Title: Structural origins of protein conformational entropy

Authors: José A. Caro^{1,2}, Kathleen G. Valentine², A. Joshua Wand^{1,2*}

Affiliations:

¹Department of Biochemistry & Biophysics, College of Agriculture and Life Sciences, Texas A&M University, College Station, Texas 77845

²Department of Biochemistry & Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104

* email: wand@tamu.edu

Abstract

The thermodynamics of molecular recognition by proteins is a central determinant of complex biochemistry. For over a half-century detailed cryogenic structures have provided deep insight into the energetic contributions to ligand binding by proteins¹. More recently, a dynamical proxy based on NMR-relaxation methods has revealed an unexpected richness in the contributions of conformational entropy to the thermodynamics of ligand binding^{2,3,4,5}. There remains, however, a discomfiting absence of an understanding of the structural origins of fast internal motion and the conformational entropy that this motion represents. Here we report the pressure-dependence of fast internal motion within the ribonuclease barnase and its complex with the protein barstar. Distinctive clustering of the pressure sensitivity correlates with the presence of small packing defects or voids surrounding affected side chains. Prompted by this observation, we performed an analysis of the voids surrounding over 2,500 methyl-bearing side chains having experimentally determined order parameters. We find that changes in unoccupied volume as small as a single water molecule surrounding buried side chains greatly affects motion on the subnanosecond timescale. The discovered relationship begins to permit construction of a united view of the relationship between changes in the internal energy, as exposed by detailed structural analysis, and the conformational entropy, as represented by fast internal motion, in the thermodynamics of protein function. [214/200]

Body

The change in the Gibbs free energy underlying molecular recognition and other complex protein functions such as allosteric regulation has, in principle, contributions from both entropy and enthalpy. The latter is comprised of the internal energy and a pressure-volume work term. Detailed analysis of static low-temperature structural models has historically provided great insight into the internal energy and has promoted significant advances in understanding protein functions such as ligand binding through simulation and theory⁶. Nevertheless, the origins of protein conformational entropy and its contribution to functions such as allostery remains less well defined⁵. Measurement of equilibrium fluctuations offers a powerful way to describe transitions between and occupancy of states that cannot be observed with classical methods of structural biology and NMR relaxation has proven particularly useful in this regard^{4,7,8}. Over the past two decades, numerous studies of fast internal side chain motion by NMR methods, particularly that of methyl-bearing amino acids, have revealed an unexpected complexity without distinguishing structural correlates⁹.

Here we take advantage of the fact that the Gibbs free energy change associated with a change in state contains a pressure-volume work term. Volume changes represent the natural variable and application of pressure can illuminate otherwise unobservable details of the thermodynamics of protein functions such as ligand binding and allostery. Protein molecules respond, both dynamically and structurally, to pressure in a complicated way. Pressure can compress proteins^{10,11,12}, remodel active sites¹² and facilitate excursions to higher lying^{13,14}, locally unfolded^{15,16,17,18} or globally unfolded states^{19,20}, thereby revealing various aspects of the ensemble nature of proteins. Here, we use high pressure NMR²¹ relaxation to probe fast internal motion of methyl-bearing side chains in the small enzyme barnase and use this motion as a proxy for conformational entropy (S_{conf}).

Barnase-barstar is one of the strongest protein-protein interactions known in biology. Its fM affinity derives from large enthalpic contributions at the interface²². Individual entropic contributions sum to yield a negligible (~zero) contribution to the binding free energy²². At ambient pressure, we find that complexation is accompanied by an overall rigidification of the

methyl-bearing side chains of barnase (Fig. 1; Supplementary Table 1). We quantify the disorder of the methyl symmetry axis in terms of the Lipari-Szabo squared generalized order parameter (O^2_{axis})²³ obtained using deuterium NMR relaxation methods²⁴. The O^2_{axis} can range from a value of one, corresponding to complete rigidity within the molecular frame, to zero, which effectively corresponds to isotropic disorder. Importantly, only motion faster than the overall molecular reorientation of the protein contributes to O^2_{axis} . The dynamical proxy for conformational entropy⁴ indicates that the overall rigidification corresponds to an unfavorable contribution (ΔS_{conf}) to the binding free energy amounting to $+11.7 \pm 1.2$ kJ/mol at room temperature.

Application of high hydrostatic pressure on free barnase yields an unexpected clustering of changes in motion (ΔO^2_{axis}) into two spatial regions, one that rigidifies with pressure and one that activates dynamically (Fig. 1; Supplementary Table 1). Region I, which becomes more rigid with pressure, is defined by twenty-one methyl-bearing side chains that are largely localized to the N-terminal domain of the protein. Eight of these side chains are fully buried. Region II, which becomes more dynamic with applied pressure, is comprised of seventeen methyl-bearing side chains in the C-terminal domain and eleven of these are fully buried. Nine methyl-bearing side chains are outside of these regions. All methyl probes are 7 Å or more from the barnase-barstar interface, which is highly polar and extensively hydrated²⁵.

A thermodynamic cycle from free barnase was created with barnase either bound to its inhibitor barstar, subjected to high hydrostatic pressure (3 kbar) or both (Fig. 1). At 3 kbar, binding of barnase to barstar results in an opposite response from the N- & C-terminal groups of side chains. Motion in Region I is activated by elevated pressure, which is opposite to the response of free barnase. Region II rigidifies upon barnase binding barstar, both at ambient and high pressure (Fig. 1; Supplementary Table 1). Application of high pressure to the barnase-barstar complex leads to a general increase in the internal motion of barnase, with the largest change centered in Region II. As might be expected, the pressure sensitivity of a side chain's motion is reduced as the O^2_{axis} approaches the rigid limit of one at ambient pressure. Of the four states of barnase examined, the complexed state at ambient pressure is the most rigid ($\langle O^2_{\text{axis}} \rangle = 0.693$) (Supplementary Table 1). The heterogeneous response of side chain dynamics to pressure is in

stark contrast with other metrics examined. For example, fast backbone motions are generally suppressed in response to pressure and without apparent grouping to regions I & II (Table S1). The heterogeneity and complexity in the dynamical character of the protein is also not apparent from the more usual tactic of characterizing the pressure dependence of NMR chemical shifts and peak intensities (Supplementary Fig 2).

The localized response of motion to pressure is not easily explained by $\langle O^2_{\text{axis}} \rangle$ values at ambient pressure (Region I: $\langle O^2_{\text{axis}} \rangle$ ($\text{var} O^2_{\text{axis}}$) = 0.721 (0.031), $n=16$; Region II: $\langle O^2_{\text{axis}} \rangle$ ($\text{var} O^2_{\text{axis}}$) = 0.621 (0.030), $n=12$). Elevation to 3 kbar, results in $\langle \Delta O^2_{\text{axis}} \rangle$ ($\text{var} \Delta O^2_{\text{axis}}$) of +0.040 (0.002) and -0.053 (0.003) for Regions I and II, respectively. The effect of pressure is fundamentally related to changes in system volume. To examine potential contributions to the pressure sensitivity by the protein itself, we carried out a fine-grained volumetric analysis of the crystal structure (see Methods). Focusing on methyl-bearing side chains without solvent accessible surface area, we find Region I side chains in the ambient pressure structure, on average, have 35 Å³ more unoccupied volume surrounding the side chain than those of Region II (111 ± 23 Å³ and 78 ± 33 Å³, respectively; $p < 0.022$). Compression of voids explains the rigidification by pressure observed in Region I. In contrast, a more densely packed Region II may not be able to further compress and will respond to pressure through other mechanisms that decrease the system volume, such as local structural transitions or changes in hydration. These initial observations prompted us to compile a database of experimentally determined methyl symmetry axis order parameters and investigate the influence of surrounding void volume more generally.

The technical challenges of bridging crystallography and NMR relaxation are numerous. On the NMR side, relaxation datasets are sometimes compromised by contamination of relaxation from nearby nuclei, inadequate tumbling analysis, incomplete depositions, misassignments, and inclusion of data from ill-resolved resonances. Crystallography datasets often suffer from absent density maps, incorrect or incomplete modeling, low resolution, extreme experimental conditions (cryogenic temperatures), ambiguity in solvent and cosolvent identification, and poor representation of solvent and alternative conformations. There are approximately 2,568 experimentally determined O^2_{axis} values of proteins with cryogenic crystal structures determined

to high-resolution. We performed a fine-grained volumetric analysis of 42 crystal structures with corresponding 1027 measured O_{axis}^2 values derived from fully buried side chains that meet the quality criteria outlined in the methods section. All of these methyl-bearing side chains were observed to have surrounding void volumes less than $180 \text{ \AA}^3$. An intriguing trend reinforces the notion that there exists a lower limit on the O_{axis}^2 values imposed by the available surrounding unoccupied volume (Fig. 2). Side chains surrounded by less than $70 \text{ \AA}^3$ of void volume retain O_{axis}^2 values above 0.5 with the large majority (75%) retaining high values above 0.75. Surrounding void volumes between 70 and $140 \text{ \AA}^3$ correlate with the entire range of possible O_{axis}^2 values. The range of void volumes over which low O_{axis}^2 values become populated represents a window of about $50 \text{ \AA}^3$, which is a remarkably small fraction of the total unoccupied volume in proteins. Indeed, such a change in void volume is on the order of that made accessible by a single alternative rotamer or the removal of a single water molecule (Fig. 3; see text below).

What defines O_{axis}^2 values above the identified lower limit of void volume clearly has a complicated origin. For example, many side chains with $>100 \text{ \AA}^3$ of surrounding void volume can still display very high O_{axis}^2 values. To identify the reasons for this behavior, we analyzed the role of three factors: the dependence on side chain topology (number of “soft” torsion angles); the presence of crystallographically defined solvent; and the presence of alternative side chain conformations. The distribution of O_{axis}^2 values (Fig 3a) reflects the long-appreciated feature that the *accessible* range depends on the number of torsion angles available in the context of a rigid backbone though the origin of the distribution within that range has remained a mystery^{9,26}. It is also apparent that it is important to consider the presence of crystallographically defined solvent water molecules. For example, Leu-42 in barnase abuts two deeply buried water molecules, present in all 35 deposited crystal structures of barnase, which appreciably reduce the volume that would be otherwise available to the side chain (Fig 3b; $83 \text{ \AA}^3$ versus $95 \text{ \AA}^3$). Analysis of intermolecular dipolar relaxation from hydrogen-bonding partners (Trp-35 indole and Ile-51) to the two waters confirm their rigid presence in solution, characterized by long-residence times and restricted motion⁴. In general, underrepresentation of solvent molecules in crystal structures will result in overestimation of the void volume of surrounding side chains. Consideration of hydration water, which experience a broad range of residence times and motion while bound to

the protein²⁷, will likely be particularly important to understanding the influences of specific solvation.

Alternative conformations can also influence the effective void volume available. For example, including the alternative conformation observed for Ile-96 of barnase results in a large increase of its available volume, but also decreases the available volume of nearby side chains Ala-11, Leu-14, and Ile-88 (black arrows, Fig. 3c). The potentially large impact of alternative conformations is further illustrated by the multi-conformer model for the room temperature crystal structure of dihydrofolate reductase (DHFR)²⁸. Including multiple side chain rotamers to represent the electron density results in an overall decrease of apparent void volume for the more rigid side chains and an increase for the more dynamic side chains (Fig. 3d). In other words, the underrepresentation of alternative conformations in cryogenic crystal structures leads to less accurate calculation of void volumes and deteriorates the correspondence with side chain dynamics. This implies that the broad distribution of low (<0.5) O_{axis}^2 values with surrounding void volume (Fig. 2a) is likely due to inadequate representation of alternative conformations in single-conformer cryogenic crystal structures.

The remarkable sensitivity of protein conformational entropy to packing density has been predicted by simulation, albeit involving somewhat greater changes than identified here²⁹. Indeed, analysis of hard-sphere liquids also indicates a dependence of dynamics on available free-space³⁰ and reinforces the notion derived from the pressure sensitivity of fast aromatic ring motion that proteins interiors are quite liquid-like³¹. A heterogeneous response of protein domains to pressure was also reflected in the sole previous study of the pressure sensitivity of fast methyl-bearing side chain motion³². The influence of elevated hydrostatic pressure on the internal motion (conformational entropy) in the barnase-barstar complex, the first complex studied by these methods, is striking. High affinity binding selectively alters the response of barnase to pressure. Region I rigidifies as free barnase is compressed, indicating that there is conformational heterogeneity involving less efficiently packed alternative conformation(s). Interestingly, this region corresponds to a late-folding intermediate³³. In contrast, pressure favors increased disorder on the sub-nanosecond timescale in Region II. Very localized spatial clustering

of the response of fast motion to pressure was observed in ubiquitin³², but the degree of structural segregation in barnase is striking. The femtomolar affinity of the barnase-barstar complex exists despite a penalty by $-T\Delta S_{\text{conf}}$ of +11.7 kJ/mol. But at high pressure, the *overall* change in side chain dynamics is zero and binding occurs with no conformational entropy penalty. The response of side chains to pressure is consistent with an important role of conformational dynamics in the adaptation of protein function to extreme environments²³. Furthermore, these results make clear that changes in both the magnitude and the sign of regional contributions of conformational entropy to the thermodynamics of protein function are possible. In a very real sense, the establishment of a connection between extremely subtle structural features (imperfect packing) and conformational entropy provides a mechanism for Nature to install the 35 year-old speculation of allostery without (detectable) structural change³⁴.

Methods

Sample preparation

pET-DUET expression plasmids containing the genes for barnase and barstar under the control of their own T7 promoter were obtained from GenScript Biotech Corporation (Piscataway, NJ, USA). An N-terminal 6xHis-tag followed by a Factor Xa cleavage site (MGSSHHHHHSQAPIEGR) was added to barnase while barstar remained untagged. Expression was carried out in BL21-(DE3) E. coli cells. Barstar expressed and purified with the N-terminal Met residue present. NMR relaxation samples were prepared largely as described elsewhere⁴. Deuterium and ¹⁵N relaxation experiments of the free proteins were performed on a 1:2 mixture of uniformly ¹⁵N-labeled protein and uniformly ¹³C-labeled protein expressed in 60% D₂O media to generate the ¹³CH₂D isotopomer. The complex was studied by combining ¹⁵N-labeled protein (barnase or barstar) with ¹³CH₂D-labeled binding partner (barstar or barnase). Prochiral methyl assignments samples were expressed during growth on 10% ¹³C₆-glucose and 90% unlabeled glucose and uniform ¹⁵N labeling³⁵.

The barnase-barstar complex was isolated by Ni-NTA affinity chromatography and the complex dissociated with 6 M guanidine HCl, pH 7.9. Barstar was collected in the flow-through (20 mL) and refolded by dilution into 1 L of water. Refolded barstar was further purified by DEAE ion

exchange column that included a wash with 25 mM imidazole, pH 7.9, 10 mM KCl and elution with 500 mM NaCl, spin-concentrated (3 kDa cutoff) and further purified by size exclusion chromatography on Superdex SEC-75 equilibrated with 25 mM imidazole, pH 7.9, 10 mM KCl. Barnase was eluted from the Ni-NTA column with 500 mM imidazole, spin-concentrated (3 kDa cutoff), and buffer exchanged into 25 mM imidazole, pH 6.2, 10 mM KCl and 5 mM CaCl₂. The His-tag was cleaved with four µg of Factor Xa per mg of barnase added and mixed overnight at room temperature. The solution was passed through a 1 mL Ni-NTA column coupled to a SEC-75 in 50 mM imidazole, pH 7.9, 50 mM KCl, spin-concentrated and buffer exchanged to 25 mM imidazole, pH 6.2, 10 mM KCl. NMR experiments were performed with samples prepared in 25 mM imidazole, pH 6.2, 10 mM KCl, 5% D₂O, and 0.02% NaN₃ (w/v). Samples were stable about 1 month for the free proteins and several months for the complex at 35 °C.

NMR assignment and relaxation of free and bound barnase

All experiments were carried out at 35 °C. Assignment experiments were done on a uniformly ¹³C, ¹⁵N-labeled sample, with only one protein in the complex labeled to reduce spectral crowding. Non-uniform sampling was used extensively for triple resonance assignment spectra³⁶. Assignments were mapped to high pressure by collecting ¹³C and ¹⁵N HSQC spectra every 500 bar. These experiments were collected either at 500 or 600 MHz. Carbon and nitrogen HSQC spectra were collected at 1, 50, 500, 1000, 1500, 2000, 2500, and 3000 bar with a waiting period of 1h between each. Spectra were collected during ramp up and ramp down of pressure, with no detectable difference observed between them. Chemical shift analysis utilized the gyromagnetic ratio weighted change in chemical shift of ¹H and ¹⁵N (or ¹³C) of bonded atoms resolved in two-dimensional correlation spectra.

Resonance peak heights and volumes were obtained using NMRFAM-SPARKY³⁷. Only fully resolved cross peaks that fitted well to a Lorentzian lineshape to provide intensities were included. Peaks were integrated using the fitted Lorentzian function. The pressure dependence of gyromagnetic weighted chemical shifts were analyzed using a second-order Taylor expansion³⁸. The correlation between second and first order coefficients was fitted with a linear regression with intercepts (6.5 ± 0.8) and $(-16 \pm 6) \times 10^{-10}$ ppm bar⁻², and slopes (-7.2 ± 0.2) and $(-2.2$

$\pm 0.2) \times 10^{-5} \text{ bar}^{-1}$ for amide proton and nitrogen, respectively, for free barnase, and intercepts (-2.2 ± 0.4) and $(10 \pm 2) \times 10^{-10} \text{ ppm bar}^{-2}$, and slopes (-6.7 ± 0.1) and $(-3.2 \pm 0.1) \times 10^{-5} \text{ bar}^{-1}$ for amide proton and nitrogen, respectively, for complexed barnase (Supplementary Fig. 1).

Longitudinal and transverse relaxation was measured using HSQC spectra with 9 interleaved delay points and 3 duplicates (delays 2, 5 and 8) for uncertainty estimation (each applied to itself and neighboring delay points 1-3, 4-6, and 7-9)³⁹. Maximum peak intensities and uncertainties were used to fit single-exponential decay curves with 3 parameters. ^1H - ^{15}N nuclear Overhauser enhancement (NOE) experiments were measured with a 5 s mixing time with and without irradiation of ^1H . Relaxation was measured at 500 and 600 MHz (^1H) for high-pressure experiments, and 500, 600, and 750 MHz (^1H) for ambient pressure experiments. Deuterium relaxation employed Dz and Dy experiments with on-the-fly IzCz-compensation⁴⁰. High pressure NMR relaxation experiments were carried out in a 3 kbar rated 5 mm o.d./3 mm i.d. ceramic NMR tube connected to a high-pressure Xtreme-60 pressure-stat syringe pump (Daedalus Innovations LLC, Aston, PA). The pressure medium was degassed water with a mineral oil interface with the sample. The effect of pressure on imidazole's pK_a is small⁴¹. Relaxation measurements on the complex (24 kDa) were performed on uniformly ^{15}N -labeled barstar and uniformly ^{13}C -labeled barnase grown in 60% D_2O .

The macromolecular rotational correlation model and Lipari-Szabo squared generalized order parameters of the amide N-H bond vectors (O_{NH}^2) were determined from ^{15}N relaxation experiments⁴². Tumbling models for the complex were obtained from ^{15}N -labeled barstar backbone relaxation measurements. Tumbling models were chosen according to the Akaike and Bayesian information criteria and F-tests to obtain probability values for each model⁴³. Simple model-free parameters were determined using a grid-search in a C++/AMP implementation of Relxn2A⁴⁴. The analysis used an effective N-H bond length of $1.04 \text{ \AA}$ ²⁶, a general ^{15}N tensor breadth of -170 p.p.m. ⁴⁵, a quadrupolar coupling constant of 167 kHz ⁴⁶, and a methyl rotation order parameter O_{rot}^2 of 0.1107 assuming perfect tetrahedral geometry of the methyl carbon used to extract the O_{axis}^2 .

Void volume calculations of protein structures

Voronoi volumes⁴⁷ are ideally suited to investigate the volume of buried atoms as “the sum of polyhedral volumes is exactly equal to the total space occupied by the points”⁴⁸. Voronoi volumes were determined with an in-house Cython program. Only structural models with a nominal resolution of < 2.5 Å based on data obtained at cryogenic temperature for the protein with the same biological context as that of the NMR experiment (e.g., free vs. complexed) were considered. Available room temperature crystal structures were analyzed separately. When multiple copies of the protein were present in the asymmetric unit, the copy with the strongest electron density and highest quality model was identified by visual inspection. The calculation finds the edges of the coordinates and defines a box with a 5 Å padding. A cubic grid is created with a step size of 0.01 Å. Each voxel is interrogated for the nearest heavy atom and assigned to it. Voxels within the van der Waals radius⁴⁹ of any atom were excluded. The sum of all voxels assigned to an atom represents the atom’s Voronoi polyhedron and is used to calculate its volume. Side chain volumes were summed starting at the C β and ending with the atoms with the same number of dihedral angles as the methyl group of interest. For example, Ile C γ 2 methyls will include the volume of both C γ carbons and one C β atom, while Ile C δ methyls will include one C δ , both C γ , and one C β atom. The surface was defined using a large probe (2.4 Å radius) to avoid fitting inside any internal cavities. The surface algorithm ignored ligands and waters, so binding pockets were considered open surfaces as well. The probe was moved through the grid to find all voxels where the probe fit without steric overlap, and protein atoms that came within 1 voxel of the probe were flagged as belonging to the surface. If any atom of a side chain (starting at the C β and ignoring backbone) contained a surface atom, the side chain was not considered buried. Alternative rotamers were analyzed by including all rotamers in the calculation, summing the volumes of all rotamers of a given side chain, and subtracting the van der Waals volume only once. All unoccupied void volume is obtained by this calculation including that which remains from perfect packing of spheres.

Data Availability

The barnase relaxation data reported here has been deposited to the BMRB under accession number [@](#).

References

1. Berman H. M., *et al.* The Protein Data Bank *Nucleic Acids Res.* **28**, 235-242. (2000).
2. Frederick K. K., Marlow M. S., Valentine K. G. & Wand A. J. Conformational entropy in molecular recognition by proteins. *Nature* **448**, 325-329 (2007).
3. Marlow M. S., Dogan J., Frederick K. K., Valentine K. G. & Wand A. J. The role of conformational entropy in molecular recognition by calmodulin. *Nat. Chem. Biol.* **6**, 352-358 (2010).
4. Caro J. A., *et al.* Entropy in molecular recognition by proteins. *Proc. Natl Acad. Sci. USA* **114**, 6563-6568 (2017).
5. Wand A. J. & Sharp K. A. Measuring entropy in molecular recognition by proteins. *Annu. Rev. Biophys.* **47**, 41-61 (2018).
6. Zhou H. X. & Gilson M. K. Theory of free energy and entropy in noncovalent binding. *Chem. Rev.* **109**, 4092-4107 (2009).
7. Mittermaier A. & Kay L. E. Review - New tools provide new insights in NMR studies of protein dynamics. *Science* **312**, 224-228 (2006).
8. Sekhar A. & Kay L. E. An NMR view of protein dynamics in health and disease. *Annu. Rev. Biophys.* **48**, 297-319 (2019).
9. Igumenova T. I., Frederick K. K. & Wand A. J. Characterization of the fast dynamics of protein amino acid side chains using NMR relaxation in solution. *Chem. Rev.* **106**, 1672-1699 (2006).
10. Kundrot C. E. & Richards F. M. Crystal-structure of hen egg-white lysozyme at a hydrostatic-pressure of 1000 atmospheres. *J. Mol. Biol.* **193**, 157-170 (1987).
11. Collins M. D., Quillin M. L., Hummer G., Matthews B. W. & Gruner S. M. Structural rigidity of a large cavity-containing protein revealed by high-pressure crystallography. *J. Mol. Biol.* **367**, 752-763 (2007).

12. Yamada H., Nagae T. & Watanabe N. High-pressure protein crystallography of hen egg-white lysozyme. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **71**, 742-753 (2015).
13. Bezsonova I., Korzhnev D. M., Prosser R. S., Forman-Kay J. D. & Kay L. E. Hydration and packing along the folding pathway of SH3 domains by pressure-dependent NMR. *Biochemistry* **45**, 4711-4719 (2006).
14. Korzhnev D. M., *et al.* Probing the transition state ensemble of a protein folding reaction by pressure-dependent NMR relaxation dispersion. *J. Am. Chem. Soc.* **128**, 5262-5269 (2006).
15. Nucci N. V., Fuglestad B., Athanasoula E. A. & Wand A. J. Role of cavities and hydration in the pressure unfolding of T-4 lysozyme. *Proc. Natl Acad. Sci. USA* **111**, 13846-13851 (2014).
16. Fuglestad B., Stetz M. A., Belnavis Z. & Wand A. J. Solution NMR investigation of the response of the lactose repressor core domain dimer to hydrostatic pressure. *Biophys. Chem.* **231**, 39-44 (2017).
17. Roche J., *et al.* Cavities determine the pressure unfolding of proteins. *Proc. Natl Acad. Sci. USA* **109**, 6945-6950 (2012).
18. Fuentes E. J. & Wand A. J. Local stability and dynamics of apocytochrome b(562) examined by the dependence of hydrogen exchange on hydrostatic pressure. *Biochemistry* **37**, 9877-9883 (1998).
19. Hawley S. A. Reversible pressure-temperature denaturation of chymotrypsinogen. *Biochemistry* **10**, 2436-2442 (1971).
20. Lesch H., Hecht C. & Friedrich J. Protein phase diagrams: The physics behind their elliptic shape. *J. Chem. Phys.* **121**, 12671-12675 (2004).
21. Caro J. A. & Wand A. J. Practical aspects of high-pressure NMR spectroscopy and its applications in protein biophysics and structural biology. *Methods* **148**, 67-80 (2018).
22. Schreiber G. & Fersht A. R. Energetics of protein-protein interactions - Analysis of the barnase-barstar interface by single mutations and double mutant cycles. *J. Mol. Biol.* **248**, 478-486 (1995).
23. Lipari G. & Szabo A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546-4559 (1982).

24. Muhandiram D. R., Yamazaki T., Sykes B. D. & Kay L. E. Measurement of ^2H T1 and T1rho relaxation times in uniformly ^{13}C -labeled and fractionally ^2H -labeled proteins in solution. *J. Am. Chem. Soc.* **117**, 11536-11544 (1995).
25. Buckle A. M., Schreiber G. & Fersht A. R. Protein-protein recognition - Crystal structural analysis of barnase barstar complex at 2.0 Å resolution. *Biochemistry* **33**, 8878-8889 (1994).
26. Mittermaier A., Kay L. E. & Forman-Kay J. D. Analysis of deuterium relaxation-derived methyl axis order parameters and correlation with local structure. *J. Biomol. NMR* **13**, 181-185 (1999).
27. Nucci N. V., Pometun M. S. & Wand A. J. Site-resolved measurement of water-protein interactions by solution NMR. *Nat. Struct. Mol. Biol.* **18**, 245-249 (2011).
28. Fenwick R. B., van den Bedem H., Fraser J. S. & Wright P. E. Integrated description of protein dynamics from room-temperature X-ray crystallography and NMR. *Proc. Natl Acad. Sci. USA* **111**, E445-E454 (2014).
29. Bromberg S. & Dill K. A. Side-chain entropy and packing in proteins. *Protein Sci.* **3**, 997-1009 (1994).
30. Doolittle A. K. Studies in Newtonian flow. II. The dependence of the viscosity of liquids on free-space. *J. Appl. Phys.* **22**, 1471-1472 (1951).
31. Kasinath V., Fu Y. N., Sharp K. A. & Wand A. J. A sharp thermal transition of fast aromatic-ring dynamics in ubiquitin. *Angew. Chem. Int. Ed.* **54**, 102-107 (2015).
32. Fu Y. N., Kasinath V., Moorman V. R., Nucci N. V., Hilser V. J. & Wand A. J. Coupled motion in proteins revealed by pressure perturbation. *J. Amer. Chem. Soc.* **134**, 8543-8550 (2012).
33. Bycroft M., Matouschek A., Kellis J. T., Serrano L. & Fersht A. R. Detection and characterization of a folding intermediate in barnase by NMR. *Nature* **346**, 488-490 (1990).
34. Cooper A. & Dryden D. T. F. Allostery without conformational change - A plausible model. *Eur. Biophys. J Biophys. Lett.* **11**, 103-109 (1984).
35. Neri D., Szyperski T., Otting G., Senn H. & Wuthrich K. Stereospecific nuclear magnetic-resonance assignments of the methyl-group of valine and leucine in the DNA-binding domain of the 434-repressor by biosynthetically directional fractional ^{13}C labeling. *Biochemistry* **28**, 7510-7516 (1989).

36. Robson S., Arthanari H., Hyberts S. G. & Wagner G. Nonuniform sampling for NMR spectroscopy. *Methods Enzymol.* **614**, 263-291 (2019).
37. Tonelli L. W. & Markley J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325-1327 (2015).
38. Erlach M. B., Koehler J., Moeser B., Horinek D., Kremer W. & Kalbitzer H. R. Relationship between nonlinear pressure-induced chemical shift changes and thermodynamic parameters. *J. Phys. Chem. B* **118**, 5681-5690 (2014).
39. Stetz M. A., *et al.* Characterization of internal protein dynamics and conformational entropy by NMR relaxation. *Methods Enzymol.* **615**, 237-284 (2019).
40. Millet O., Muhandiram D. R., Skrynnikov N. R. & Kay L. E. Deuterium spin probes of side-chain dynamics in proteins. 1. Measurement of five relaxation rates per deuteron in C-13-labeled and fractionally H-2-enriched proteins in solution. *J. Am. Chem. Soc.* **124**, 6439-6448 (2002).
41. Quinlan R. J. & Reinhart G. D. Baroresistant buffer mixtures for biochemical analyses. *Anal. Biochem.* **341**, 69-76 (2005).
42. Tjandra N., Feller S. E., Pastor R. W. & Bax A. Rotational diffusion anisotropy of human ubiquitin from ¹⁵N NMR relaxation. *J. Am. Chem. Soc.* **117**, 12562-12566 (1995).
43. Moorman V. R., *et al.* Dynamic and thermodynamic response of the Ras protein Cdc42Hs upon association with the effector domain of PAK3. *J. Mol. Biol.* **426**, 3520-3538 (2014).
44. Dellwo M. J. & Wand A. J. Model-independent and model-dependent analysis of the global and internal dynamics of cyclosporine-A. *J. Am. Chem. Soc.* **111**, 4571-4578 (1989).
45. Lee A. L. & Wand A. J. Assessing potential bias in the determination of rotational correlation times of proteins by NMR relaxation. *J. Biomol. NMR* **13**, 101-112 (1999).
46. Mittermaier A. & Kay L. E. Measurement of methyl H-2 quadrupolar couplings in oriented proteins. How uniform is the quadrupolar coupling constant? *J. Am. Chem. Soc.* **121**, 10608-10613 (1999).
47. Voronoi G. New applications of continuous parameters to the theory of quadratic forms. - First memoir - Some properties of perfect positive quadratic forms. *J. Reine Angewandte Mathematik* **133**, 97-178 (1908).

48. Richards F. M. Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**, 151-176 (1977).
49. Bondi A. van der Waals volumes and radii. *J. Phys. Chem.* **68**, 441-451 (1964).

Acknowledgements

We thank Kim Sharp and Ken Dill for helpful discussion. Funding from NIH (R01 GM102447), the University of Pennsylvania and Texas A&M University.

Author contributions

A.J.W. and J.A.C conceived and designed the project. K.G.V. and J.A.C. carried out the relaxation experiments and data analysis. J.A.C. wrote the necessary software and carried out the volumetric analyses reported. J.A.C., K.G.V. and A.J.W. analyzed the results. A.J.W. and J.A.C. wrote the manuscript.

Competing interests

A.J.W. is a founding member of Daedalus Innovations, LLC, a manufacturer of high-pressure NMR apparatus.

Additional information

Supplementary information The online version contains supplementary material available at [@](#)

Correspondence and requests for materials should be addressed to A.J.W.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

Figures

Fig. 1 | Dynamic response of barnase to high pressure and to binding barstar. The locations of reporter methyl groups are indicated by spheres on the ribbon-diagram of barnase (PDB 1B2X & 1B27) and color coded according to the ΔO^2_{axis} where red and blue correspond to decreases and increases in ΔO^2_{axis} upon the indicated change of state, respectively. The inset summarizes the average change of ΔO^2_{axis} values in barnase as a whole and within the two regions indicated by the dashed lines, region I as dashed lines and region II as dotted lines (see text for details).

Values in parentheses are the variance of changes within the indicated group and not error.

Contributions of conformational entropy of the indicated state interconversion were obtained using the dynamical proxy⁴ and included all methyl probes. See Supplementary Table 1 for further details. [2 columns]

Figure 2 | Dependence of fast side chain motion on void volume. **a**, Summary of the database of 50 cryogenic crystal structures and 2,568 O^2_{axis} parameters obtained in solution near room temperature. Each filled circle represents one O^2_{axis} value plotted as a function of void volume calculated from the side chain's corresponding crystal structure. Circles are colored from blue to red (sparse to dense) according to the contour lines. Sources are summarized in Supplementary Table 3. **b**, Distributions of experimentally determined O^2_{axis} used in the analysis (see main text). The distributions for $N\chi = 0$ (Ala βCH_3), $N\chi = 1$ (Thr, Val and Ile γCH_3), $N\chi = 2$ (Leu and Ile δCH_3), and $N\chi = 3$ (Met ϵCH_3) were comprised of 273, 1059, 1111 and 125 values, respectively. Data sources are identified in Supplementary Table 1. [1 column]

Fig. 3| Relationship between surrounding unoccupied volume and fast methyl-bearing side chain motion. **a**, Methyl-bearing side chain O^2_{axis} values correlated with the surrounding void volumes in the cryogenic structure of barnase (1B2X) calculated with (closed circles) and without (open circles) crystallographic water molecules. **b**, Void volumes of methyl groups in the room temperature structure of barnase (1A2P) with (closed squares) and without (open squares) the

alternative conformation of Ile-96. **c**, The changes in void volume resulting from incorporation of a multi-conformational model of the room temperature structure of dihydrofolate reductase (4PST). See text. [1 column]

Supplementary Information

Structural origins of protein conformational entropy

José A. Caro^{1,2}, Kathleen G. Valentine², A. Joshua Wand^{1,2*}

¹Department of Biochemistry & Biophysics, College of Agriculture and Life Sciences, Texas

A&M University, College Station, Texas 77845

²Department of Biochemistry & Biophysics, Perelman School of Medicine, University of

Pennsylvania, Philadelphia, Pennsylvania 19104

* email: wand@tamu.edu

Supplementary Data Table 1: Fast internal motion in barnase^a

Barnase ^b	P (bar)	$\langle O_{axis}^2 \rangle$	$var O_{axis}^2$ ^d	n _{NH} ^e	$\langle O_{NH}^2 \rangle$	$var O_{NH}^2$	τ_m (ns)
Free (all) ^c	1	0.653	0.046	96	0.851	0.002	4.78 ^h
Free (Region I) ^f	1	0.721	0.031				
Free (Region II) ^g	1	0.621	0.030				
Free (all)	3,000	0.647	0.054	92	0.866	0.001	5.94 ⁱ
Free (Region I)	3,000	0.762	0.029				
Free (Region II)	3,000	0.568	0.039				
Complex (all)	1	0.693	0.055	97	0.847	0.001	8.89 ^h
Complex (Region I)	1	0.750	0.030				
Complex (Region II)	1	0.667	0.040				
Complex (all)	3,000	0.649	0.051	90	0.899	0.001	11.48 ⁱ
Complex (Region I)	3,000	0.725	0.023				
Complex (Region II)	3,000	0.589	0.048				

^a all data obtained at 308 K

^b number of probes spectrally resolved in all four states are 33, 16, and 12 for all, Region I, and Region II, respectively.

^c includes all spectrally resolved methyl probes in barnase

^d variance is shown to indicate the variation of the dynamics within each region

^e only amide ¹⁵N-¹H probes in barnase were used to characterize tumbling and backbone motion

^f Region I includes the buried side chains of Val-10, Leu-14, Leu-42, and Ile-51 and the surface exposed side chains of Val-3, Ile-4, Thr-6, Leu-20, Thr-26, Ile-25, Leu-33, Val-45, and Thr-79

^g Region II includes the buried side chains of Leu-63, Ile-76, Ile-88, Leu-89, Ile-96, and the surface exposed side chains of Leu-95, Thr-99, Thr-100, Thr-107, and Ile-109

^h Fully anisotropic tumbling model, effective τ_m reported.

ⁱ Axially symmetric tumbling model, effective τ_m reported.

Supplementary Table 2: Changes in fast internal motion in barnase^a

		$\langle \Delta O_{axis}^2 \rangle (var \Delta O_{axis}^2)$				
		Barnase		Barnase-Barstar		
		1 Bar	3,000 Bar	1 Bar	3,000 Bar	
$\langle \Delta O_{NH}^2 \rangle (var \Delta O_{NH}^2)$	Barnase	1 Bar	-	-0.006 (0.004)	+0.039 (0.006)	-0.004 (0.011)
		3,000 Bar	-0.014 (0.001)	-	+0.045 (0.011)	+0.002 (0.012)
	Barnase-Barstar	1 Bar	+0.009 (0.002)	+0.023 (0.001)	-	-0.042 (0.004)
		3,000 Bar	-0.043 (0.002)	-0.027 (0.001)	-0.052 (0.001)	-

^a Values correspond to the change in order parameter calculated as the column state minus the row state.

^b Upper diagonal of table corresponds to ΔO_{axis}^2 and lower diagonal to ΔO_{NH}^2 .

Supplementary Table 3: Summary of database used in analysis of surrounding unoccupied volume and local fast side chain motion.

Protein	Ligand	PDB	BMRB	Ref
Ubiquitin	-	1UBQ	-	50
HIV-1 Protease	DMP323	1QBS	-	51
Oxidized Flavodoxin	-	1FLV	-	52
SAP (SLAM-associated protein) SH2	-	1D1Z	-	53
SAP (SLAM-associated protein) SH2	Y281 peptide	1D4T	-	53
SAP (SLAM-associated protein) SH2	pY281 peptide	1D4W	-	53
Protein L, B1 immunoglobulin binding domain	-	1HZ6	-	54
Tenascin; third fnIII	-	1TEN	-	55
Fibronectin; tenth fnIII	-	1FNF	-	55
Dihydrofolate reductase	Folate	1RX7	-	56
Dihydrofolate reductase	Folate, NADP	4NX7	-	56
Dihydrofolate reductase	NADPH	1RA1	-	57
Dihydrofolate reductase	NADPH, methotrexate	1RX3	-	57
Malate Synthase G	Pyruvate, acetyl-CoA	1P7T	-	58
Calmodulin	Ca ²⁺	1UP5	15188	2
Calmodulin	Ca ²⁺ , smMLCK peptide	1CDL	15186	2
Calmodulin	Ca ²⁺ , CaMKI peptide	1MXE	15184	2
Calmodulin	Ca ²⁺ , eNOS peptide	1NIW	15185	2
Calmodulin	Ca ²⁺ , nNOS peptide	2O60	15191	2
PhoB, DNA binding domain	-	1GXQ	-	59
PhoB, DNA binding domain	DNA	1GXP	-	59
Chymotrypsin Inhibitor 2	-	2CI2	-	60
SAP90, PDZ3 domain	-	1BFE	-	61
SAP90, PDZ3 domain	CRIP1	1BE9	-	61
Galectin-3, carbohydrate domain	lactose	2NN8	-	62
Galectin-3, carbohydrate domain	L2	2XG3	-	62
Galectin-3, carbohydrate domain	L3	1KJR	-	62
Galectin-3, carbohydrate domain	S (carbohydrate)	6QGE	27721	63
Galectin-3, carbohydrate domain	R (carbohydrate)	6QGF	27722	63
Hen egg white lysozyme	-	1LZA	18304	64
Hen egg white lysozyme	chitotriose	1LZB	18305	64
Thymidylate synthase	-	2FTQ	-	65
Thymidylate synthase	dUMP	1BID	-	65
Histamine Binding Protein	-	3GAQ	25728	4
Histamine Binding Protein	histamine	3G7X	25727	4
Barnase	-	1B2X	26619	4
Barnase	DNA	1BRN	26620	4
Adhesin-competence repressor	Zn(II)	3TGN	27447	66
CzrA	Zn(II)	6CDA	27459	67
Catabolyte activator protein	2xcAMP	2GZW		68
Catabolyte activator protein	2xcAMP, DNA	1ZRF		68
Barnase	Barstar	1B27		this work

Supplementary Fig 1 a, Dependence of free barnase ^1H - ^{15}N HSQC resonance peak volumes on pressure. **b**, Correlation between first- and second- order terms of the polynomial used to fit ^1H chemical shifts as a function of pressure. The line is a linear fit to the error-weighted data points and has an intercept of $(6.5 \pm 0.8) \times 10^{-10}$ ppm bar $^{-2}$ and a slope of $(-7.2 \pm 0.2) \times 10^{-5}$ bar $^{-1}$.

Additional References

50. Lee A. L., Flynn P. F. & Wand A. J. Comparison of ^2H and ^{13}C NMR relaxation techniques for the study of protein methyl group dynamics in solution. *J. Am. Chem. Soc.* **121**, 2891-2902 (1999).
51. Ishima R., Petkova A. P., Louis J. M. & Torchia D. A. Comparison of methyl rotation axis order parameters derived from model-free analyses of ^2H and ^{13}C longitudinal and transverse relaxation rates measured in the same protein sample. *J. Am. Chem. Soc.* **123**, 6164-6171 (2001).
52. Liu W., Flynn P. F., Fuentes E. J., Kranz J. K., McCormick M. & Wand A. J. Main chain and side chain dynamics of oxidized flavodoxin from *Cyanobacterium anabaena*. *Biochemistry* **40**, 14744-14753 (2001).
53. Finerty P. J., Muhandiram R. & Forman-Kay J. D. Side-chain dynamics of the SAP SH2 domain correlate with a binding hot spot and a region with conformational plasticity. *J. Mol. Biol.* **322**, 605-620 (2002).

54. Millet O., Mittermaier A., Baker D. & Kay L. E. The effects of mutations on motions of side-chains in protein L studied by 2H NMR dynamics and scalar couplings. *J. Mol. Biol.* **329**, 551-563 (2003).
55. Best R. B., Rutherford T. J., Freund S. M. & Clarke J. Hydrophobic core fluidity of homologous protein domains: relation of side-chain dynamics to core composition and packing. *Biochemistry* **43**, 1145-1155 (2004).
56. Schnell J. R., Dyson H. J. & Wright P. E. Effect of cofactor binding and loop conformation on side chain methyl dynamics in dihydrofolate reductase. *Biochemistry* **43**, 374-383 (2004).
57. Mauldin R. V., Carroll M. J. & Lee A. L. Dynamic dysfunction in dihydrofolate reductase results from antifolate drug binding: Modulation of dynamics within a structural state. *Structure* **17**, 386-394 (2009).
58. Tugarinov V. & Kay L. E. Quantitative ¹³C and ²H NMR relaxation studies of the 723-residue enzyme malate synthase G reveal a dynamic binding interface. *Biochemistry* **44**, 15970-15977 (2005).
59. Okamura H., Makino K. & Nishimura Y. NMR dynamics distinguish between hard and soft hydrophobic cores in the DNA-binding domain of PhoB and demonstrate different roles of the cores in binding to DNA. *J. Mol. Biol.* **367**, 1093-1117 (2007).
60. Whitley M. J., Zhang J. & Lee A. L. Hydrophobic core mutations in CI2 globally perturb fast side-chain dynamics similarly without regard to position. *Biochemistry* **47**, 8566-8576 (2008).
61. Petit C. M., Zhang J., Sapienza P. J., Fuentes E. J. & Lee A. L. Hidden dynamic allostery in a PDZ domain. *Proc. Natl. Acad. Sci. USA* **106**, 18249 (2009).
62. Diehl C., *et al.* Protein flexibility and conformational entropy in ligand design targeting the carbohydrate recognition domain of galectin-3. *J. Am. Chem. Soc.* **132**, 14577-14589 (2010).
63. Verteramo M. L., *et al.* Interplay between conformational entropy and solvation entropy in protein-ligand binding. *J. Am. Chem. Soc.* **141**, 2012-2026 (2019).
64. Moorman V. R., Valentine K. G. & Wand A. J. The dynamical response of hen egg white lysozyme to the binding of a carbohydrate ligand. *Protein Sci.* **21**, 1066-1073 (2012).

65. Sapienza P. J. & Lee A. L. Widespread perturbation of function, structure, and dynamics by a conservative single-atom substitution in thymidylate synthase. *Biochemistry* **55**, 5702-5713 (2016).
66. Capdevila D. A., Huerta F., Edmonds K. A., Le M. T., Wu H. & Giedroc D. P. Tuning site-specific dynamics to drive allosteric activation in a pneumococcal zinc uptake regulator. *eLife* **7**, e37268 (2018).
67. Capdevila D. A., Edmonds K. A., Campanello G. C., Wu H., Gonzalez-Gutierrez G. & Giedroc D. P. Functional role of solvent entropy and conformational entropy of metal binding in a dynamically driven allosteric system. *J. Am. Chem. Soc.* **140**, 9108-9119 (2018).
68. Tzeng S.-R. & Kalodimos C. G. Protein activity regulation by conformational entropy. *Nature* **488**, 236-240 (2012).