

Cite this: *Chem. Sci.*, 2021, 12, 5944

All publication charges for this article have been paid for by the Royal Society of Chemistry

# Time-dependent communication between multiple amino acids during protein folding†

Song-Ho Chong  and Sihyun Ham \*

Cooperativity is considered to be a key organizing principle behind biomolecular assembly, recognition and folding. However, it has remained very challenging to quantitatively characterize how cooperative processes occur on a concerted, multiple-interaction basis. Here, we address how and when the folding process is cooperative on a molecular scale. To this end, we analyze multipoint time-correlation functions probing time-dependent communication between multiple amino acids, which were computed from long folding simulation trajectories. We find that the simultaneous multiple amino-acid contact formation, which is absent in the unfolded state, starts to develop only upon entering the folding transition path. Interestingly, the transition state, whose presence is connected to the macrostate cooperative behavior known as the two-state folding, can be identified as the state in which the amino-acid cooperativity is maximal. Thus, our work not only provides a new mechanistic view on how protein folding proceeds on a multiple-interaction basis, but also offers a conceptually novel characterization of the folding transition state and the molecular origin of the phenomenological cooperative folding behavior. Moreover, the multipoint correlation function approach adopted here is general and can be used to expand the understanding of cooperative processes in complex chemical and biomolecular systems.

Received 25th December 2020  
Accepted 23rd March 2021

DOI: 10.1039/d0sc07025d

rsc.li/chemical-science

## Introduction

Biomolecular assembly, recognition and folding are complex processes in which building blocks, such as amino acids in proteins, search for favorable inter- or intra-molecular interactions in intricate manners.<sup>1–3</sup> Cooperativity has been recognized to be a key concept associated with these processes.<sup>4–6</sup> However, cooperativity in macromolecular systems is typically described at a phenomenological, macrostate level, and is broadly defined as a characteristic of processes in which intermediate states are disfavored, *i.e.*, only the extreme states are significantly populated. Such all-or-none behavior, corresponding to switching between “on” and “off” states, is critical in regulation and signaling to avoid undesirable effects. The all-or-none character in ligand binding—receptor binding sites are either empty or fully occupied—is the basis for the Hill equation, which provides a commonly adopted measure of cooperativity.<sup>7</sup> The cooperativity concept in protein folding was also introduced at the macrostate level,<sup>8</sup> conveying that folding proceeds in a two-state, all-or-none fashion.

Such a macrostate cooperativity concept, however, does not reveal underlying molecular mechanisms. In this regard, we

notice that the cooperativity between two events *A* and *B* can in general be captured by the correlation,  $\chi = P(A, B) - P(A)P(B)$ , defined in terms of the joint probability  $P(A, B)$  and the product  $P(A)P(B)$  of the probabilities of individual events:<sup>9,10</sup>  $\chi > 0$  or  $\chi < 0$  corresponds to positive or negative cooperativity, respectively. For example, when *A* and *B* refer to ligand binding events to receptor sites *i* and *j*,  $\chi > 0$  indicates that the conditional probability  $P(B/A) = P(A, B)/P(A)$  is larger than  $P(B)$ , *i.e.*, the ligand binding to site *i* enhances the binding affinity to site *j* from what it would be in isolation. Thus, the cooperativity formulated with  $\chi$  is able to uncover the existence of a certain communication between molecular events occurring at distinct sites (the term “communication” is used here only in this sense, *i.e.*, when the correlation or cooperativity quantified by  $\chi \neq 0$  is present). Owing to the recent advances in experimental and computational technologies, the folding transition path that was previously inaccessible has now become within our reach.<sup>11–16</sup> The folding transition path is a small fraction of equilibrium folding trajectories where the folding process actually takes place. The transition path thus contains, in principle, all the mechanisms of protein folding, and there must be certain concerted molecular processes that underlie the macrostate folding cooperativity.

Here, we investigate the folding cooperativity through the correlation  $\chi$  defined with microscopic events occurring in the transition path. This is done for a number of small globular proteins displayed in Fig. 1 (see also Table S1†), whose all-atom

Department of Chemistry, The Research Institute of Natural Sciences, Sookmyung Women's University, Cheongpa-ro-47-gil 100, Yongsan-ku, Seoul 04310, Korea.  
E-mail: sihyun@sookmyung.ac.kr

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0sc07025d

simulations were reported by Shaw and coworkers.<sup>17–20</sup> Since protein folding requires the establishment of native amino-acid contacts, we will choose the formations of those contacts as the relevant microscopic events. Of particular interest in the present work is the timing (early, intermediate, or late stage) at which the cooperativity sets in during the transition path. To achieve this goal,  $\chi(t)$  carrying the time-dependence shall be introduced, which hence probes time-dependent cooperativity or communication between amino acids. Thereby, we would like to address how and when the folding process is cooperative on a molecular scale. We will then argue how such microscopic cooperativity is connected to the emergence of the macrostate cooperative folding behavior.

## Results

We start from surveying the folding behavior of the systems studied here. To succinctly describe our results, we will mainly deal with the  $\alpha$ -helical villin headpiece subdomain (HP-35) in the following; the results for the  $\beta$ -sheet WW domain (FIP35) are also included in the main text, and those for the other eight systems are presented in Fig. S1 to S8.† The folding process is monitored by the fraction of native amino-acid contacts  $Q$  ( $0 \leq Q \leq 1$ ), which was reported to be a good reaction coordinate of folding.<sup>21</sup> We computed  $Q(\mathbf{r}(t))$  for each protein configuration  $\mathbf{r}(t)$  along the trajectory (Fig. 2A), and constructed the probability distribution  $P(Q)$  of sampled  $Q(\mathbf{r}(t))$  values. The folding reaction free energy profile is then obtained from  $F(Q) = -k_B T \log P(Q)$  with Boltzmann's constant  $k_B$  and temperature  $T$  (Fig. 2B). It is observed that the system stays most of the time either in the folded or unfolded state (Fig. 2A) and that the unfolded- ( $Q_u$ ) and folded-state minima ( $Q_f$ ) are separated by a transition-state maximum ( $Q^*$ ), whose locations are indicated by the dashed lines (Fig. 2B). These results represent a typical two-state behavior in the sense of the original, macrostate cooperativity concept.

The transition path is a portion of the trajectory that starts from an unfolded configuration ( $Q(\mathbf{r}) < Q_u$ ) and ends at a folded one ( $Q(\mathbf{r}) > Q_f$ ) without recrossing the  $Q = Q_u$  line. To detect

cooperativity among multiple amino acids, we introduce a time-dependent correlation,

$$\chi_{ij;kl}(t) = \langle \sigma_{ij}(0) \sigma_{ij}(t) \sigma_{kl}(0) \sigma_{kl}(t) \rangle - \langle \sigma_{ij}(0) \sigma_{ij}(t) \rangle \langle \sigma_{kl}(0) \sigma_{kl}(t) \rangle. \quad (1)$$

Here, the time  $t$  is measured relative to the beginning of the transition path (*i.e.*,  $Q(\mathbf{r}(t)) = Q_u$  at  $t = 0$ );  $\sigma_{ij}(t)$  is equal to 1 when there is a contact between a pair of amino acids  $i$  and  $j$  at time  $t$ , and equal to  $-1$  otherwise;  $\sigma_{ij}(0) \sigma_{ij}(t)$  therefore varies from 1 to  $-1$  when a contact absent at time  $t = 0$  is formed at time  $t$ ; and the angular brackets denote an average over the configurations at  $t = 0$ . By definition,  $\chi_{ij;kl}(t) = 0$  when the contact formations of  $(i, j)$  and  $(k, l)$  amino-acid pairs occur independently. Therefore,  $\chi_{ij;kl}(t) > 0$  indicates the existence of positive cooperativity between  $(i, j)$  and  $(k, l)$  amino-acid pairs at time  $t$ . We also introduce  $\chi(t) = (1/N) \sum_{(ij),(kl)} \chi_{ij;kl}(t)$  averaged over all the pairs

forming native amino-acid contacts, with  $N$  denoting the number of those pairs, which is a measure of an overall strength of the cooperativity present in a protein at time  $t$ . The time-dependent correlation  $\chi(t)$ , when viewed as a multipoint time-correlation function, is an analog of the dynamic susceptibility used for probing cooperative dynamics in glass-forming supercooled systems.<sup>22–24</sup>

We computed  $\chi(t)$  for the transition path (cyan curve in Fig. 2C) by averaging over all the transition paths identified in each system. We also computed  $\chi(t)$  for the unfolded state (magenta curve in Fig. 2C) using the trajectory parts that are close to  $Q = Q_u$  (painted magenta in Fig. 2B). We find that, while  $\chi(t)$  for the unfolded state remains small at all the times, the one for the transition path develops a significant peak. We confirmed that the peak indeed originates from the correlation of distinct amino-acid pairs by comparing the diagonal ( $(i, j) = (k, l)$ ) and off-diagonal ( $(i, j) \neq (k, l)$ ) contributions to  $\chi(t)$  (Fig. S9†), to be denoted as  $\chi_{\text{diag}}(t)$  and  $\chi_{\text{off-diag}}(t)$  in the following. Thus, the growth of the amino-acid correlation is a distinguishing characteristic unique to the transition path. Such a behavior of  $\chi(t)$  as a function of time closely resembles that of a microscopic measure of “thermodynamic

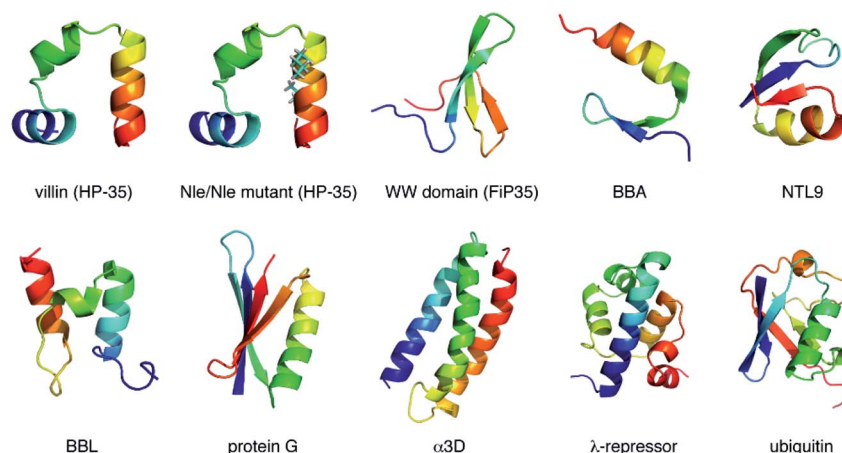
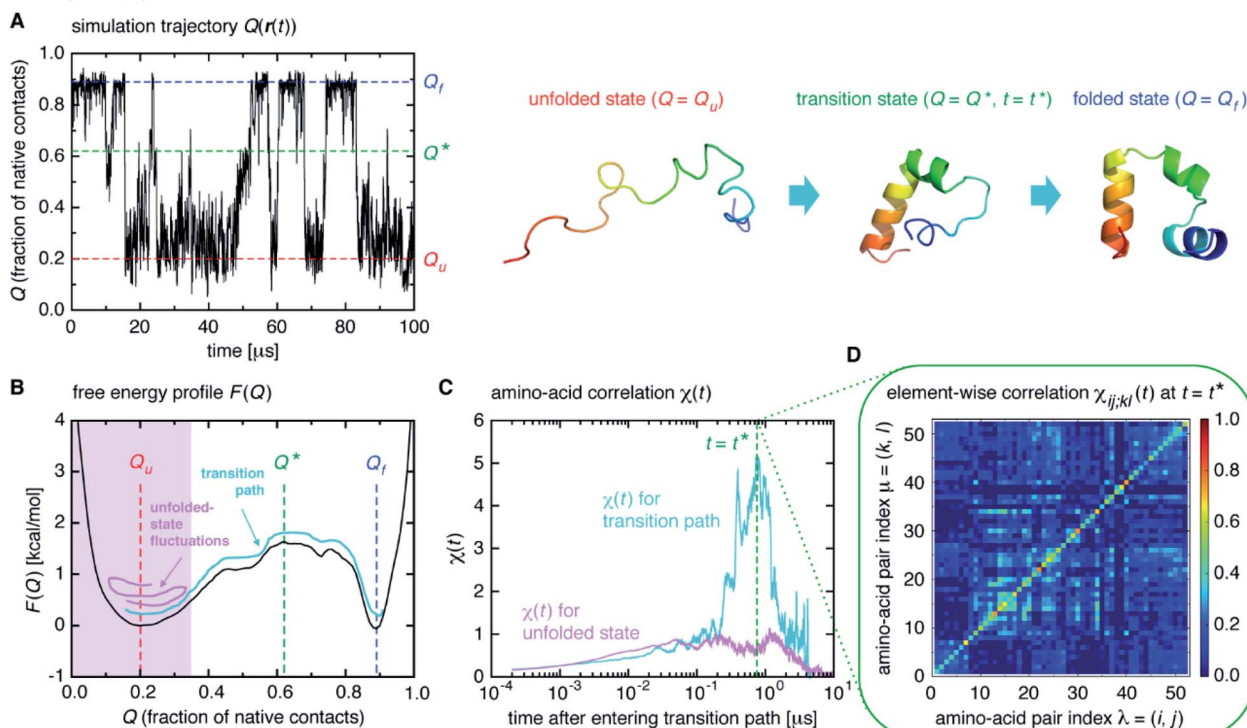


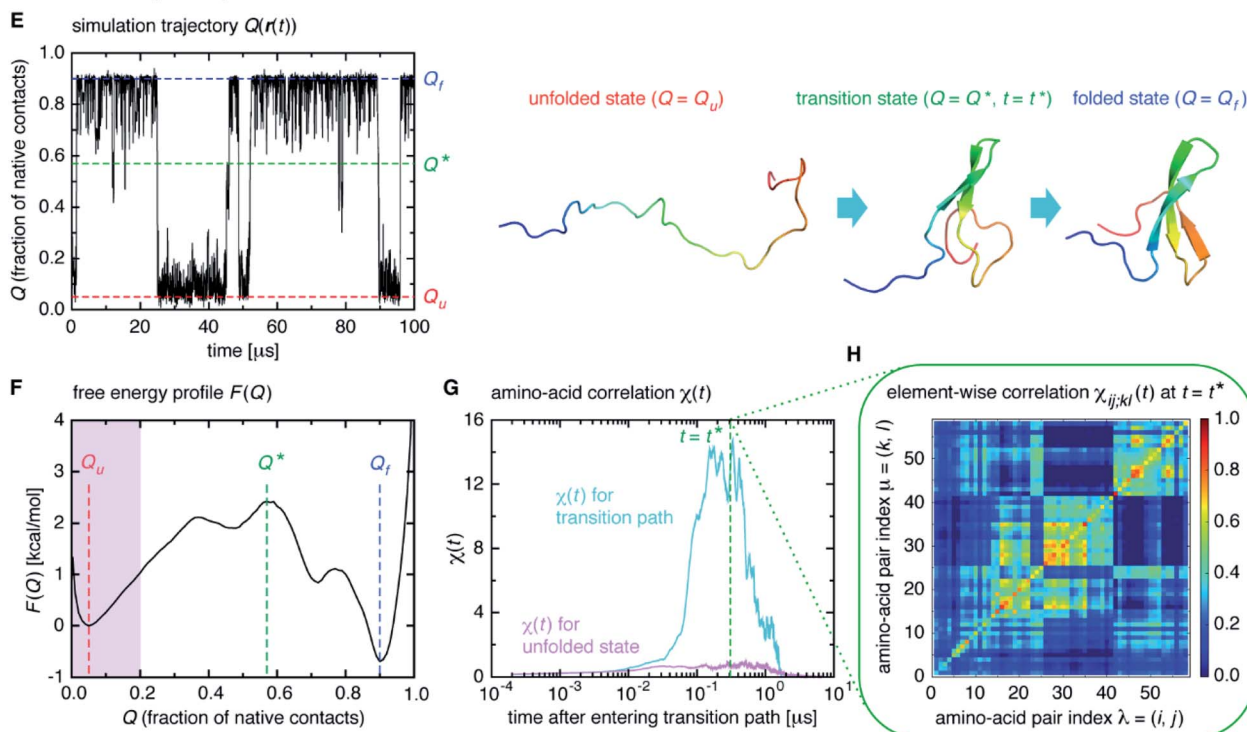
Fig. 1 Proteins studied in the present work. Proteins are color coded according to the sequence, ranging from blue (N-terminus) to red (C-terminus). Norleucine (Nle) residues in the Nle/Nle mutant are indicated by the stick representations.



## villin (HP-35)



## WW domain (FiP35)



**Fig. 2** Time-dependent cooperativity between multiple amino acids. (A) Fraction of the native amino-acid contacts  $Q(r(t))$  for the protein configuration  $r(t)$  at time  $t$  for a 100  $\mu$ s portion of the simulation trajectory of the villin headpiece subdomain (HP-35). (B) Folding free energy profile  $F(Q)$  versus  $Q$ . (C)  $\chi(t)$  for the transition path (colored cyan) and for the unfolded state (colored magenta) on a logarithmic timescale. (D) Element-wise  $\chi_{ij;kl}(t)$  at  $t = t^*$ . (E–H) Corresponding results for the WW domain (FiP35).

cooperativity" versus temperature,<sup>25</sup> and the cooperativity described by  $\chi(t)$  may be termed the dynamic cooperativity. Our observation is also consistent with the recent NMR

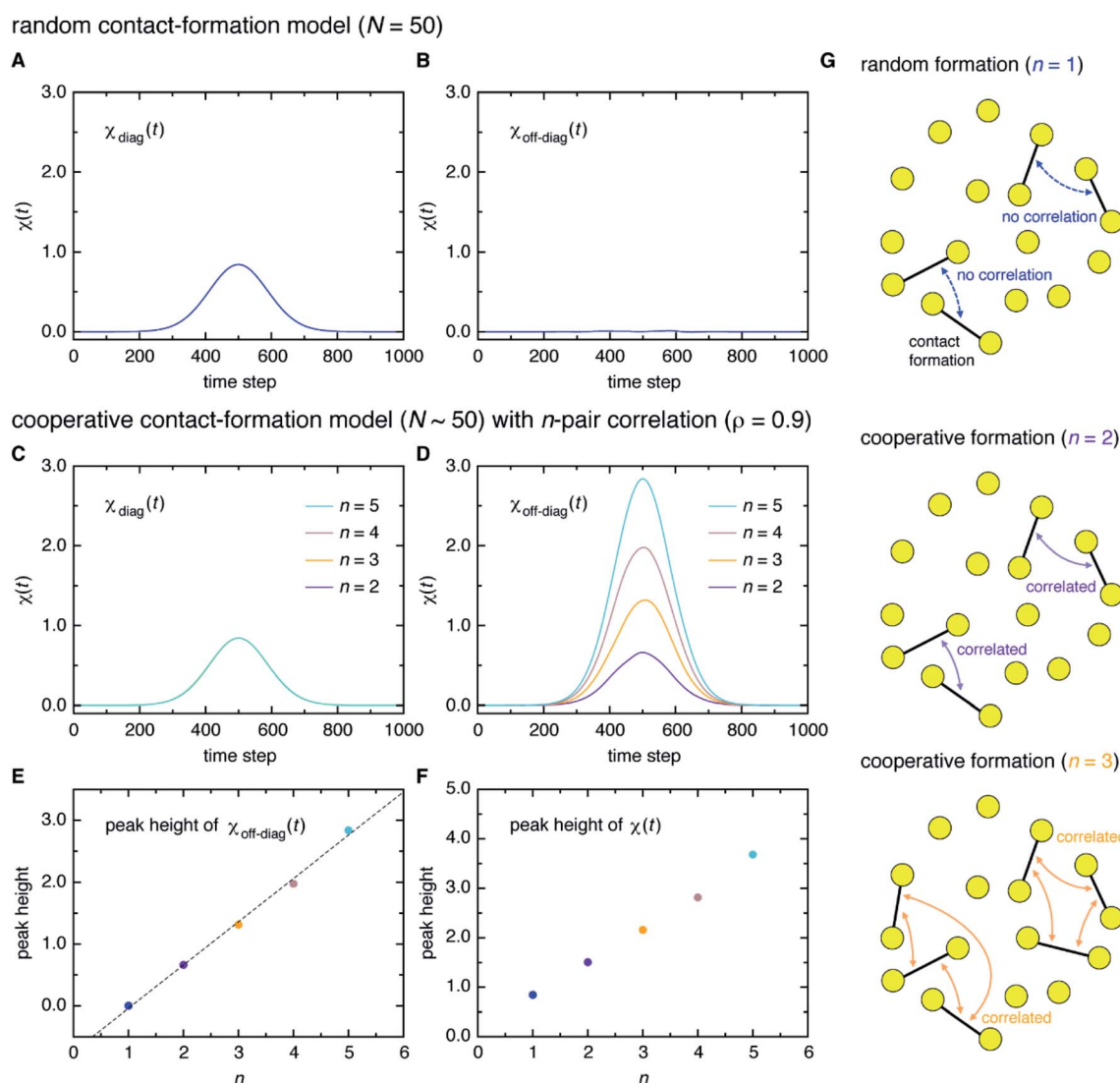
measurements demonstrating that the amino acids forming key contacts in the transition state interact not simultaneously in the denatured state.<sup>26</sup>



Here, a digression might be useful to better understand the nature of  $\chi(t) = \chi_{\text{diag}}(t) + \chi_{\text{off-diag}}(t)$  since a peak in  $\chi(t)$  may arise from a trivial reason, *i.e.*, just from the fact that a number of amino-acid contacts are formed roughly at the same time (in fact, the folding occurs within quite a short duration of time as can be inferred from Fig. 2A). We introduce a simple random model in which amino-acid pair contact formations are assumed to occur at random, Gaussian distributed times about the middle of the transition path. We find that  $\chi(t)$  of this model exhibits a peak whose height is about 1. However, since this model does not incorporate any correlations between distinct amino-acid pairs, such a peak entirely reflects the “self” term, *i.e.*,  $\chi(t) \approx \chi_{\text{diag}}(t) \approx 1$  and  $\chi_{\text{off-diag}}(t) \approx 0$  (Fig. 3A and B). Thus, the mere presence of a peak in  $\chi(t)$  does not warrant the existence of cooperative processes. We next consider an extended

model in which correlations (characterized by the correlation coefficient  $\rho$ ) are imposed between contact formation times of  $n$  amino-acid pairs. This model can be implemented by using the  $n$ -variate Gaussian distribution.<sup>9</sup> (We notice that  $n = 1$  corresponds to the random model.) We find for the model of  $\rho = 0.9$  that, whereas  $\chi_{\text{diag}}(t)$  remains the same as that of the random model, the peak of  $\chi_{\text{off-diag}}(t)$  increases linearly with  $n$  and that the peak height of  $\chi(t)$  provides a very rough estimate of the average number of correlated contact pairs (Fig. 3C to F). Thus,  $\chi(t)$  conforming to  $\chi_{\text{off-diag}}(t) \gg 1$ , which holds in the protein systems studied here (Fig. S9†), indeed indicates the presence of highly cooperative amino-acid contact formation.

Interestingly, we find that the time at which the amino-acid cooperativity attains its maximum size corresponds to the time when the system crosses the transition state. Not only can this



**Fig. 3** Random versus cooperative contact-formation models. (A and B) The self-component  $\chi_{\text{diag}}(t)$  (A) and the distinct component  $\chi_{\text{off-diag}}(t)$  (B) computed from the random contact-formation model of  $N = 50$ . (C and D) Corresponding results from the cooperative contact-formation model that incorporates correlation ( $\rho = 0.9$ ) between  $n$  amino-acid pairs. (E and F) The peak height of  $\chi_{\text{off-diag}}(t)$  along with a linear fit denoted by the dashed line (E), and the peak height of  $\chi(t)$  (F) for  $n = 1$  to  $5$  ( $n = 1$  corresponds to the random contact-formation model). (G) Schematic illustration of the random ( $n = 1$ ) and cooperative ( $n = 2$  and  $3$ ) contact formation in which vertices (yellow circles) represent amino acids and edges (black lines) represent the formation of native amino-acid contacts.





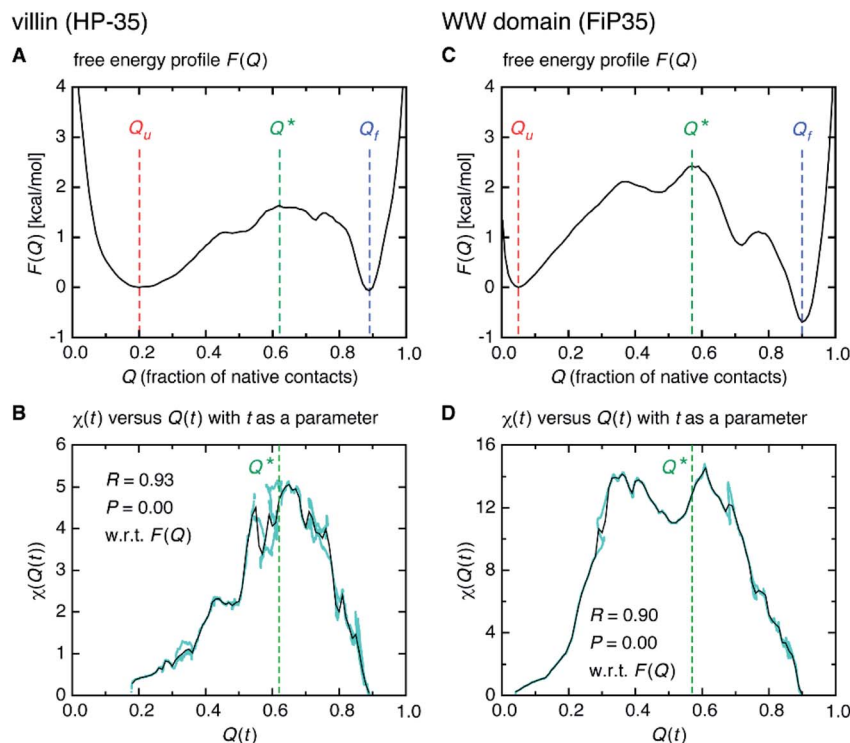


Fig. 4 Connection between the macrostate (thermodynamic) and microscopic (dynamic) cooperativity. (A) Folding free energy profile  $F(Q)$  versus  $Q$  of the villin headpiece subdomain (HP-35). (B) Parametric plot of  $\chi(t)$  versus  $Q(t)$  with  $t$  as a parameter (cyan filled circles). The black solid line was obtained after taking the average along the vertical direction for each  $Q = Q(t)$ . (C and D) Corresponding results for the WW domain (FiP35).

be identified in Fig. 2C, but it can be also observed in the corresponding figures for the other systems, in which the average time  $t^*$  the transition state is reached (i.e.,  $Q(t^*) = Q^*$  with  $Q(t) = \langle Q(r(t)) \rangle$ ) is indicated by the vertical dashed line. This implies that the transition state can be characterized as the state in which the amino-acid cooperativity is maximal. To further corroborate this observation, we plotted  $\chi(t)$  as a function of  $Q(t)$  with  $t$  as a parameter. The resulting  $\chi(Q(t))$  profile is shown and compared with the free energy profile  $F(Q)$  in Fig. 4A and B. We find that  $\chi(Q(t))$  closely traces  $F(Q)$  not only in the transition-state region ( $Q = Q^*$ ), but also in the whole  $Q$  range ( $Q_u \leq Q \leq Q_f$ ) it is defined (Pearson's correlation coefficient is  $R = 0.93$ ; corresponding results for the other systems are shown in Fig. 4C, D and in Fig. S10†). This is a nontrivial result since  $\chi(Q(t))$  is purely a dynamic quantity, and provides evidence demonstrating that the macrostate, thermodynamic cooperativity (brought about by the presence of the transition-state barrier) is connected to the microscopic, dynamic cooperativity (characterized by  $\chi(t)$ ).

The element-wise correlation  $\chi_{ij,kl}(t)$  at  $t = t^*$  (Fig. 2D) quantifies the strength of communication between individual amino-acid pairs. To facilitate its visual understanding, we present in Fig. 5A network representations of protein configurations during the transition path. In the upper section, the vertices (yellow circles) refer to amino acids and the edges (black lines) represent the formation of native amino-acid contacts; the folding process implies an increase in the number of black edges. In the lower section, the vertices and

edges are colored cyan when  $\chi_{ij,kl}(t) > 0.3$  for those amino acids in  $(i, j)$  and  $(k, l)$  pairs (this criterion was chosen since such large amino-acid correlation is barely observed in the unfolded state, as shown in Fig. S11†). The growth of the amino-acid correlation toward the transition state and its subsequent diminution are clearly visible in the network graphs.

Further insights into the amino-acid cooperativity, which are smeared in  $\chi(t)$  after summing over all the pairs, can be gained through the analysis of individual  $\chi_{ij,kl}(t)$  elements. For example,  $\chi(t)$  can be decomposed into the main-chain and side-chain contributions by examining which of the main-chain and side-chain contacts is mainly involved in the  $(i, j)$  and  $(k, l)$  amino-acid pairs, and we find that the magnitude of those contributions is comparable (Fig. S12†). The peak time  $t_{ij,kl}^*$  for each  $\chi_{ij,kl}(t)$  element can also be introduced. We observe that  $t_{ij,kl}^*$  values are dispersed around the average peak time  $t^*$  (Fig. S13†). Again, this is a dynamical analog of the thermodynamic transition in which residue-dependent variations were identified in the transition midpoint temperature.<sup>27</sup>

## Discussion

The fact that the folding transition state can be characterized as the state of maximum cooperativity is, to the best of our knowledge, a novel view. However, it is in fact quite natural once the existence of such cooperativity is cognized. This is because the protein configurations exhibiting the maximum internal correlations will be the ones with the lowest probability of



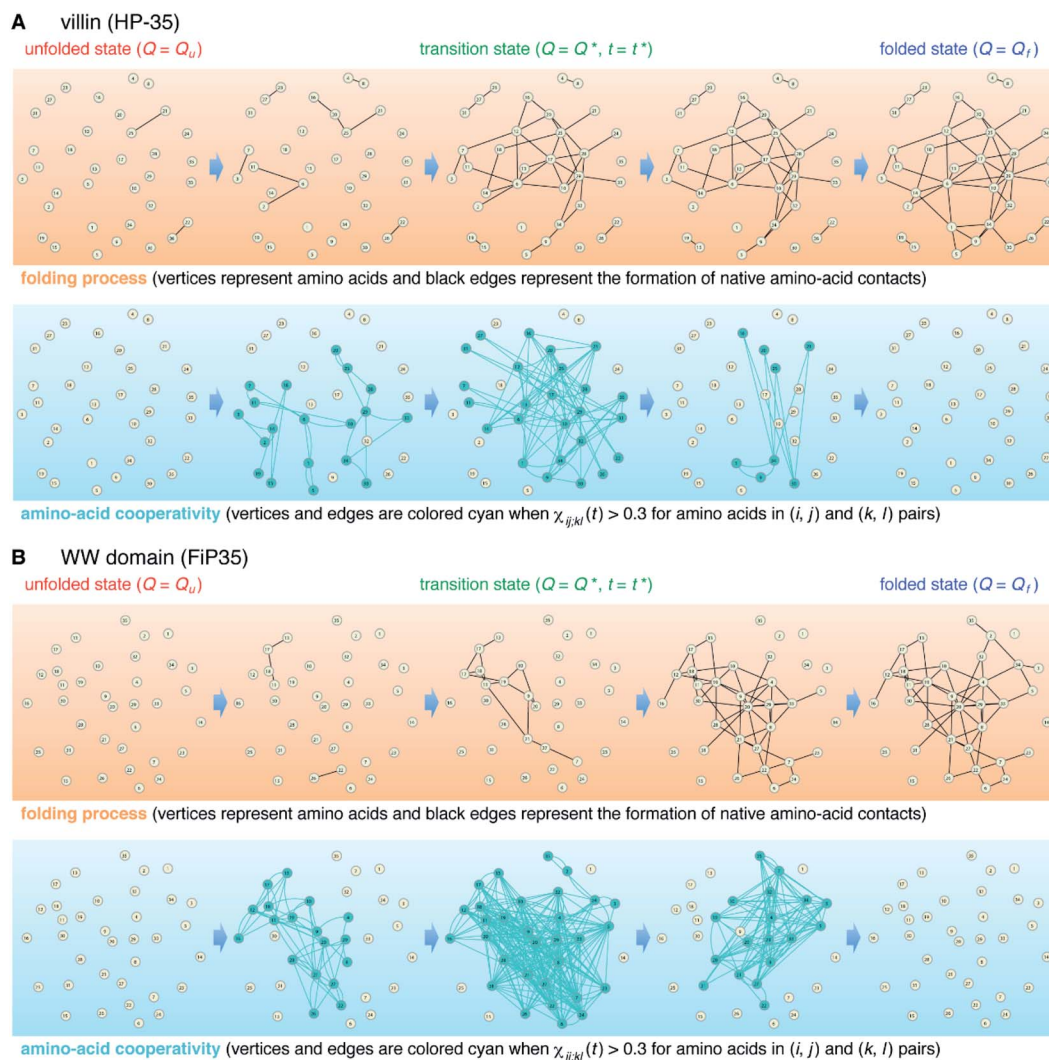


Fig. 5 Network representation of the folding transition path. (A) Upper section: network representation of the folding process of the villin headpiece subdomain (HP-35) in which vertices (yellow circles) represent amino acids and edges (black lines) represent the formation of native amino-acid contacts. Lower section: network representation of the time-dependent amino-acid cooperativity in which vertices and edges are colored cyan when  $\chi_{ijkl}(t) > 0.3$  for amino acids in  $(i, j)$  and  $(k, l)$  pairs. (B) Corresponding results for the WW domain (FIP35).

forming spontaneously. This new view in turn implies that the transition state barrier height should be an increasing function of the strength of the cooperativity. This is indeed the case as demonstrated in Fig. 4, which connects the microscopic cooperativity (characterized by  $\chi(t)$ ) and the macrostate two-state folding cooperativity (brought about by the presence of the transition-state barrier in  $F(Q)$ ).

Our current view of protein folding owes much to the funneled energy landscape perspective.<sup>28–30</sup> This perspective asserts that, in order to resolve Levinthal's paradox,<sup>31,32</sup> folding should not be a random conformation search; it must be energetically biased. However, the landscape perspective does not provide a clear picture of the transition-state barrier responsible for the emergence of cooperative two-state folding: the barrier is ascribed as being due to a “mismatch” between the energy gain and the entropy loss at the middle of the funneled landscape.<sup>33</sup> As we argued here, the folding transition state comes out naturally as the state of the maximum

microscopic cooperativity by realizing that the amino acid contact formation is not a random process, but occurs on a multiple-interaction basis. In this sense, the new view for the folding transition state represents an extension of the landscape perspective.

While native contacts are of primary interest in protein folding studies, non-native contacts can in principle contribute to the time-dependent amino-acid cooperativity discussed in the present work. This is because  $\chi_{ijkl}(t)$  defined in eqn (1) is invariant under the sign change,  $\sigma_{ij}(t) \rightarrow -\sigma_{ij}(t)$ :  $\sigma_{ij}(0)\sigma_{ij}(t)$  varies from 1 to  $-1$  not only when a contact absent at time  $t = 0$  ( $\sigma_{ij}(0) = -1$ ) is formed at time  $t$  ( $\sigma_{ij}(t) = 1$ ), but also when a contact present at time  $t = 0$  ( $\sigma_{ij}(0) = 1$ ) is broken at time  $t$  ( $\sigma_{ij}(t) = -1$ ). Therefore, if there exist a number of non-native contacts that are highly populated in the unfolded state but are broken during the folding process, they would contribute to  $\chi(t)$ . For the systems studied here, the existence of highly populated non-native contacts was not detected, and we cannot illustrate such



a possibility. Nevertheless, it is important to realize that the amino-acid cooperativity does not necessarily refer to the formation of contacts; the breaking of contacts can also occur cooperatively.

Finally, we present a possible experimental method for the detection of the cooperative contact formation of multiple amino acids by using a kind of Kirkwood relation that connects fluctuations and response. For this purpose, we introduce a two-point time correlation function  $F(t) = \langle q(t) \rangle$  with  $q(t) = (1/N) \sum_{(ij)} \sigma_{ij}(0) \sigma_{ij}(t)$ . This function describes how on

average the native contacts are being formed as the folding proceeds. The multipoint function  $\chi(t)$  capturing the time-dependent cooperativity can be written as the fluctuations around the average folding dynamics:  $\chi(t) = N \langle \delta q(t)^2 \rangle$  in which  $\delta q(t) = q(t) - \langle q(t) \rangle$ . Let us introduce a susceptibility defined as the response of  $F(t)$  to a perturbation  $\varphi$  (such as a change in temperature):  $\chi_\varphi(t) = \partial F(t) / \partial \varphi$ . It was demonstrated for dielectric and density fluctuations in glass-forming systems that  $\chi_\varphi(t)^2$  exhibits essentially the same dynamics as  $\chi(t)$ .<sup>34</sup> Since the average function  $F(t)$  is intimately related to the “shape” function of the transition path that is now experimentally accessible,<sup>35</sup> measuring  $\chi_\varphi(t)$  by varying experimental conditions will provide experimental evidence of the microscopic cooperativity in protein folding.

## Conclusions

Cooperativity in complex systems is typically described at a macrostate level, and its characterization in molecular terms has been very challenging. In the present work, we succeed in identifying time-dependent cooperativity among multiple amino acids concealed in the folding transition path, and argue how it might be connected to the macrostate cooperative behavior. The use of the multipoint correlation functions is essential in this regard, since a cooperative nature of fluctuating processes occurring at two distinct sites cannot be disclosed by conventional, two-point correlation functions. Since cooperativity pervades complex biological phenomena—the most notable example being allostery<sup>36</sup>—the multipoint correlation function approach will bring out novel microscopic insights into those complex processes.

## Author contributions

S.-H. C. and S. H. designed the research, conducted the research, and wrote the manuscript.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We acknowledge D. E. Shaw Research for providing us with the simulation trajectories. This work was supported by the

Samsung Science and Technology Foundation under Project Number SSTF-BA1401-52.

## References

- 1 K. A. Dill and J. L. MacCallum, *Science*, 2012, **338**, 1042–1046.
- 2 S. E. Ahnert, J. A. Marsh, H. Hernández, C. V. Robinson and S. A. Teichmann, *Science*, 2015, **350**, aaa2245.
- 3 A. C. Steven, W. Baumeister, L. N. Johnson and R. N. Perham, *Molecular Biology of Assemblies and Machines*, Garland Science, New York, 2016.
- 4 A. Whitty, *Nat. Chem. Biol.*, 2008, **4**, 435–439.
- 5 C. A. Hunter and H. L. Anderson, *Angew. Chem., Int. Ed.*, 2009, **48**, 7488–7499.
- 6 L. K. S. von Krbek, C. A. Schalley and P. Thordarson, *Chem. Soc. Rev.*, 2017, **46**, 2622–2637.
- 7 T. L. Hill, *Cooperativity Theory in Biochemistry, Steady State and Equilibrium Systems*, Springer-Verlag, New York, 1985.
- 8 P. L. Privalov, *Adv. Protein Chem.*, 1979, **33**, 167–241.
- 9 W. Feller, *Introduction to Probability Theory and Its Applications*, Wiley, New York, 1968, vol. 1.
- 10 A. Ben-Naim, *Cooperativity and Regulation in Biochemical Processes*, Kluwer Academic, New York, 2010.
- 11 M. Perutz, *Q. Rev. Biophys.*, 1989, **22**, 139–236.
- 12 H. S. Chung, K. McHale, J. M. Louis and W. A. Eaton, *Science*, 2012, **335**, 981–984.
- 13 H. S. Chung and W. A. Eaton, *Nature*, 2013, **502**, 685–688.
- 14 K. Neupane, D. A. N. Foster, D. R. Dee, H. Yu, F. Wang and M. T. Woodside, *Science*, 2016, **352**, 239–242.
- 15 P. L. Freddolino, C. B. Harrison, Y. Liu and K. Schulten, *Nat. Phys.*, 2010, **6**, 751–758.
- 16 T. J. Lane, D. Shukla, K. A. Beauchamp and V. S. Pande, *Curr. Opin. Struct. Biol.*, 2013, **23**, 58–65.
- 17 D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan and W. Wriggers, *Science*, 2010, **330**, 341–346.
- 18 K. Lindorff-Larsen, S. Piana, R. O. Dror and D. E. Shaw, *Science*, 2011, **334**, 517–520.
- 19 S. Piana, K. Lindorff-Larsen and D. E. Shaw, *Proc. Natl. Acad. Sci. U.S.A.*, 2012, **109**, 17845–17850.
- 20 S. Piana, K. Lindorff-Larsen and D. E. Shaw, *Proc. Natl. Acad. Sci. U.S.A.*, 2013, **110**, 5915–5920.
- 21 R. B. Best, G. Hummer and W. A. Eaton, *Proc. Natl. Acad. Sci. U.S.A.*, 2013, **110**, 17874–17879.
- 22 C. Donati, S. C. Glotzer and P. H. Pool, *Phys. Rev. Lett.*, 1999, **82**, 5064–5067.
- 23 C. Bennemann, C. Donati, J. Baschnagel and S. C. Glotzer, *Nature*, 1999, **399**, 246–249.
- 24 A. Cavagna, *Phys. Rep.*, 2009, **476**, 51–124.
- 25 D. K. Klimov and D. Thirumalai, *Folding Des.*, 1998, **3**, 127–139.
- 26 P. Kukic, Y. Pustovalova, C. Camilloni, S. Gianni, D. M. Korzhnev and M. Vendruscolo, *J. Am. Chem. Soc.*, 2017, **139**, 6899–6910.
- 27 E. P. O'Brien, B. R. Brooks and D. Thirumalai, *Biochemistry*, 2009, **48**, 3743–3754.



- 28 J. D. Bryngelson, J. N. Onuchic, N. D. Socci and P. G. Wolynes, *Proteins*, 1995, **21**, 167–195.
- 29 P. G. Wolynes, J. N. Onuchic and D. Thirumalai, *Science*, 1995, **267**, 1619–1620.
- 30 K. A. Dill and H. S. Chan, *Nat. Struct. Biol.*, 1997, **4**, 10–19.
- 31 C. Levinthal, *Mössbauer Spectroscopy in Biological Systems Proceedings*, 1969, vol. 67, pp. 22–24.
- 32 R. Zwanzig, A. Szabo and B. Bagchi, *Proc. Natl. Acad. Sci. U.S.A.*, 1992, **89**, 20–22.
- 33 M. Oliveberg and P. G. Wolynes, *Q. Rev. Biophys.*, 2005, **38**, 245–288.
- 34 L. Berthier, G. Biroli, J.-P. Bouchaud, L. Cipelletti, D. E. Masri, D. L'Hôte, F. Ladieu and M. Pierno, *Science*, 2005, **310**, 1797–1800.
- 35 N. Q. Hoffer, K. Neupane, A. G. T. Pyo and M. T. Woodside, *Proc. Natl. Acad. Sci. U.S.A.*, 2019, **116**, 8125–8130.
- 36 J.-P. Changuex, *Nat. Rev. Mol. Cell Biol.*, 2013, **14**, 819–829.

