May the driving force be with you — whatever it is

John Cavanagh and Mikael Akke

Changes in the atomic coordinate fluctuations contribute to the entropy of biomolecular processes such as complex formation. Characterizations of such changes in proteins reveal that the response may be dramatically different between the backbone and the side chains, and further resolve enthalpy-entropy compensation at the molecular level.

Countless biological processes, including macromolecular interactions and protein folding, occur with concomitant and often compensating changes in enthalpy and entropy[1]. This common phenomenon reflects the opposing effects of the fundamental tendency of a system to seek out the lowest energy levels and, at the same time, populate as many levels as possible given the available thermal energy. A thorough understanding of biological processes requires a detailed description on the molecular level of the driving thermodynamic forces involved. Inspection of high resolution structures of protein complexes often reveals a large number of favorable interactions (for example, hydrogen bonds, 'salt bridges', and close packing of hydrophobic residues) that offer intuitively appealing explanations for the enthalpic contribution to binding. In contrast, the entropic contributions are impossible to assess directly from static molecular coordinates.

On page 72 of this issue of *Nature Structural Biology*, Wand and coworkers[2] use NMR relaxation methods to characterize the change in dynamic behavior of calcium-loaded calmodulin resulting from the binding of a peptide from the calmodulin-binding domain of smooth muscle myosin light chain kinase. Their investigations reveal that while the dynamic behavior of the protein backbone is essentially unaltered, the motions of several side chains are dramatically affected. Upon target binding, methyl-bearing side chains throughout the protein exhibit slightly decreased mobility on average, while certain methionines in and around the binding site become much more rigid. The authors use their results to provide a description of compensatory changes in enthalpy and entropy associated with the formation of a protein–protein interface. This work is noteworthy because it delineates such enthalpy-entropy compensation at the level of individual side chain groups.

All molecular processes are governed by the difference in Gibbs free energy between the initial and final states. This difference can be divided further into enthalpic and entropic contributions as stated by the familiar expression $\Delta G = \Delta H - T\Delta S$. Calorimetric methods, such as isothermal titration calorimetry, provide an experimental means to determine the changes in each of the component state function[3]. Knowing the relative importance of enthalpic and entropic contributions is a first step towards characterizing the dominant driving forces of a given biomolecular recognition process. However, to understand the molecular basis for biological function, these thermodynamic parameters must be related to the detailed interactions between the molecules, as well as to any changes in solvation, structure and structural fluctuations of the interacting molecules that result from the binding event. To this extent, atomic resolution structural information is essential.

In principle, the energetics of ligand binding can be assessed from high resolution structures and model energy functions, but full thermodynamic treatments are still computationally demanding[4]. Mutational analyses targeting recognition surfaces have enabled the identification of 'hot spots' on proteins that dominate the binding energy[5]. However, these types of studies have not revealed why certain residues appear particularly important for recognition. Because proteins are inherently dynamic, analyses of time- or ensemble-averaged structures do not provide a complete description of protein behavior and function. Complementary work that characterizes the dynamic properties of protein–protein interaction surfaces has underscored the potential importance of protein flexibility in governing affinity and specificity[6,7]. In this respect, significant advances have been made toward

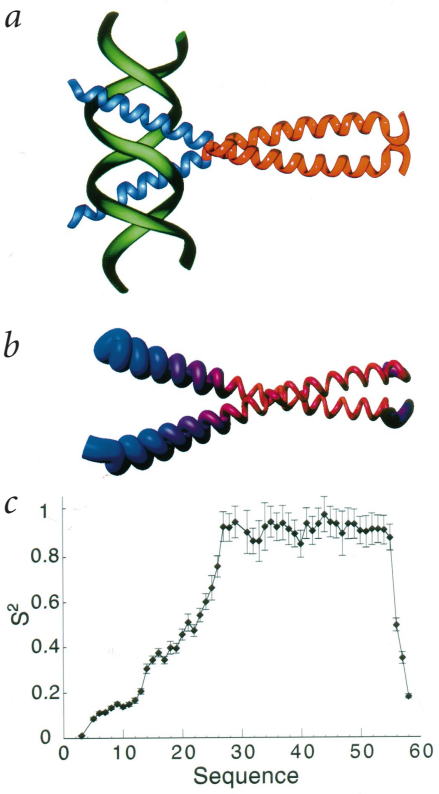

Fig. 1 ***a***, Ribbon diagram of GCN4-58 complexed with DNA. The section in blue is the basic DNA-binding domain and the section in red (residues 27–58) is the leucine zipper dimerization domain (residues 3–26). ***b***, Representation of the backbone motions of GCN4-58 at 310 K in the absence of DNA. The width and color are representative of motions (picosecond–nanosecond regime), with the thicker and deeper purple regions indicating the highest degree of disorder. ***c***, Approximate backbone order parameters derived from spectral density mapping of GCN4-58 plotted against amino acid residue number. The amino acid sequence of (*c*) is aligned with the structure shown in (*b*). It can be seen that the order parameters begin low, indicating a highly disordered structure, and gradually move to a value of $S^2 \sim 0.9$ at residue 27. Order parameters of ~0.86 are typical of well ordered residues in helices in globular proteins. There is a clear partition between the DNA-binding domain and the dimerization domain.

understanding the subtle relationships between biomolecular recognition and the conformational space of proteins.

Another very important question concerns the decomposition of the entropy changes into contributions from solvation and from the configurational entropy of the macromolecules themselves. Semi-empirical models based on changes in the solvent accessible surface area have been proposed that provide estimates of the relative importance of these two contributions[8]. Ideally, one would like to obtain experimental data that report directly on the configurational entropy of all components of the system: the protein, the ligand, the solvent water, and any co-solutes such as buffer ions. In the last few years, NMR relaxation experiments have offered some initial, albeit partial, data of this kind that promise to provide important insights into the entropy of biomolecular processes.

NMR spectroscopic investigations can provide specific information about both the amplitudes and time scales of intramolecular motions. Heteronuclear (^{13}C and ^{15}N) spin relaxation is caused by time-dependent magnetic fields, the most dominant of which are usually those generated by the magnetic dipoles of the covalently attached proton(s). The time dependence of the magnetic field arises as a consequence of intramolecular motions and rotational tumbling of the entire molecule. A suite of NMR techniques is now available that allows one to investigate and distinguish intramolecular motions in different 'time windows'; in solution NMR, these windows are commonly divided into the picosecond–nanosecond, microsecond–millisecond and second time scales[9,10]. Important results have been compiled over the last five to six years showing that motions on each of these time scales may be of fundamental biological significance[11–14].

NMR relaxation experiments sensitive to the picosecond–nanosecond time scale probe the amplitudes of bond vector fluctuations, which can be expressed in terms of an order parameter[15,16], commonly reported as the square of its value (S^2), that varies between 0 for unrestricted internal motion to 1 for no internal motion. The order parameter can be directly related to the entropy of the corresponding molecular degrees of freedom, as follows. The order parameter determined using NMR spin relaxation experiments in solution provides information on the equilibrium probability distribution, $P(\Omega)$, of the bond vector orientation on a time scale shorter than overall rotational diffusion, which typically is on the order of 10 ns[9,10]:

$$S^2 = \frac{4\pi}{5} \sum_{m=-2}^{2} \left| \sum_{\Omega} P(\Omega) Y_{2,m}(\Omega) \right|^2 \qquad (1)$$

where $Y_{2,m}(\Omega)$ is a second order spherical harmonic function. Using model energy functions, $E(\Omega)$, the probability distribution measured through the order parameter can be obtained and related to the partition function, Q, of the system in question[17–20]:

$$P(\Omega) = \frac{\exp[-E(\Omega)/(kT)]}{\sum_{\Omega} \exp[-E(\Omega)/(kT)]} = \exp[-E(\Omega)/(kT)]/Q \qquad (2)$$

which in turn enables calculation of thermodynamic parameters, such as the free energy (G) and entropy (S)[17–20]:

$$G = kT\ln Q$$
$$S = k \sum_{\Omega} P(\Omega) \ln P(\Omega) \qquad (3)$$

where k is the Boltzmann constant and T is the temperature. Reassuringly, the various physically reasonable models of $E(\Omega)$ that have been applied in this context yield similar results when the difference in thermodynamic parameters between two different protein states (for example, free and bound) is evaluated.

Currently, this approach has several limitations: (i) the order parameter is sensitive only to motions on a time scale shorter than overall rotational diffusion; (ii) the order parameter is sensitive only to motions that reorient the relevant bond vector; (iii) present methods do not account for possible correlations between motions of different bond vectors; and (iv) only a subset of bond vectors are characterized. It should be noted that even the slowest vibrational modes of proteins tend to fall within the time window dictated by the overall rotational diffusion, implying that limitation (i) may not be severe[21]. Also, the errors introduced by (ii) and (iii) tend to cancel each another to some extent.

The interpretation of NMR relaxation experiments is also associated with potential pitfalls that appear at an earlier stage of the analysis. Commonly, the calculation of order parameters is based on only three different relaxation parameters: the longitudinal relaxation rate (R_1), the transverse relaxation rate (R_2) and the heteronuclear NOE. Notably, the transverse relaxation rate is sensitive to motions on both the picosecond–nanosecond and the microsecond–millisecond time scales, but only the former is related to the order parameter. Fluctuations on the former time scale decrease the relaxation rate, whereas fluctuations on the latter time scale (known as 'chemical exchange') increase the relaxation rate, so that these two effects tend to offset each other. Accurate extraction of order parameters is therefore critically dependent on the separation of these two opposing effects[22], which can be achieved using recent experiments that measure so called cross-correlation relaxation rates[23,24], the magnetic field dependence of R_2 (ref. 25), or rotating frame relaxation dispersion[26,27]. Clearly, the interpretation of NMR relaxation data requires careful assessment of these different contributions. All in all, while the approach for estimating conformational entropy from NMR order parameters offers only a partial picture, and is sensitive to experimental artifacts, it does provide a unique means for probing the conformational entropy on the molecular level.

Several independent studies using this approach suggest that changes in configurational entropy may make important contributions to binding processes (for reviews see refs 11–14). In the previous issue of *Nature Structural Biology*, Stone and coworkers published a report[28,29], suggesting that the conformational entropy of the protein backbone may increase globally upon binding of a small, hydrophobic ligand. Calculated estimates of the relative contributions to the binding free energy indicate that the increase in backbone conformational entropy is significant and required for binding[28]. Another representative and strikingly illustrative example is the recent work by Palmer and coworkers[30] on the dynamics of the basic leucine zipper domain of the dimeric yeast transcription factor GCN4 (GCN4-58) as it relates to DNA binding (Fig. 1*a*). GCN4-58 comprises two regions: the basic DNA-binding region and the leucine zipper region involved in dimerization. The basic region (residues 3–26) exhibits order parameters indicative of extensive conformational disorder. In stark contrast, the order parameters of the leucine zipper region (residues 27–58) indicate a highly organized structure (Fig. 1*b,c*), with values typical of helical residues in globular proteins ($S^2 = 0.85$). They calculated that the contribution from conformational entropy of the protein backbone to the thermodynamics of DNA binding was $\Delta S = 0.6$ kJ K^{-1} mol^{-1}. Remarkably, theoretical predictions based on experimentally determined calorimetric measurements for the same system predicted $\Delta S = 0.5$ kJ K^{-1} mol^{-1}. The good agreement between such different approaches in cal-

culating the conformational entropy supports the validity of each method.

With all this in mind, the paper of Wand and coworkers[2] is particularly notable since it clearly illustrates several important points. Most importantly, a definite separation is seen between backbone and side chain response upon substrate binding to calcium-loaded calmodulin. While there is little change in the motional characteristics of the backbone, both increases and decreases in fluctuations are observed among the side chains. In particular, two functionally important methionine residues Met 72 and Met 124 show large increases in order parameter of $\Delta S^2 > 0.5$. The implication is that the significantly reduced flexibility of these residues enhances their interactions with the target peptide which in turn compensates for the reduction in the local entropy upon peptide binding. This result suggests that localized regions within calmodulin show enthalpy-entropy compensation. At the same time, calmodulin exhibits extensive redistribution of side chain entropy upon binding of the target peptide, such that certain side chains remote from the binding site experience increased entropy upon binding, thereby partly offsetting the entropic penalty paid by the residues that interact tightly with the substrate. The authors discuss these results with respect to different modes of protein recognition in light of previous studies on an SH2 domain[6], where smaller and more localized order parameter changes upon peptide binding were observed than in the case of calmodulin. Wand and coworkers[2] also interpret the order parameters to obtain an approximate measure of the contribution from side chain entropy in calmodulin to the binding free energy. Their estimate is in qualitative agreement with results obtained from calorimetric measurements[31]. This finding complements the work by the Palmer[30] and Stone[28] groups.

The above NMR results all relate to the conformational entropy of the protein, but say nothing about the entropy of hydration. However, recent experiments have shown that NMR can provide information relevant to this problem as well. Magnetic field dependent NMR relaxation studies of protein hydration have been used to characterize the order parameters of water molecules that are bound to specific sites in proteins[20]. As a particularly elegant application of the method outlined by the equations above, the order parameters can be directly related to the rotational entropy of the water molecules. These results suggest that the change in entropy may in some cases favor binding of water molecules to proteins — in direct contrast to commonly accepted views.

Taken together, the results obtained to date are encouraging and indicate that further development in this area should be worthwhile.

John Cavanagh is in the Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, USA. Mikael Akke is in the Division of Physical Chemistry 2, Lund University, P.O. Box 124, S-221 00 Lund, Sweden. Correspondence should be addressed to J.C. email: johnc@purdue.edu or M.A. email: mikael.akke@fkem2.lth.se.

1. Dunitz, J.D. *Chem. Biol.* **2**, 709–712 (1995).
2. Lee, A.L., Kinnear, S.A. & Wand, A.J. *Nature Struct. Biol.* **7**, 72–77 (2000).
3. Ladbury, J.E. & Chowdhry, B.Z. *Chem. Biol.* **3**, 791–801 (1996).
4. Levy, R.M. & Gallicchio, E. *Ann. Rev. Phys. Chem.* **49**, 531–567 (1998).
5. Clackson, T. & Wells, J.A. *Science* **267**, 383–386 (1995).
6. Kay, L.E., Muhandiram, D.R., Farrow, N.A., Aubin, Y. & Forman-Kay, J.D. *Biochemistry* **35**, 361–368 (1996).
7. Kay, L.E., Muhandiram, D.R., Wolf, G., Shoelson, S.E. & Forman-Kay, J.D. *Nature Struct. Biol.* **5**, 156–163 (1998).
8. Spolar, R.S. & Record, M.T. *Science* **263**, 777–784 (1994).
9. Palmer, A.G., Williams, J. & McDermott, A. *J. Phys. Chem.* **100**, 13293–13310 (1996).
10. Fischer, M.W.F., Majumdar, A. & Zuiderweg, E.R.P. *Progr. NMR Spectrosc.* **33**, 207–272 (1998).
11. Kay, L.E. *Nature Struct. Biol.* **5 suppl**, 513–517 (1998).
12. Feher, V.A. & Cavanagh, J. *Nature* **400**, 289–293 (1999).
13. Palmer, A.G. *Curr. Opin. Struct. Biol.* **7**, 732–737 (1997).
14. Roberts, G.C.K. *Curr. Opin. Biotechn.* **10**, 42–47 (1999).
15. Halle, B. & Wennerström, H. *J. Chem. Phys.* **75**, 1928–1943 (1981).
16. Lipari, G. & Szabo, A. *J. Am. Chem. Soc.* **104**, 4546–4559 (1982).
17. Akke, M., Brüschweiler, R. & Palmer, A.G. *J. Am. Chem. Soc.* **115**, 9832–9833 (1993).
18. Yang, D. & Kay, L.E. *J. Mol. Biol.* **263**, 369–382 (1996).
19. Li, Z., Raychaudhuri, S. & Wand, A.J. *Protein Sci.* **5**, 2647–2650 (1996).
20. Denisov, V.P., Venu, K., Peters, J., Hörlein, H.D. & Halle, B. *J. Phys. Chem.* **101**, 9380–9389 (1997).
21. Brüschweiler, R. *J. Chem. Phys.* **102**, 3396–3403 (1995).
22. Evenäs, J., Forsén, S., Malmendal, A. & Akke, M. *J. Mol. Biol.* **289**, 603–617 (1999).
23. Kroenke, C.D., Loria, J.P., Lee, L.K., Rance, M. & Palmer, A.G. *J. Am. Chem. Soc.* **120**, 7905–7915 (1998).
24. Felli, I.C., Desvaux, H. & Bodenhausen, G. *J. Biomol. NMR* **12**, 509–521 (1998).
25. Phan, I.Q.H., Boyd, J. & Campbell, I.D. *JBNMR* **8**, 369–378 (1996).
26. Meiboom, S. *J. Chem. Phys.* **34**, 375–388 (1961).
27. Deverell, C., Morgan, R.E. & Strange, J.H. *Mol. Phys.* **18**, 553–559 (1970).
28. Zidek, L., Novotny, M.V. & Stone, M.J. *Nature Struct. Biol.* **6**, 1118–1121 (1999).
29. Forman-Kay, J.D. *Nature Struct. Biol.* **6**, 1086–1087 (1999).
30. Bracken, C., Carr, P.A., Cavanagh, J. & Palmer, A.G. *J. Mol. Biol.* **285**, 2133–2146 (1999).
31. Wintrode, P.L. & Privalov, P.L. *J. Mol. Biol.* **266**, 1050–1062 (1997).