

Covariation of backbone motion throughout a small protein domain

Kristen L Mayer¹, Matthew R Earley¹, Sonia Gupta¹, Kumar Pichumani¹, Lynne Regan² & Martin J Stone¹

The synchronization (correlation) of conformational fluctuations in folded proteins may influence the rates of enzyme catalysis and ligand binding as well as the stabilities of native proteins and their complexes. However, experimental detection of correlated motions remains difficult. Herein, we present an analysis of the covariation of NMR-derived backbone dynamical parameters among a family of ten mutants of a small protein. Both the spatial restriction and the time scales of backbone motions exhibit a higher degree of covariation than would be expected if the internal motions of each group were independent, providing experimental support for correlated dynamics. Application of this approach to other proteins may reveal dynamical correlations that influence catalysis, ligand-binding and/or protein stability.

Protein structures can fluctuate on time scales ranging from femtoseconds to weeks. The rates of enzyme catalysis, ligand binding and solvent hydrogen exchange may be limited by the kinetics of protein conformational rearrangements on time scales corresponding to these processes (microseconds to hours). Much faster (subnanosecond) dynamics can influence enzyme-catalyzed quantum tunneling by modulating tunneling distances and/or pathways^{1,2}. In addition, synchronization (correlation) of fast internal motions may lead to the occasional concerted motions necessary for conventional enzymes to overcome activation barriers, hence influencing the kinetics of ligand binding or catalysis³. For example, correlated motions of residues distant from the active site of dihydrofolate reductase are proposed to have a profound effect on the rate of catalysis^{3,4}. Similarly, acrylamide appears to noncompetitively inhibit phosphorylase *b* by quenching the fast internal motions of the enzyme⁵. Thus, microscopic motions on very fast time scales (<1 ns) can potentially regulate macroscopic chemical phenomena with much longer apparent time scales (>1 μ s)^{6,7}.

Protein dynamics also influence the thermodynamics of protein chemistry. Fluctuations on the subnanosecond time scale involve interconversion of many conformational substates. Thus, modification of these dynamics may result in substantial changes in conformational entropy (and possibly also enthalpy⁶), with obvious consequences for the thermodynamics of folding and ligand binding, including allosteric effects. Synchronization of the internal motions would reduce the residual conformational entropy of a protein. Thus, the correlated motions of residues in, for example, an antibody combining site⁸ may minimize the entropic losses upon antigen binding, resulting in higher affinity.

Although both kinetic and thermodynamic considerations point to the importance of fast correlated motions in proteins, the experimental

observation of correlated motions remains a considerable challenge. The subnanosecond rotational motions of individual bond vectors in proteins can be characterized based on NMR relaxation data, providing measures for both the degree of restriction (order parameter, S^2) and the time scale (internal correlation time, τ_c) of such motion^{9–11}. Comparisons of these motions for free and ligated forms of proteins have revealed both decreases and increases in flexibility upon binding¹². In addition, these parameters have been correlated with enzymatic activity⁴, binding affinity¹³ and binding cooperativity¹⁴ and provide a plausible molecular description for the observed glass transition of protein structures¹⁵. However, approaches to relate NMR-derived order parameters to conformational entropy^{16–18} suffer from several limitations, including an inability to account for correlated subnanosecond motions of different bond vectors¹².

Herein we propose a general approach to the detection of correlated changes in internal protein motions, which are expected to reflect the underlying influence of correlated dynamics. If the motions of two bond vectors are coupled, then a perturbation of the protein that modifies the dynamics of one vector is likely to influence the coupled vector in a similar manner. Correlated changes in motional amplitudes and/or timescales may be detected by covariation of the S^2 and/or τ_c values for a series of protein perturbations. The perturbations may be chemical (such as mutations or different ligands) or physical (such as temperature or pressure). They must be sufficiently severe that they result in measurable dynamical changes but must be sufficiently mild that they do not disrupt existing networks of correlated motions. As a demonstration of this approach, we have analyzed the covariation of backbone N-H dynamics data for ten mutants of the immunoglobulin (IgG)-binding domain (B1 domain) of *Streptococcal* protein G (ref. 19; Fig. 1).

¹Department of Chemistry, Indiana University, Bloomington, Indiana 47405-0001, USA. ²Department of Chemistry and Department of Molecular Biophysics and Biochemistry, Yale University, PO Box 208114, 266 Whitney Avenue, New Haven, Connecticut 06520-8114, USA. Correspondence should be addressed to M.J.S. (mastone@indiana.edu).



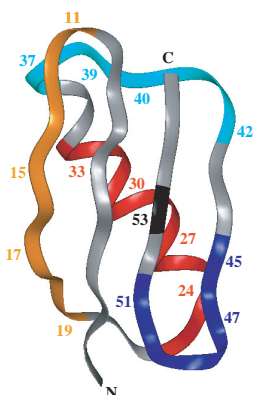


Figure 1 Ribbon representation of the B1 domain of protein G. Regions with the highest prevalence of correlated dynamics (see text) are colored: β 2-strand, orange; α -helix, red; extended loop, cyan; and β 3- β 4 turn, blue. The guest position (residue 53) is black.

RESULTS

Relaxation and dynamics parameters

^{15}N longitudinal (R_1) and transverse (R_2) auto-relaxation rates, transverse cross-relaxation rates (η_{xy}) and heteronuclear ^1H - ^{15}N NOEs were measured at 30 °C for all resolved NH resonances in mutants of the B1 domain (the background construct is triple mutant T2Q I6A T44A of the wild-type B1 domain) substituted with each of ten amino acids (alanine, aspartate, histidine, isoleucine, leucine, methionine, serine, threonine, tryptophan and valine) at position 53, on the solvent-exposed surface of the β -sheet²⁰. Surface mutations were used because they are likely to be less disruptive than core mutations to the networks of correlated motion. The stabilities of the ten mutants vary over a range of 2.0 kcal mol⁻¹ (all are >99% folded under the experimental conditions) and differences in backbone dynamics are

detectable^{20,21}. Data were analyzed according to the Lipari-Szabo model-free formalism using the same rotational diffusion tensor for all mutants and the same mathematical model for a particular residue in all mutants. The covariation analyses were carried out using dynamics parameters for all residues whose relaxation data could be adequately fit using a single time scale of internal motions faster than the rotational correlation time (3.43 ns).

Covariation analysis

The covariation of motions between each pair of N-H bond vectors was probed by correlating the S^2 or τ_e values of the two vectors for all the available mutant data. A number of strong correlations were observed (Fig. 2a–d), although many poor correlations and some negative correlations were also found (Fig. 2e,f). The normalized frequency of good correlations (high r^2 values) was substantially higher than randomly expected for both S^2 and τ_e data (Fig. 2g,h). Moreover, in the absence of correlated motions, one would expect distributions of both Pearson correlation coefficients (r) and slopes to be distributed approximately evenly around zero. The observed distributions of both parameters are clearly shifted toward positive values (Fig. 2i–l), providing further evidence for the existence of nonrandom correlations and indicating that increased motional amplitudes or timescales at one position are most often associated with increased (rather than decreased) amplitudes or time scales, respectively, at the correlated position. These data provide support for the proposal that correlated backbone motions occur in the B1 domain.

Structural distribution of dynamical covariation

The structural distribution of the correlated dynamical parameters is represented by the two-dimensional color map in Figure 3. In theory,

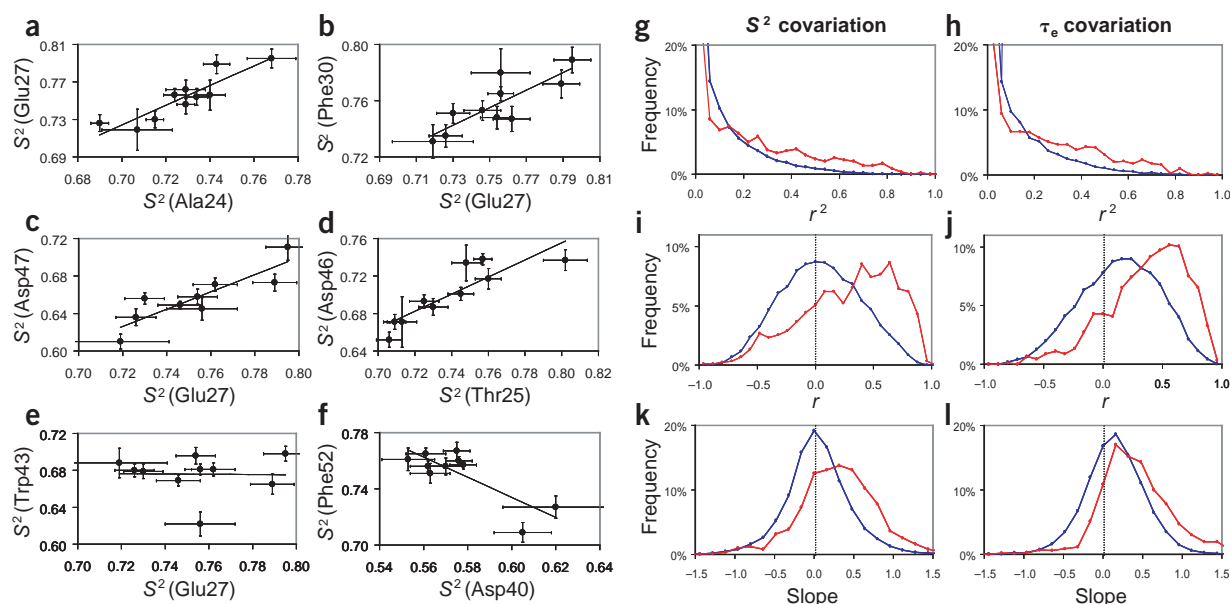


Figure 2 Covariation of dynamics parameters for the B1 domain. (a–d) Correlations between the order parameters of several NH groups in the α -helix (Ala24, Thr25, Glu27 and Phe30) and the β 3- β 4 hairpin turn (Asp46 and Asp47). (e) An example of a poor order parameter correlation (between Glu27 in the α -helix and Trp43 in the β 3-strand). (f) An example of a negative order parameter correlation (between Asp40 in the α - β 3 loop and Phe52 adjacent to the mutation site in the β 4 strand). (g–l) Normalized distributions of squared Pearson correlation coefficients (r^2), raw Pearson correlation coefficients (r) and slopes for covariation analyses of order parameters (g, i, k) and effective internal correlation times (h, j, l). Observed distributions are red, whereas randomly expected distributions (in the absence of correlated motions) are blue. The randomly expected distributions of slopes and r values should be centered close to zero if the discrete distributions of dynamical parameters are very similar for all mutants (as observed here), but will be shifted from zero if there are substantial differences between these distributions (as would occur for variable temperature data).

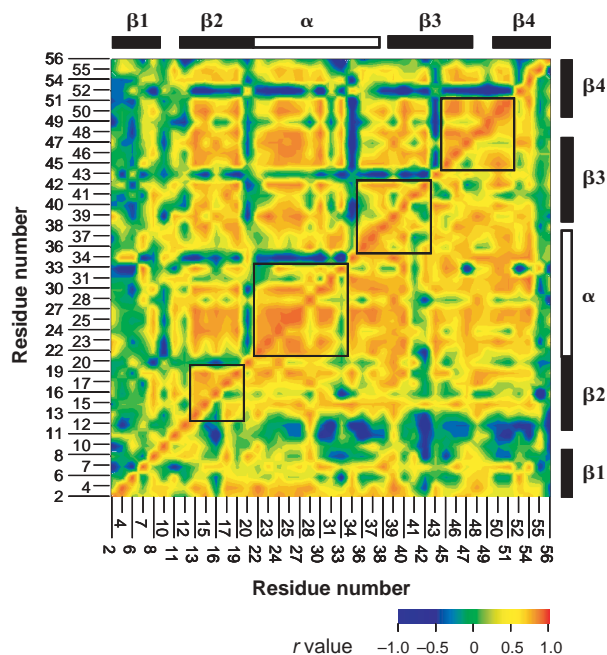


Figure 3 Structural distribution of dynamical covariations. Color map showing the strengths of the observed correlations (r values) for each pair of residues. Data for S^2 covariations are above the diagonal whereas those for τ_e correlations are below the diagonal. Black boxes indicate the positions of the structural elements for which correlated changes in dynamics are most prevalent: the second β -strand (residues 11–19), the first ~ 3 turns of the α -helix (residues 22–33), the extended loop (residues 36–42) and the $\beta 3$ – $\beta 4$ hairpin turn (residues 45–51). The positions of secondary structure elements are indicated at the top and on the right as closed or open bars (β -strands and α -helix, respectively).

true synchronous motions could lead to either positive or negative covariation of order parameters but would always yield a positive covariation of time scales. **Figure 3** indicates that there is extensive positive covariation of both S^2 and τ_e values within the second β -strand, the first ~ 3 turns of the α -helix, the extended loop and the $\beta 3$ – $\beta 4$ hairpin turn, and also from each of these regions to the others. These correlations can be rationalized by the spatial proximity of amino acids within these regions (**Fig. 1**): Asn37 and Val39 (extended loop) pack against Thr11 and Leu12 ($\beta 2$ -strand); Thr16 and Thr18 ($\beta 2$ -strand) pack against Val29, Phe30 and Tyr33 (α -helix); Ala23 and Glu27 (α -helix) interact with Tyr45 ($\beta 3$ – $\beta 4$ hairpin); and Lys31 (α -helix) closely approaches Asp40 (extended loop). It is noteworthy that one of the highly correlated regions (the $\beta 3$ – $\beta 4$ hairpin) is the folding nucleus of the B1 domain, which retains its hairpin structure as an isolated peptide^{22,23}. Notably, few strong correlations are observed for the β -strand residues surrounding the variable amino acid (**Fig. 1**), suggesting that any correlated motions in these regions may have been disrupted by the mutations. Taken together, the B1 domain data are consistent with a long-range network of correlated motions in which the dynamics of residues on opposite sides of the protein are sensitive to each other by virtue of a series of intervening noncovalent interactions. This study does not address whether the dynamic network is involved in the IgG-binding function of the B1 domain. However, the present observations add credence to the proposal that dynamic networks may exist in other proteins and could influence enzymatic activity^{3–5}.

DISCUSSION

Current methods to interpret NMR-derived order parameters in terms of changes in conformational entropy rely on the assumption that different bond vectors move independently of each other^{12,14}. If the motions of several groups are coupled, these methods will overestimate the residual entropy of the system. Thus, there has been interest in determining whether the internal motions of proteins are strongly coupled and, if so, how markedly the motional coupling affects the conformational entropy. Recent simulations by Lee *et al.*²⁴ indicate that relatively strong motional coupling within a cluster of four side chains leads to moderate ($\sim 40\%$) reductions in residual entropy. The present data indicate that motional covariations in the backbone of the B1 domain are above randomly expected levels, but the majority of pairwise correlations are still very weak (**Fig. 2g–j**); thus, $r^2 < 0.25$ for 64% and 58% of S^2 and τ_e correlations, respectively, and $r^2 < 0.50$ for 84% and 85% of S^2 and τ_e correlations, respectively. Taken together, these results suggest that the conformational entropy values derived from NMR order parameters are likely to be only slightly overestimated by assuming independent motions. Furthermore, if the extent of motional correlations remains the same in the two states being compared, some cancellation of systematic errors arising from motional correlation is to be expected when calculating the change in conformational entropy from order parameters. Assuming independent motions, the backbone entropy estimated from NH order parameters for the ten B1 domain mutants varies by more than 4 kcal mol⁻¹, more than the variation in global stability^{20,21}, but is poorly correlated with folding free energies (data not shown). Thus, the variations of backbone entropy in the folded state appear to make a substantial, but not dominant, contribution to B1 domain stability.

In summary, the data presented herein have begun to reveal the presence and nature of correlated motions throughout a small protein domain. The existence of correlated motions within a single amino acid residue or between neighboring residues can be inferred from comparison of dynamics data for different internuclear vectors in the same amino acid^{25,26}, from cross-correlated NMR relaxation experiments¹¹, from statistical variations of experimental dynamics data²⁷ and from anisotropic motional models derived using relaxation data for two bond vectors in the same peptide plane²⁸. The approach used herein reveals both short- and long-range dynamical correlations throughout a protein domain. Although this approach requires extensive measurements for many mutants or physical states of the same protein, it is experimentally and conceptually straightforward, and can be applied to studying the correlated motions of side chain groups and other backbone groups that are sensitive to different dynamic modes²⁸. Thus, application of this method to additional bond vectors and additional proteins, and comparison with molecular dynamics simulations²⁹ will yield a more detailed understanding of correlated protein motions.

METHODS

Sample conditions, experimental methods and data analysis methods were as described for the wild-type B1 domain³⁰, unless noted. The rotational diffusion tensor was initially determined independently for each mutant, using the energy-minimized average coordinates of the wild-type B1 domain (PDB entry 2GB1)³¹. Nine of the ten mutants were best described by prolate axially symmetric diffusion tensors, whereas mutant His53 was slightly more consistent with an isotropic diffusion model. Subsequent model selection and model-free analyses were carried out using the consensus prolate axially symmetric diffusion tensor with best agreement to the appropriate R_2/R_1 ratios for all ten mutants ($\tau_m = (2D_{||} + 4D_{\perp})^{-1} = 3.43$ ns; $2D_{zz} / (D_{xx} + D_{yy}) = 1.29$; $\theta = 86.3^\circ$ and $\phi = 198.8^\circ$; angles are relative to the 2GB1 coordinates). A possible source of

error is that the average structures of particular mutants could differ from the 2GB1 coordinates and could even lie outside the higher-resolution ensemble of NMR structures (3GB1)³². Initial dynamics model selection was carried out independently for each residue in each mutant, using the five standard models³³: model 1, S^2 only; model 2, S^2 and τ_c (<400 ps); model 3, S^2 and exchange broadening term R_{ex} ; model 4, S^2 , τ_c (<400 ps), and R_{ex} ; and model 5, two time scales of internal motion faster than τ_m . Final calculations were then made for each residue using the simplest model consistent with the data for that residue in all ten mutants; models 2, 4 and 5 were selected for 4, 39 and 12 residues, respectively. Alternative data analyses using independently optimized diffusion tensors for each mutant, independently selected dynamics models for the same residue in the different mutants, or model 2 for all residues in all mutants gave qualitatively similar results to those presented here. Residues fit to model 5 data were excluded from the covariation analyses owing to the low reliability of this model. Covariations were analyzed for the remaining 43 residues. Randomly expected distributions were obtained from 10,000 correlations between sets of S^2 or τ_c values (one value for each mutant) in which the individual values were randomly generated using the discrete probability distribution observed for the relevant parameter for that mutant.

ACKNOWLEDGMENTS

We thank J.L. Vaughn for discussions. This work was supported by grants from the US National Science Foundation and the American Chemical Society Petroleum Research Fund.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 21 April; accepted 25 August 2003

Published online at <http://www.nature.com/naturestructuralbiology/>

- Knapp, M.J., Rickert, K. & Klinman, J.P. Temperature-dependent isotope effects in soybean lipoxygenase-1: correlating hydrogen tunneling with protein dynamics. *J. Am. Chem. Soc.* **124**, 3865–3874 (2002).
- Balabin, I.A. & Onuchic, J.N. Dynamically controlled protein tunneling paths in photosynthetic reaction centers. *Science* **290**, 114–117 (2000).
- Bruice, T.C. & Benkovic, S.J. Chemical basis for enzyme catalysis. *Biochemistry* **39**, 6267–6274 (2000).
- Osborne, M.J., Schnell, J., Benkovic, S.J., Dyson, H.J. & Wright, P.E. Backbone dynamics in dihydrofolate reductase complexes: role of loop flexibility in the catalytic mechanism. *Biochemistry* **40**, 9846–9859 (2001).
- Matko, J. *et al.* Correlation between activity and dynamics of the protein matrix of phosphorylase b. *Biochemistry* **19**, 5782–5786 (1980).
- Wand, A.J. Dynamic activation of protein function: a view emerging from NMR spectroscopy. *Nat. Struct. Biol.* **8**, 926–931 (2001).
- Wand, A.J. On the dynamic origins of allosteric activation. *Science* **293**, 1395 (2001).
- Viswanathan, M., Linthicum, D.S. & Subramaniam, S. Analysis of correlated motion in antibody combining sites from molecular dynamics simulations. *Methods* **20**, 362–371 (2000).
- Palmer, A.G. III. NMR probes of molecular dynamics: overview and comparison with other techniques. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 129–155 (2001).
- Kay, L.E. Protein dynamics from NMR. *Nat. Struct. Biol. Suppl.* **5**, 513–517 (1998).
- Fischer, M.W.F., Majumdar, A. & Zuiderweg, E.R.P. Protein NMR relaxation: theory, applications and outlook. *Prog. Nucl. Magn. Reson. Spectrosc.* **33**, 207–272 (1998).
- Stone, M.J. NMR relaxation studies of the role of conformational entropy in protein stability and ligand binding. *Acc. Chem. Res.* **34**, 379–388 (2001).
- Kay, L.E., Muhandiram, D.R., Wolf, G., Shoelson, S.E. & Forman-Kay, J.D. Correlation between binding and dynamics at SH2 domain interfaces. *Nat. Struct. Biol.* **5**, 156–163 (1998).
- Maler, L., Blankenship, J., Rance, M. & Chazin, W.J. Site-site communication in the EF-hand Ca^{2+} -binding protein calbindin D9k. *Nat. Struct. Biol.* **7**, 245–250 (2000).
- Lee, A.L. & Wand, A.J. Microscopic origins of entropy, heat capacity and the glass transition in proteins. *Nature* **411**, 501–504 (2001).
- Li, Z., Raychaudhuri, S. & Wand, A.J. Insights into the local residual entropy of proteins provided by NMR relaxation. *Protein Sci.* **5**, 2647–2650 (1996).
- Akke, M., Bruschweiler, R. & Palmer, A.G. III. NMR order parameters and free energy: an analytical approach and its application to cooperative Ca^{2+} binding by calbindin D9k. *J. Am. Chem. Soc.* **115**, 9832–9833 (1993).
- Yang, D.W. & Kay, L.E. Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: application to protein folding. *J. Mol. Biol.* **263**, 369–382 (1996).
- Boyle, M. *Bacterial Immunoglobulin-Binding Proteins* (Academic, San Diego, California, USA, 1990).
- Smith, C.K., Withka, J.M. & Regan, L. A thermodynamic scale for the β -sheet forming tendencies of the amino acids. *Biochemistry* **33**, 5510–5517 (1994).
- Stone, M.J., Gupta, S., Snyder, N. & Regan, L. Comparison of protein backbone entropy and β -sheet stability: NMR-derived dynamics of protein G B1 domain mutants. *J. Am. Chem. Soc.* **123**, 185–186 (2001).
- Sheinerman, F.B. & Brooks, C.L. Calculations on folding of segment B1 of *Streptococcal* protein G. *J. Mol. Biol.* **278**, 439–456 (1998).
- Blanco, F.J. & Serrano, L. Folding of protein G B1 domain studied by the conformational characterization of fragments comprising its secondary structure elements. *Eur. J. Biochem.* **230**, 634–649 (1998).
- Lee, A.L., Sharp, K.A., Kranz, J.K., Song, X.-J. & Wand, A.J. Temperature dependence of the internal dynamics of a calmodulin-peptide complex. *Biochemistry* **41**, 13814–13825 (2002).
- Daragan, V.A. & Mayo, K.H. A simple approach to analyzing protein side-chain dynamics from ^{13}C NMR relaxation data. *J. Magn. Reson.* **130**, 329–334 (1998).
- Lemaster, D.M. NMR relaxation order parameter analysis of the dynamics of protein side chains. *J. Am. Chem. Soc.* **121**, 1726–1742 (1999).
- Goodman, J.L., Pagel, M.D. & Stone, M.J. Relationships between protein structure and dynamics from a database of NMR-derived backbone order parameters. *J. Mol. Biol.* **295**, 963–978 (2000).
- Fischer, M.W., Zeng, L., Majumdar, A. & Zuiderweg, E.R. Characterizing semilocal motions in proteins by NMR relaxation studies. *Proc. Natl. Acad. Sci. USA* **95**, 8016–8019 (1998).
- Prompers, J.J. & Bruschweiler, R. General framework for studying the dynamics of folded and nonfolded proteins by NMR relaxation spectroscopy and MD simulation. *J. Am. Chem. Soc.* **124**, 4522–4534 (2002).
- Seewald, M.J. *et al.* The role of backbone conformational heat capacity in protein stability: temperature dependent dynamics of the B1 domain of *Streptococcal* protein G. *Protein Sci.* **9**, 1177–1193 (2000).
- Gronenborn, A.M. *et al.* A novel, highly stable fold of the immunoglobulin binding domain of *Streptococcal* protein G. *Science* **253**, 657–661 (1991).
- Kuszewski, J., Gronenborn, A.M. & Clore, G.M. Improving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration. *J. Am. Chem. Soc.* **121**, 2337–2338 (1999).
- Clore, G.M., Driscoll, P.C., Wingfield, P.T. & Gronenborn, A.M. Analysis of the backbone dynamics of interleukin-1 β using two-dimensional inverse detected heteronuclear ^{15}N - 1H NMR spectroscopy. *Biochemistry* **29**, 7387–7401 (1990).