. Pavlichin for helping to create data analysis software. This work was funded by the US National Institutes of Health through grant GM49243 to D.H.

AUTHOR CONTRIBUTIONS

S.V.S., M.G. and D.H. contributed to the experimental design and writing of the manuscript; M.G. performed the experiments; and M.G. and S.V.S. carried out data analysis.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/nsmb/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. De Meyts, P., Roth, J., Neville, D.M., Gavin, J.R. & Lesniak, M.A. *Biochem. Biophys. Res. Commun.* **55**, 154–161 (1973).
2. Levitzki, A. & Koshland, D.E. *Proc. Natl. Acad. Sci. USA* **62**, 1121–1128 (1969).
3. Lohman, T.M. & Ferrari, M.E. *Annu. Rev. Biochem.* **63**, 527–570 (1994).
4. Perutz, M.F. *Q. Rev. Biophys.* **22**, 139–237 (1989).
5. Hill, A.V. *J. Phys.* **40**, iv–vii (1910).
6. Weiss, J.N. *FASEB J.* **11**, 835–841 (1997).
7. Wyman, J. *Adv. Protein Chem.* **19**, 223–286 (1964).
8. Lu, H.P., Xun, L.Y. & Xie, X.S. *Science* **282**, 1877–1882 (1998).
9. Zhuang, X. *et al. Science* **296**, 1473–1476 (2002).
10. Okumus, B., Wilson, T.J., Lilley, D.M.J. & Ha, T. *Biophys. J.* **87**, 2798–2806 (2004).
11. Qu, X.H. *et al. Proc. Natl. Acad. Sci. USA* **105**, 6602–6607 (2008).
12. Solomatin, S.V., Greenfeld, M., Chu, S. & Herschlag, D. *Nature* **463**, 681–684 (2010).
13. Sattin, B.D., Zhao, W., Travers, K., Chu, S. & Herschlag, D. *J. Am. Chem. Soc.* **130**, 6085–6087 (2008).
14. Draper, D.E., Grilley, D. & Soto, A.M. *Annu. Rev. Biophys. Biomol. Struct.* **34**, 221–243 (2005).
15. Draper, D.E. *Biophys. J.* **95**, 5489–5495 (2008).
16. Das, R., Travers, K.J., Bai, Y. & Herschlag, D. *J. Am. Chem. Soc.* **127**, 8272–8273 (2005).
17. Das, R. *et al. J. Mol. Biol.* **332**, 311–319 (2003).
18. Greenfeld, M., Solomatin, S.V., Herschlag, D. *J. Biol. Chem.* doi:10.1074/jbc.M111.235465 (08 April 2011).
19. Bartley, L.E., Zhuang, X.W., Das, R., Chu, S. & Herschlag, D. *J. Mol. Biol.* **328**, 1011–1026 (2003).
20. Tawfik, D.S. *Nat. Chem. Biol.* **6**, 692–696 (2010).
21. Travers, K.J., Boyd, N. & Herschlag, D. *RNA* **13**, 1205–1213 (2007).